

# To-do list

- Process the data
- Write the backtesting framework and run a simple strategy
- Output the PnL Curve and performance
  - Annual Return
  - Annual Vol
  - Max DD
  - Sharpe ratio
  - Sortino ratio

## Process the data

Use double-index function "unstack" in pandas to get two tables: last price & volume.

Dealing with NA value.

## Strategy Backtesting

Due to the limited time, I choose a simple momentum strategy:

Denote  $\mu$  as the expected return,  $\sigma$  as the volatility and  $A$  is the risk aversion, we use

$$\text{Mom} = \mu - A\sigma$$

as the momentum signal, choose the top  $n = 10$  stocks and invest them with equal weights. A fixed lookback window length and a rebalance period are predetermined.

There are many kinds of trading cost: comission, bid-ask spread, tax, market impact, and intra-day volatility... But I assume a market with no trading cost just for simplicity, and a possible improvement is to add a fixed trading cost in proportion to the turnover.

## Output and Analysis

### Parameters

Initial budget = 10e6 dollars

Lookback window length = 120 trading days

Risk aversion = 1

Rebalance every 20 trading days

## Performance

```
Annualized Return: 11.832964562928803 %  
Annualized Volatility: 23.480414915911798 %  
Downside Deviation: 1.194872015854983 %  
Max Drawdown(in percentage): 43.70012148554696 %  
Max Drawdown(in dollars): 1055904.9004372242  
Sharpe Ratio: 0.503950403146839  
Sortino Ratio: 9.90312301728968
```

```
Process finished with exit code 0
```

## PnL Curve



## Analysis

The strategy seems to work well in comparison with risk-free assets in general, especially in the aspect of downside risk. However, if the data set is from US market, the performance is just ordinary, which did not out perform Dow Jones Index significantly.

Consider that I assume there is no transaction cost and stocks have perfect divisibility, the practical result would be worst than this.

As for the potential improvement, I could try to remove these two assumptions "no transaction cost" and "perfect divisibility" firstly. If the time is enough, I'd like to make use of the "volume" data. And I believe if there are more available data (like OHLC, intra-day, financial report data..), I could write a better strategy.