Problem Statement

This project is an independent project, which is not listed in the suggested projects document. The objective of this project is to let a multi-arm agent select tickers that will make up an investment portfolio. The portfolio will then be detailed in terms of distribution of initial funds into each of selected stocks, and the portfolio that lies on the efficient frontier will be selected. Its performance will be compared to the benchmark SPY stock over the same number of trading days.

Background

The bandit problem is a classic decision-making scenario in which an agent must repeatedly choose from a set of options, each associated with an uncertain reward. The agent aims to maximize its cumulative reward over time while facing the challenge of balancing exploration and exploitation. Multi-Armed Bandits can be used to model and solve such decision problems. In the context of this project, MAB was applied to allocate resources among various assets, and used to created an investment portfolio.

The investment portfolio is then used in the minimum variance allocation problem. The goal of this problem is to minimize a portfolio's overall volatility. The variance that needs to be minimizes refers to the statistical measure of the portfolio's deviation from its expected return. The covariance between the returns of different assets is analyzed to achieve a diversified portfolio, and one that provides the lowest overall risk for a given level of expected return.

Methodology

The dataset used is the daily open high low close dataset from 01/03/2023 to 11/17/2023, for 500 S&P companies. 25 companies or tickers are randomly selected. The world chooses a ticker from the collection 25 stocks, and asks whether or not the return of this ticker at the end of today was greater or equal to the risk-free rate, and updates the world accordingly. 10000 agents are deployed and simulated. For each agent, their preferences for each of the 25 stocks are ranked based on the reward: if the daily return is greater than or equal to the risk-free rate, the agent receives a reward of r = 1. Otherwise, the agent receives a r = 0 reward. All 10000 agents' preferences were evaluated and normalized based on the frequency a stock was selected to be top 3. The top 6 stocks that were chose most often makes up the portfolio.

After the portfolio was chosen, it gets analyzed in the minimum variance allocation problem. The initial fund is $2500. A minimum desire reward level is selected, then efficient frontier that represents a set of portfolios that offer the highest expected return for that desired return. The expected value of the return is calculated, discounted to present value, and then compared to the benchmark SPY stock over the same period of time, which is 120 trading days.

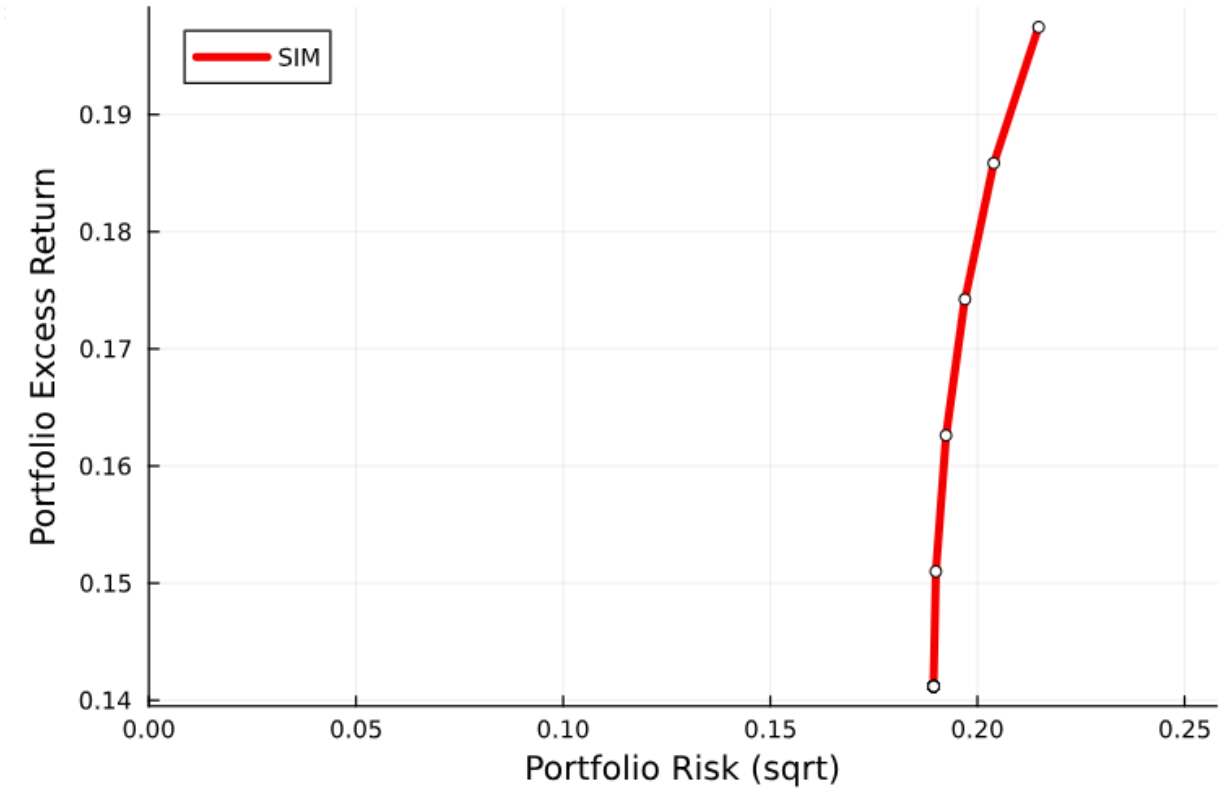
Results

The bandit-problem ticker picker outputs the following top preference array with tickers and their respective frequency:

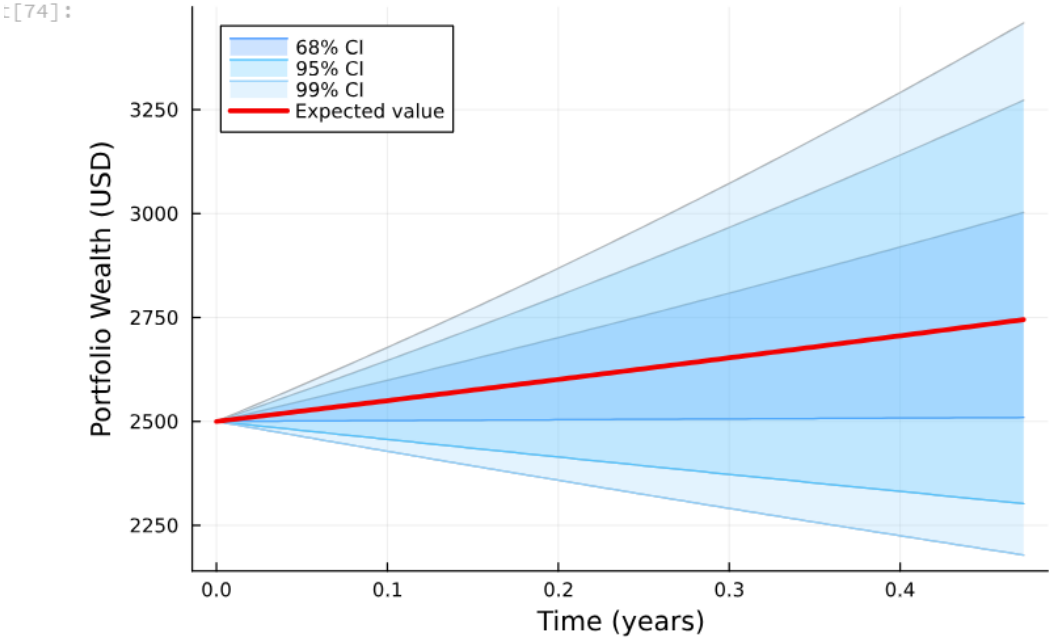
"MSFT" => 0.4586
"JPM" => 0.3211
"AEP" => 0.3595
"C" => 0.1915
"MRK" => 0.2982
"V" => 0.3558
"TMUS" => 0.2351
"OKE" => 0.4432
"CRL" => 0.3829
"TSLA" => 0.3568
"CZR" => 0.3598
"VRTX" => 0.3778
"AMD" => 0.3254
"BSX" => 0.3641
"TXT" => 0.2721
"INTC" => 0.239
"IVZ" => 0.272
"IT" => 0.3723
"DIS" => 0.165
"ECL" => 0.2771
"URI" => 0.297
"FIS" => 0.253
"PEP" => 0.317
"MET" => 0.361
"AAPL" => 0.3907

Thus, the 6 stocks that were selected were["MSFT", "IT","BSX","AAPL","VRTX","CRL"]

The efficient frontier calculated from historic data look like this:



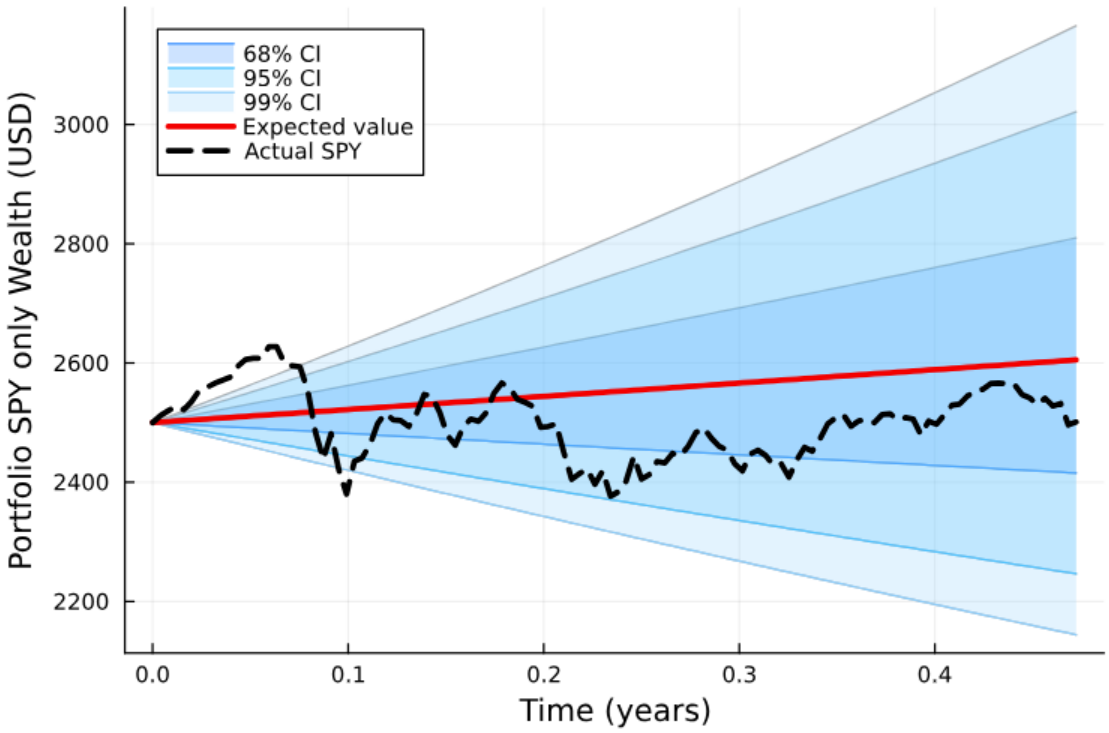
The Expected Return for the portfolio over the 120 trading days are



The selected portfolio should be worth (L,E,U) = (2509.51, 2745.03, 3002.64) USD by the end of term. The expected NPV for this portfolio is $184.27.

Compared to the benchmark SPY, the portfolio performs like so

]:



SPY only portfolio should be worth (L,E,U) = (2415.48, 2605.17, 2809.76) USD. The expected NPV for this portfolio is NPV = $47.5

Thus, Although the NPV is lower than the initial investment. This portfolio still outperformed the benchmark SPY's NPV.