**Momentum Investing Project**

Weijie Cao, Shirley Duan, Haoyang Shang, Regan Zhu

1. **Introduction**

Momentum investing is a well-known trading strategy that suggests investors first compare and rank the historical performances among various stocks, then purchase the past winner stocks and short the past loser stocks to maximize their profits. In 1993, Jegadeesh and Titman(1993) used this strategy and generated significant positive returns over 3- to 12-month holding periods. In our project, we will use Python to replicate this system and backtest it to see whether momentum investing is a robust trading strategy or not.

To replicate this system, we need to select our “tracking universe.” There are numerous stocks in the financial market; therefore, to improve the efficiency of our project, we applied several filters on S&P 500 stocks to obtain 100 stocks as candidates for our momentum portfolio. Next, by using yfinance(Yahoo Finance), we collected stock price data through a certain lookback window(e.g., 12 months, 6 months, etc.). Based on the data, we calculated the returns of all candidates and ranked them accordingly from the worst return to the best return. We selected top 5 winners to form the momentum portfolio. We also tried several weighting strategies that are used to assign our wealth into every component of the portfolio. After a certain holding period(e.g., 1 month, 3 month, etc.), we used our model to rebalance the portfolio. We checked our final wealth after one year to see whether momentum strategy is robust or not.

Lookback window and holding period are two critical components in our project. We developed eight strategies with different lookback windows and holding periods to test which combination works the best in momentum investing.

Our paper contains the following parts: Introduction, Data Collection, Data Cleaning, Model Implementation, Model Validation, Conclusion, Limitations and Improvements, Appendix and Reference.

1. **Data Collection**

The major data we acquired for this project is S&P 500 stock tickers and their historical daily stock price. We were aiming to construct a 100-stock universe from S&P500 for our momentum investing portfolio to choose from.

First, we noticed the dynamic of the market and the change in S&P 500 stock tickers throughout the years. Stocks could be delisted from publicly traded platforms due to various reasons and tickers could be added or deleted from the S&P 500 index. If we used current S&P 500 stock, we would encounter survivor bias in our ticker selection. We were using the stocks that already survived in the market. Yet, from the perspective of the past with past information, we were not able to assess which stock could be delisted and which stock could survive. Therefore, we acquired the historical list of S&P 500 tickers by date. (Github fja05680) When we construct our portfolio universe, we use the S&P 500 tickers list at the initial date and focus only on the selected universe. In such a way, we choose stocks with past information(to avoid survivor bias) and will not be distracted by the change in S&P 500 tickers.

The data source we used is Yahoo Finance and the stock price column we used is“Adjusted Close Price”. This price reflects the value changed due to corporate action, such as dividends, saving extra steps for our data cleaning.

1. **Data Cleaning**

The major steps for our data cleaning process are filtration and dealing with null values. For filtration, we set up a minimum stock price (~$10) to filter out “penny stock”. Some of these stocks have strong momentum signals yet very little price change. These stocks could be too suspicious to trade.

A null value in our data frame is caused by delisted stocks. The system will automatically fill in the past highest value when being used for calculation during our portfolio rebalance phase. To avoid this, we choose to fill in the null value with the last good value and modify our algorithm such that our portfolio will only pick up positive momentum signals. An equal value will cause 0 momentum thus we would avoid disturbance from the null value.

1. **Model Implementation**

Like moving physical objects have formulas to calculate their momentum, we assume stocks have an underlying series of mechanics carrying their ups and downs. If a stock had been moving up for 12 months, then we say it had enough momentum to keep moving up for the next month. Therefore, the design focus of the model was around the lookback window and portfolio holding period. Different combinations of the two time periods formed the basic structure of the strategies that we want to put to test.

To better compare the effectiveness of the strategies, we set them to have the same portfolio start date with a one-year duration. At the initial date of the portfolio, the program generates one hundred random stocks from the S&P 500 lists of that date. Each strategy picks a certain look-back window to calculate the momentum during the past time frame and decides which five stocks with the highest momentum to hold for the following period. When each strategy comes to the end of its holding period, it recalculates the momentum from the updated lookback window and picks five new stocks to hold for the next portfolio holding period. After a one-year duration from the portfolio's initial date, we get the YTY payoff of the strategies. Due to the random nature of haphazardly selecting one hundred stocks from the S&P 500 index, we repeat the process fifty times to get the average portfolio performance for the year to minimize sample bias.

Another parameter we want to investigate is how to distribute the one million dollar initial capital into the five stocks to maximize our return. We came up with three algorithms to compare their effectiveness, which are equal weight, inverse volatility, and optimal portfolio. Intuitively, equal weights mean the capital is evenly invested in the five stocks in each period and is a perfect base case for comparison. The inverse volatility weighting method could potentially reduce portfolio returns while decreasing drawdowns and increasing the Sharpe ratio. The optimal portfolio is theoretically the strongest candidate as it allocates capital based on the generated efficient frontier.

After finishing setting up all the input parameters, it is time to test the outputs. We would like to compare the annualized returns with SP500 index returns to see the general effectiveness of the model as well as tweak with individual parameters to observe the impact each one does on the system. For realistic purposes, we had to take into consideration the portfolio path, max drawbacks, and variance of the returns as much as the average final returns.

1. **Model Validation**

**Appropriate loop times to test the strategies:**

We loop 50 times with 50 different tracking universes of the same start date to test our momentum strategies. In the testing, we apply our momentum strategies from the same universe to 100 other tracking universes. However, since the 100 tickers of the universe are randomly taken from S&P 500 at that date, there will be a significant difference with varying universes of tracking. We found that only using the strategy once with one tracking universe can lead to substantial volatility. Hence to improve our estimation accuracy, we need to test for the best number of tracking universes we need. We recorded all payoffs of different tracking universes. For example, with one universe(100 tickers) starting at the date 2010-01-01, we will have one payoff, and with 100 universes, we will have 100 different payoffs and calculate their volatility. With data from 1 to 100 universe, the payoff volatility result is shown in Appendix 9. The plot shows that the volatility is enormous when we only use a few different universes. However, when the number of universes goes after 50, the volatility starts to converge, and as a result, we decided to use 50 as our test times for our momentum strategies.

**The difference among eight strategies with various holding periods and lookback windows:**

In validation, we tried eight strategies with different lookback windows and holding periods. In choosing the length of the lookback window and holding period, we learned from the paper “Value and Momentum Everywhere”(Asness, Clifford) and added some strategies. The most representative three strategies are strategy1 with 12 months lookback window and one month holding period, strategy8 with an 11-month lookback window, a 1-month gap and three month holding period, and strategy6 with a 1-month lookback window and one month holding period.

Strategy1 is the most common strategy in choosing the lookback window and holding period. This is the most stable strategy in our testing (Appendix 1-6 ). The ranking of this strategy is only sometimes the first, but it showed in the first three most times from 2010 to 2020.

Strategy 8 is the optimal strategy we learned from paper, and its strength is that it can get the most payoff sometimes, like in the year 2018. However, the disadvantage is that with a long period of holding, the stock price variation can be significant, which is why this strategy sometimes shows in the middle ranking or bottom half. Hence, we can use this strategy to maximize our profit during a stable year like 2019, but we will not choose this strategy in the year with some uncommon accidents and fluctuations like the year 2011 and year 2015 (S&P 500 showed negative returns in these two years).

Strategy6 is a test for using a short-length lookback window. In the testing, this strategy showed diverse results in different years. For instance, in the year 2014,2018, with both the “equal” and “inverse volatility” weighting methods, this strategy showed up in the bottom half of the rank, but in 2010, this strategy performed well. As a result, we don’t recommend using this strategy since the risk is too high. However, a shorter length of lookback window and holding period can be our future investing direction since, with a shorter period, people can see the outcome and payoff fast.

**Results’ difference from 10 different years:**

Because of the length limitation of the paper, we didn’t put all the testing results in the Appendix, but in the test from 2010 to 2020, the momentum strategy1, the most stable strategy, beat the market return most times except the year 2019. Also, in the testing, we can find a correlation between the market return and the return of our momentum strategy. However, we believe that by adding a short position to our strategy, this impact will become lower, and our strategy will have less risk.

**Difference of 3 weighting strategies:**

As mentioned above, we used three different weighting strategies for our momentum strategies, and one strategy can not be used due to the data problem (abnormal data showed in testing). We found that the other two strategies show similar features in different years. Hence, we use one-year data as an example. In 2014 (Appendix 2, Appendix 5)， we can find that the two different weighting strategies do not significantly impact the ranking of each strategy, and only strategies in the middle rank are affected. The first two strategies and the last two strategies stay the same rank.

Though these two strategies do not vary by year, they showed a relatively strong power in the average payoff of each year. In 2014, the average payoff of 8 strategies with the “equal” weighting method led to an unacceptable result. Only the first two strategies in the ranking have a positive payoff. As a comparison, the “inverse volatility” weighting method lets the first five strategies in the ranking have positive outcomes. However, the payoff of the first two strategies of different weighting methods is similar. This result tells us that the “inverse volatility” weighting method can not let us make more money, but with this method, we can try different strategies with less loss. Hence, we can use the “inverse volatility” weighting method to lower the potential risk when unsure about which weighting strategy we will choose when investing.

**Abnormal data in testing:**

In the testing, we found an abnormal result in the year 2011 with the strategy “Optimal Portfolio.” In that year, we see the wealth went from 1M to 400 M. By tracing back to the data for that year, we found that the monthly payoff from January to February and from March to April appears problematic (Appendix 7). Then we checked the portfolio we held for those two months and found that stock “TIE” jumped from 17.02 to 19000 in one day (Appendix 8), which is ridiculous. We then delete this stock from our stock tracking universe.

From the previous wrong data problem, yfinance, one of the most popular data resources, is not liable for doing data testing.

1. **Conclusion**

After implementing eight strategies and two weighted scenarios (due to data source issues, we did not include optimal portfolio weighted strategy in final conclusion) on three different initial dates, we reached several conclusions.

First, holding the weighted scenario constant, we found out that in different years, the performance of the same strategy can vary a lot. Appendix 1, 2, and 3 are results for the same equal-weighted portfolio with initial dates of 2010-01-01, 2014-01-01, and 2018-01-01. From the three tables, we can see that the best strategy in Appendix 1 is strategy 1, the best strategy in Appendix 2 is strategy 4, and the best strategy in Appendix 3 is strategy 8. Further, taking strategy 6 as an example, it ranked number two in Appendix 1, while it ranked last in both Appendix 2 and Appendix 3. Those observations suggest that the performance of the same strategy can vary when the initial date is different.

Second, comparing those three initial dates, we can see that strategies with longer lookback windows always rank higher than strategies with shorter lookback windows with few exceptions. One noticeable exception is strategy 6 (whose lookback window is only 1 month). In 2010, strategy 6 performed outstandingly in both weighted scenarios, this might suggest that the stock market was quite volatile during that period. However, most of the time, strategies with longer lookback windows should be recommended to investors.

Third, holding the initial date and strategy constant, an inverse volatility weight scenario can sometimes enhance the profit or reduce the loss, but the impact is limited. For example, comparing strategy 6 in Appendix 1 and strategy 6 in Appendix 4, the inverse volatility weight scenario gave us an average profit of 58387.56, while the equal-weighted scenario only gave us -2857.16. When comparing strategy 7 in Appendix 1 and strategy 7 in Appendix 4, the inverse volatility weight scenario gave us an average loss of -10531.73, while the equal-weighted scenario gave us a greater loss of -41322.95. However, the inverse volatility weight scenario cannot guarantee better results all the time, so we concluded that this impact is limited in our model.

Fourth, as mentioned before, the highest-ranking strategy varies in different years; therefore, we cannot simply conclude the best strategy. However, we noticed that strategy 1 (whose lookback window is 12 months, and holding period is 1 month) can always rank in the top 4, so we think that this strategy is a relatively great and stable strategy for investors.

At last, when we fix the holding period to 3 months, the 11-month lookback window always works better than the 12-month lookback-window with few exceptions. This can be explained by the reversal effect in the most recent month. As Lehmann(1990) suggests in his paper, “market makers are only intermediaries between patient and impatient traders,” hence, when stock prices move, impatient traders will make immediate adjustments in the demand for liquidity. While “market-wide proprietary trading by large broker-dealers (who are patient traders) is probably the natural source of supply of liquidity over intervals like days or weeks.” Therefore, this imbalance can create short-term reversals in the market. Additionally, Jegadeesh(1990) also mentioned this issue in his paper, and he further explained that the “thin trading phenomenon” and “the presence of bid-ask spread bias” are two main reasons for this reversal effect.

1. **Limitations and Improvements**

Our project is not perfect, it contains some limitations and needs further improvements. First of all, we can further modify our portfolio. To improve it, we can add short positions and sector filters. Second, Yahoo Finance is the main data source of our project; however, we noticed that Yahoo Finance contains some invalid data which can result in unreasonable payoffs in our model. Therefore, we need to find more reliable sources to support our model. Third, when comparing strategies, we only used the average final wealth. However, average final wealth can only provide us with limited aspects of the result, and to enhance our conclusion, we should compare more aspects, such as variance. Fourth, our conclusions were drawn without any statistical tests to further verify the results. So, we should add some statistical tests next time to make sure that the results are statistically significant and conclusions would be drawn.

1. **Appendix**

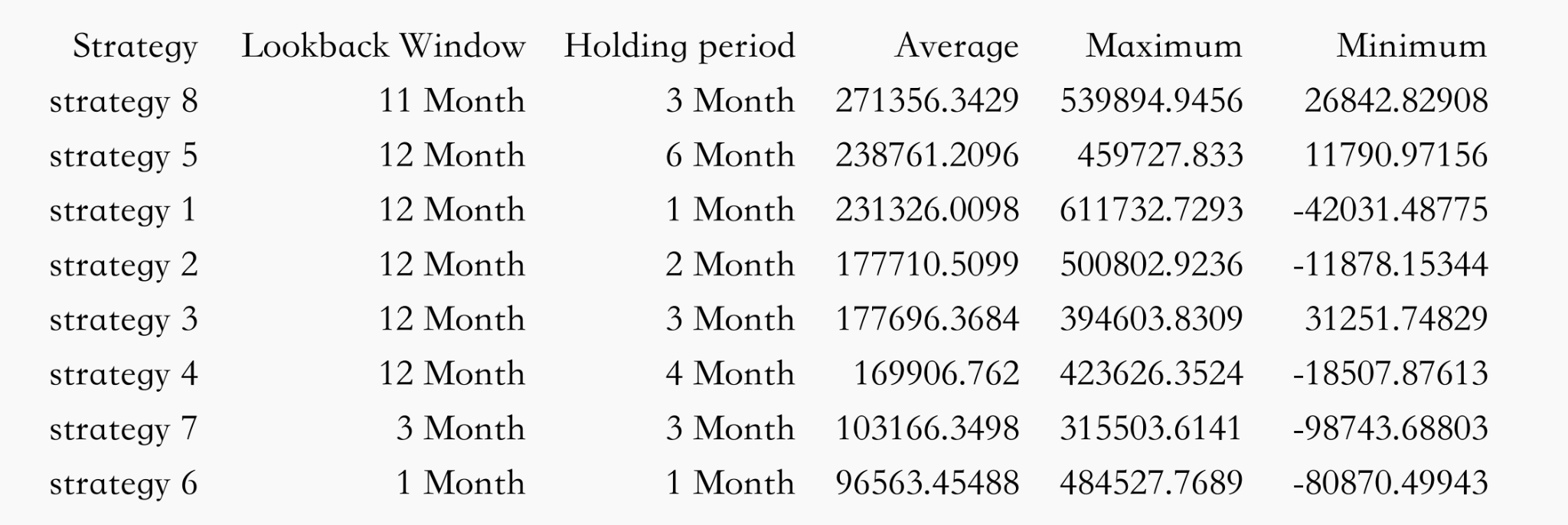
Appendix 1. Equal-weighted strategy(Ranked by Average); Initial start date: 2010-01-01



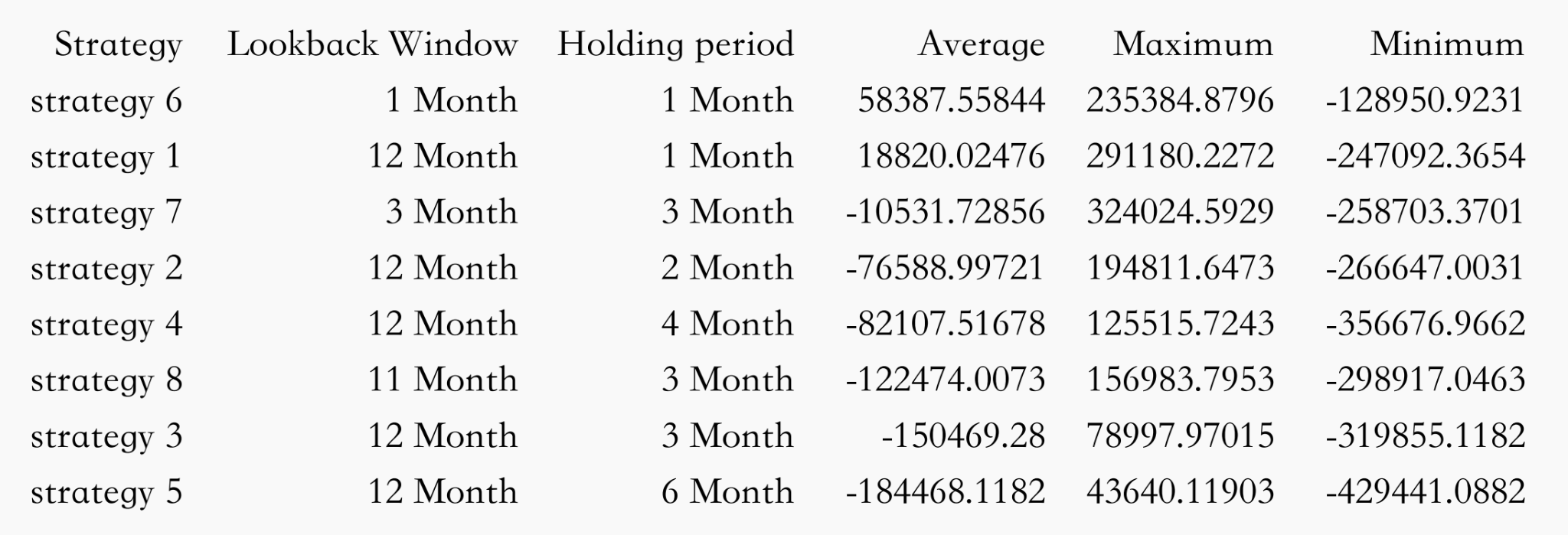
Appendix 2. Equal-weighted strategy(Ranked by Average); Initial start date: 2014-01-01



Appendix 3. Equal-weighted strategy(Ranked by Average); Initial start date: 2018-01-01



Appendix 4. Inverse volatility strategy(Ranked by Average); Initial start date: 2010-01-01



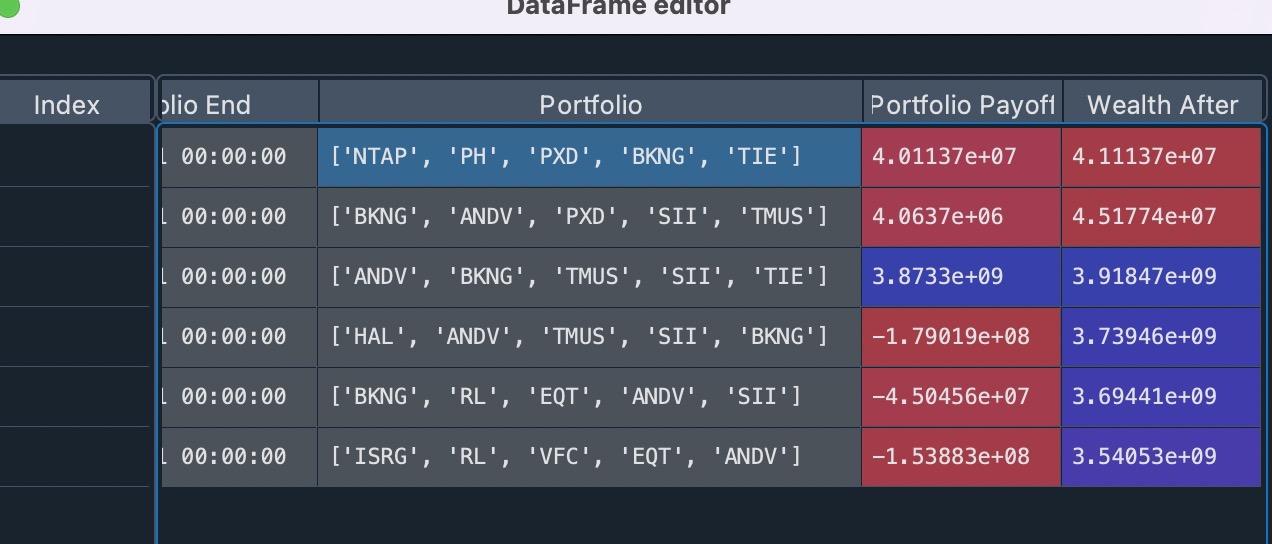
Appendix 5. Inverse volatility strategy(Ranked by Average); Initial start date: 2014-01-01



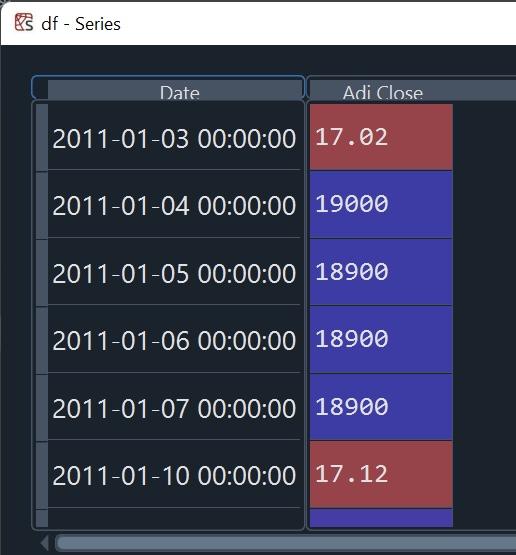
Appendix 6. Inverse volatility strategy(Ranked by Average); Initial start date: 2018-01-01



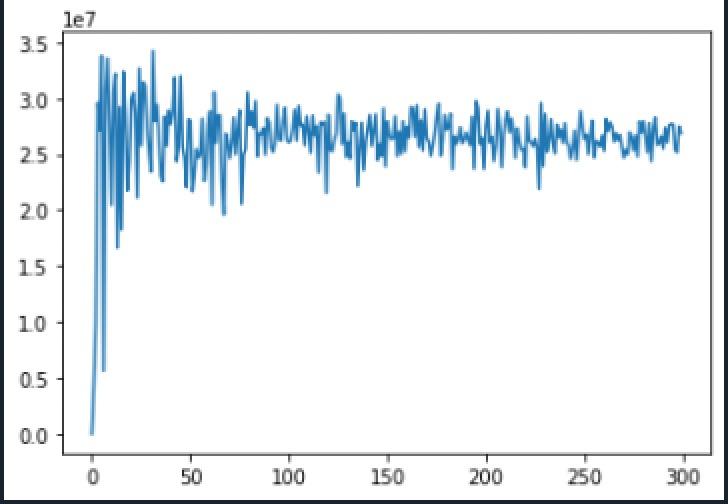
Appendix 7. Abnormal payoff in January and March, 2010



Appendix 8. Wrong data of Stock TIE in year 2010



Appendix 9. Plot of Volatility with different loop times(2010-01-01, “equal” weight, Strategy 1)



1. **Reference**

1. Asness, Clifford S., et al. “Value and Momentum Everywhere.” The Journal of Finance (New York), vol. 68, no. 3, 2013, pp. 929–985.

2. fja05680. “FJA05680/SP500: Current and Historical Lists of S&P 500 Components since 1996.” *GitHub*, https://github.com/fja05680/sp500.

3. Jack Vogel, PhD. “How to Measure Momentum?” *Alpha Architect*, 10 May 2022, https://alphaarchitect.com/2016/10/how-to-measure-momentum/.

4. Jegadeesh, N. “EVIDENCE OF PREDICTABLE BEHAVIOR OF SECURITY RETURNS.” The Journal of Finance (New York), vol. 45, no. 3, 1990, pp. 881–898.

5. Jegadeesh, Narasimhan, and Sheridan Titman. “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency.” *The Journal of Finance*, vol. 48, no. 1, 1993, pp. 65–91. JSTOR, https://doi.org/10.2307/2328882. Accessed 12 Oct. 2022.

6. Lehmann, Bn. “Fads, Martingales, and Market Efficiency.” The Quarterly Journal of Economics, vol. 105, no. 1, 1990, pp. 1–28.

7. “Momentum Portfolios.” *Varsity by Zerodha*, 12 July 2019, https://zerodha.com/varsity/chapter/momentum-portfolios/.

8. Tyagi, Parth. “Portfolio Optimization in Python: Modern Portfolio Theory in Python.” *Analytics Vidhya*, Data Science Blogathon, 27 Apr. 2021 https://www.analyticsvidhya.com/blog/2021/04/portfolio-optimization-using-mpt-in-python/.