Final Report: HEALTHCARE PORTFOLIO OPTIMIZATION USING MPT

Ashutosh Singh

College of Professional Studies, Northeastern University

EAI 6050: Finance Information Processing

Prof. Zhi He

December 17, 2022

Abstract

This study sought to determine if a fund or portfolio built using a market index of risky assets is the best possible portfolio. The study in this report examines the business strategies, corporate goals, market research, and whether these factors have been represented in the balance sheets of corporations in the healthcare industry. This was investigated by building an imaginary optimal portfolio for the US stock market's healthcare sectors and analyzing the investment weights in it (investment allocations). The research also examined the possibility of strong returns from actively managed sector-based funds in 2021. The analysis's findings can vary, but, the best-performing portfolios are those that have their asset allocation set for long-term investments. The Modern Portfolio Theory (MPT), the efficient frontier and how it is produced, and how to get an optimal portfolio value are all explained in the literature review of the study. I want to convey my sincere gratitude to my instructors, Zhi He, to whom I am indebted for the priceless information I acquired via their lessons as well as for their insightful conversations and ongoing mentoring outside of office hours.

I have examined the stocks using both business and technical analysis, which will enable me to better comprehend its future.

Introduction

Markowitz is credited with laying the groundwork for the Modern Portfolio Theory (MPT) that is used to create and choose an appropriate portfolio today. After doing thorough research of the US stock market's listed companies, several solid businesses with bright futures may be identified. When making investment selections, the next phase must consider two key issues. Which of those corporations should we invest in first? Second, how much money should we put into the firm and what is the weight of each investment? The Modern Portfolio Theory (MPT) suggests using Markowitz's portfolio theory to choose a portfolio that is suited to the investor's goals, even if there may be other traditional and original approaches to answer the problems. According to the hypothesis, there are several asset combinations that may be used to create effective portfolios. At a specific risk threshold, each portfolio produces the highest return achievable (or the lowest possible risk at a given return). To determine if purchasing the index is the best portfolio strategy to maximize the Sharpe ratio, this article conducts an empirical analysis. If not, the next consideration is whether an ideal portfolio can or would outperform an index fund. The research will also examine if an optimal portfolio choice may outperform S&P 500 index. More information on the MPT and how it aids investors in choosing the best portfolio from a variety of efficient portfolios will be provided in the literature review of this article. I have picked a dataset for my research, and I have imported the data using the pandas data reader library. The dataset has 253 rows and 14 columns. For my investigation, I just used the closing price and used the top 14 healthcare businesses. These businesses include Abbott Laboratories (ABT), AstraZeneca plc (AZN), GSK plc (GSK), Elevance Health Inc (ELV), Novartis AG(NVS), Merck & Co Inc (MRK), UnitedHealth Group Inc (UNH), Johnson & Johnson (JNJ), Eli Lilly and Co (LLY), Pfizer Inc (PFE), Novo Nordisk A/S (NVO), Novo Nordisk A/S (NVO (MRNA).

What is Modern Portfolio Theory?

When building a portfolio based on projected return and amount of risk, portfolio managers are said to use the Modern Portfolio Theory as a framework. MPT draws on several ideas, but it is mostly based on Harry Markowitz's "Portfolio Selection" article from The Journal of Financial in March 1952 and his 1959 book "Portfolio Selection Efficient Diversification." In his paper, Markowitz outlines the two steps that make up the portfolio selection process:

1. "Begin with observation and experience and concludes with views about the future performances of accessible securities." The majority of this stage is focused on financial factors. Financial analysts, accountants, and microeconomists first conduct a thorough examination of the company values, anticipated performance, and relevant markets. When they locate the businesses that are thought to have a bright future, the process is then over.
2. 2. "Starts with the pertinent belief about future performance and finishes with the selection of the portfolio". This stage begins with the findings from stage one (the bright prospects of several stocks) and concludes with the identification of a number of potentially effective portfolios that are estimated to have the highest expected return at a given risk level or the lowest risk level at a given expected return. This phase results in the Modern Portfolio Theory as we know it today.

The method to choose the best portfolio of hazardous assets from a list of efficient portfolios was developed by Markowitz. James Tobin first proposed the concept that there are two stages to the process of choosing an investment in 1958. The best portfolio of hazardous assets is chosen in the first phase using Markowitz's method. The investor's decision to divide funds between the ideal portfolio and a single risk-free asset occurs in the second phase. According to Modern Portfolio Theory, also known as the mean-variance analysis [or approach], it is possible to carry out an optimization that yields the risk/return or mean-variance efficient frontier "given estimates of the returns, volatilities, and correlations of a set of investments and constraints on investment choices [for example, the portfolio expected return to be greater than the return of risk-free asset]."

Rationale Picking Healthcare Sector and Stocks

I chose the healthcare sector for my analysis because I believe that, as a portfolio manager, I can provide my clients with greater returns with less risk in the healthcare sector. When compared to other industries, such as the aviation business, which can be significantly damaged by epidemics and war, the risk factors for the healthcare sector are quite low. Nevertheless, we can see a spike in the demand for healthcare during the pandemic and war. People are more worried about their health now than they were before COVID-19, which is why they go for routine checkups and take daily medications and multivitamins. As a result, there is a large demand for both medical supplies and medications. Also, Many Healthcare companies are working on HIV and cancer vaccines, which might be a major boost for the healthcare industry in the future. The healthcare sector is also receiving a lot of relaxation from the government. Globally, there is a significant need for modern medical technologies. The financial sheets of pharmaceutical companies like MRNA, LLY, and ABT are extremely sound, and their returns on invested capital (ROIC) are high or rising. Many investors are attracted by the sector’s impressive numbers: At the end of 2019, health spending made up approximately 18% of the US GDP, and by 2028, it is projected to reach $6 trillion annually.

I'd like to talk about the few stocks I chose for my study. My personal favorite is Moderna, which has a very solid balance sheet for 2020 and an outstanding change in revenue from 2019 to 2020 for MRNA of 4245%. With mRNA technology, they are developing cancer vaccines, and they are open about their work. Eli Lilly and Co. (LLY), the competing firm, has a 5-Year Avg. LLY's annualized return was -35.5%, whereas the average annual return for the 14 equities I picked was 24.4%. UHG is a top-ten Fortune 500 company and a market leader in health insurance and healthcare goods. Pfizer Inc. (PFE) and AstraZeneca plc. (AZN) are two other businesses that are dominating the covid vaccine industry by providing vaccines all over the world.

Build a portfolio with Modern Portfolio Theory (2020 data)

Step 01: - Importing the data: - I imported the data using the yfinance Python package, and using it, we can gather the financial information from Yahoo. I first retrieved data for the period of time from January 1, 2020, to December 31, 2020. There are 14 columns and 253 rows in the data set. For my investigation, I just used the closing price and used the top 14 healthcare businesses. These businesses include Abbott Laboratories (ABT), AstraZeneca plc (AZN), GSK plc (GSK), Elevance Health Inc (ELV), Novartis AG(NVS), Merck & Co Inc (MRK), UnitedHealth Group Inc (UNH), Johnson & Johnson (JNJ), Eli Lilly and Co (LLY), Pfizer Inc (PFE), Novo Nordisk A/S (NVO), Moderna (MRNA), Danaher Corp (DHR).

Date	ABT	AZN	DHR	ELV	GSK	JNJ	LLY	MRK	MRNA	NVO	NVS	PFE	RHHBY
2020-01-02	86.949997	50.389999	155.110001	300.869995	46.919998	145.970001	132.210007	87.824425	19.230000	58.360001	94.949997	37.134724	40.889999
2020-01-03	85.889999	50.090000	154.149994	296.880005	46.480000	144.279999	131.770004	87.070610	18.889999	57.110001	94.790001	36.935486	40.650002
2020-01-06	86.339996	49.880001	154.610001	300.450012	46.500000	144.100006	132.259995	87.442749	18.129999	57.009998	95.430000	36.888046	41.020000
2020-01-07	85.860001	50.070000	156.130005	299.540009	46.209999	144.979996	132.509995	85.114502	17.780001	56.950001	94.480003	36.764706	40.700001
2020-01-08	86.209999	49.950001	156.289993	307.480011	46.410000	144.960007	133.710007	84.541985	17.980000	56.849998	94.480003	37.058823	40.840000
...	...	...	...	...	...	...	...	...	...	...	...	...	...
2020-12-23	107.449997	48.770000	220.570007	308.350006	36.230000	151.940002	165.479996	76.106873	130.339996	69.620003	88.379997	37.439999	42.490002
2020-12-24	108.349998	48.520000	221.490005	308.670013	36.139999	152.470001	166.660004	76.469467	123.389999	69.430000	88.550003	37.270000	42.549999
2020-12-28	107.790001	49.380001	222.750000	312.890015	36.290001	153.190002	166.500000	76.765266	111.400002	70.250000	91.199997	36.820000	43.040001
2020-12-29	108.330002	49.900002	222.860001	314.329987	36.980000	154.139999	166.580002	77.690842	114.389999	70.580002	93.349998	37.049999	43.310001
2020-12-30	108.440002	50.180000	220.679993	314.049988	37.040001	156.050003	167.009995	76.898857	111.129997	70.199997	94.360001	36.740002	43.840000

252 rows × 14 columns

Step 02: - Visualizing the data: - We can see that the revelation of the vaccination has caused the price of Moderna's shares, MRNA, to skyrocket. Additionally, we can see that Danaher Corp (DHR) and United Health Group (UHG) will have healthy growth in 2020 along with Moderna (MRNA).

Step 03: - Checking null value: - There is no null value exists for any of the stock.

```
Health_portfolio_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 252 entries, 2020-01-02 to 2020-12-30
Data columns (total 14 columns):
 #   Column  Non-Null Count  Dtype  
---  ------  --------------  -----  
 0   ABT     252 non-null    float64
 1   AZN     252 non-null    float64
 2   DHR     252 non-null    float64
 3   ELV     252 non-null    float64
 4   GSK     252 non-null    float64
 5   JNJ     252 non-null    float64
 6   LLY     252 non-null    float64
 7   MRK     252 non-null    float64
 8   MRNA    252 non-null    float64
 9   NVO     252 non-null    float64
 10  NVS     252 non-null    float64
 11  PFE     252 non-null    float64
 12  RHHBY   252 non-null    float64
 13  UNH     252 non-null    float64
dtypes: float64(14)
memory usage: 29.5 KB
```

Step 04: - Statistical analysis of the data: - We can see that the MRNA, DHR, and UNH stocks have extremely large standard deviations, which indicates that these stocks have extremely high volatility. We may select these stocks for large returns since they have a superior risk to reward ratio.

	ABT	AZN	DHR	ELV	GSK	JNJ	LLY	MRK	MRNA	NVO	NVS	PFE	R
count	252.000000	252.000000	252.000000	252.000000	252.000000	252.000000	252.000000	252.000000	252.000000	252.000000	252.000000	252.000000	252.0
mean	96.060635	51.865714	185.383095	278.743293	40.125198	145.770516	148.735000	77.163304	62.337619	64.812262	87.350079	34.950254	42.8
std	10.755836	4.163520	31.500122	26.778455	3.292716	6.673327	10.454763	4.005207	33.447821	4.376666	4.848810	2.531011	1.8
min	62.820000	37.790001	121.389999	174.679993	31.850000	111.139999	119.050003	63.358780	17.780001	49.459999	70.669998	27.030361	35.2
25%	89.017498	49.895001	160.967499	265.467499	37.497500	143.869995	141.597500	74.680342	31.790000	62.754999	84.849998	33.816413	41.7
50%	94.439999	52.994999	178.419998	277.389999	40.474998	147.409996	148.885002	76.803436	64.735001	65.355000	87.035000	35.175522	43.2
75%	106.527498	54.727501	213.042500	296.980003	41.857499	149.602505	155.540001	79.191319	72.947500	67.560001	89.797499	36.423149	44.0
max	114.419998	61.099998	245.460007	333.149994	47.889999	156.050003	172.630005	87.824425	169.860001	73.800003	99.010002	42.560001	46.8

Methodology

Method 1: - Mean-Variance Optimization

The main objective of portfolio theory is to divide your money amongst various assets as efficiently as possible. You may make this allocation using a quantitative technique called mean-variance optimization (MVO), which considers the trade-off between risk and return. Your predicted return is considered when allocating your portfolio, subject to a chosen degree of risk. Markowitz came up with the idea of portfolio optimization utilizing the mean and variance initially. Finding the standard deviation of a portfolio as a gauge of its risk (the square root of the variance) was an important idea in this study.

Sharpe Ratio: - The return-to-risk ratio is known as the Sharpe ratio. The Sharpe ratio rises when the risk is low and the rewards are high. The algorithm seeks for the portfolio with the best return and lowest risk, or one with the highest Sharpe ratio. In the end, the portfolio performs better a greater the Sharpe ratio.

```
#calulating weights with max sharpe statistics
from pypfopt.efficient_frontier import EfficientFrontier

ef = EfficientFrontier(mu, S)
weights = ef.max_sharpe()

cleaned_weights = ef.clean_weights()
print(dict(cleaned_weights))
```

```
{'ABT': 0.0, 'AZN': 0.0, 'DHR': 0.29916, 'ELV': 0.0, 'GSK': 0.0, 'JNJ': 0.0, 'LLY': 0.0, 'MRK': 0.0, 'MRNA': 0.6324
3, 'NVO': 0.0, 'NVS': 0.0, 'PFE': 0.0, 'RHHBY': 0.0, 'UNH': 0.0684}
```

I am following the above approach but if the investor is looking for low volatile stock the distribution of stock will be different as below.

```
ef.portfolio_performance(verbose=True)
```

```
Expected annual return: 318.7%
Annual volatility: 64.1%
Sharpe Ratio: 4.94

(3.1874170918036686, 0.6407163815909143, 4.943555655528731)
```

Discrete allocation of shares: -

```
In [15]: #converting stocks into actual allocation values
         from pypfopt.discrete_allocation import DiscreteAllocation, get_latest_prices

         latest_prices = get_latest_prices(Health_portfolio_df)

         da = DiscreteAllocation(weights, latest_prices, total_portfolio_value=500000)

         allocation, leftover = da.greedy_portfolio()
         print("Discrete allocation:", allocation)
         print("Funds remaining: ${:.2f}".format(leftover))
```

```
Discrete allocation: {'MRNA': 2845, 'DHR': 678, 'UNH': 99}
Funds remaining: $60.11
```

According to our methodology, we ought to buy 99 shares of UNH, 678 shares of DHR, and 2845 shares of MRNA. We can observe that our portfolio performs well, with a 318.7% percent predicted annual return. This performance is a result of Moderna's explosive expansion throughout the epidemic. Additionally, the portfolio optimization technique performs effectively with our existing data, as shown by the Sharpe ratio value of 4.94. This return is obviously overstated and is not likely to persist in the future.

Method 2: - Hierarchical Risk Parity (HRP)

Hierarchical risk parity got its name since the first thing we're going to do is cluster the financial assets. Each cluster will then be broken into smaller clusters, and so on, until we reach the individual assets. As a result, this procedure will establish a hierarchy, or a tree, that symbolizes the recursive process by which we split the universe of assets into clusters that will subsequently be further divided.

Portfolio Return: -

Date	ABT	AZN	DHR	ELV	GSK	JNJ	LLY	MRK	MRNA	NVO	NVS	PFE	RHHBY	UNH
2020-01-03	-0.012191	-0.005954	-0.006189	-0.013262	-0.009378	-0.011578	-0.003328	-0.008583	-0.017681	-0.021419	-0.001685	-0.005365	-0.005869	-0.010120
2020-01-06	0.005239	-0.004192	0.002984	0.012025	0.000430	-0.001248	0.003719	0.004274	-0.040233	-0.001751	0.006752	-0.001284	0.009102	0.006942
2020-01-07	-0.005559	0.003809	0.009831	-0.003029	-0.006237	0.006107	0.001890	-0.026626	-0.019305	-0.001052	-0.009955	-0.003344	-0.007801	-0.006037
2020-01-08	0.004076	-0.002397	0.001025	0.026507	0.004328	-0.000138	0.009056	-0.006726	0.011249	-0.001756	0.000000	0.008000	0.003440	0.021084
2020-01-09	0.002668	0.002603	0.008830	-0.003480	0.005171	0.002966	0.016528	0.008804	0.023359	0.010378	0.003069	-0.004352	0.000490	-0.005678
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2020-12-23	-0.007665	0.000821	-0.010586	0.016650	0.003879	-0.005107	-0.007438	0.004281	0.035431	-0.014160	0.002268	0.019053	-0.007938	0.007701
2020-12-24	0.008376	-0.005126	0.004171	0.001038	-0.002484	0.003488	0.007131	0.004764	-0.053322	-0.002729	0.001924	-0.004541	0.001412	0.009479
2020-12-28	-0.005168	0.017725	0.005689	0.013672	0.004151	0.004722	-0.000960	0.003868	-0.097172	0.011810	0.029927	-0.012074	0.011516	0.015141
2020-12-29	0.005010	0.010531	0.000494	0.004602	0.019013	0.006201	0.000480	0.012057	0.026840	0.004698	0.023575	0.006247	0.006273	0.004047
2020-12-30	0.001015	0.005611	-0.009782	-0.000891	0.001623	0.012391	0.002581	-0.010194	-0.028499	-0.005384	0.010820	-0.008367	0.012237	-0.006794

251 rows × 14 columns

Weight Calculation, Expected annual return and Sharpe ratio: -

```
hrp.portfolio_performance(verbose=True)
print(dict(hrp_weights))
```

```
Expected annual return: 17.4%
Annual volatility: 28.8%
Sharpe Ratio: 0.53
{'ABT': 0.049555789146461665, 'AZN': 0.07301545109813028, 'DHR': 0.07885228010540335, 'ELV': 0.053989668059532266,
'GSK': 0.07767268602777815, 'JNJ': 0.08232453274207821, 'LLY': 0.08872374904903048, 'MRK': 0.1323892087232456, 'MRN
A': 0.023918583283296924, 'NVO': 0.07999376339460092, 'NVS': 0.0833487382371751, 'PFE': 0.0742599871887004, 'RHHBY'
: 0.06350891689831022, 'UNH': 0.038446646046256444}
```

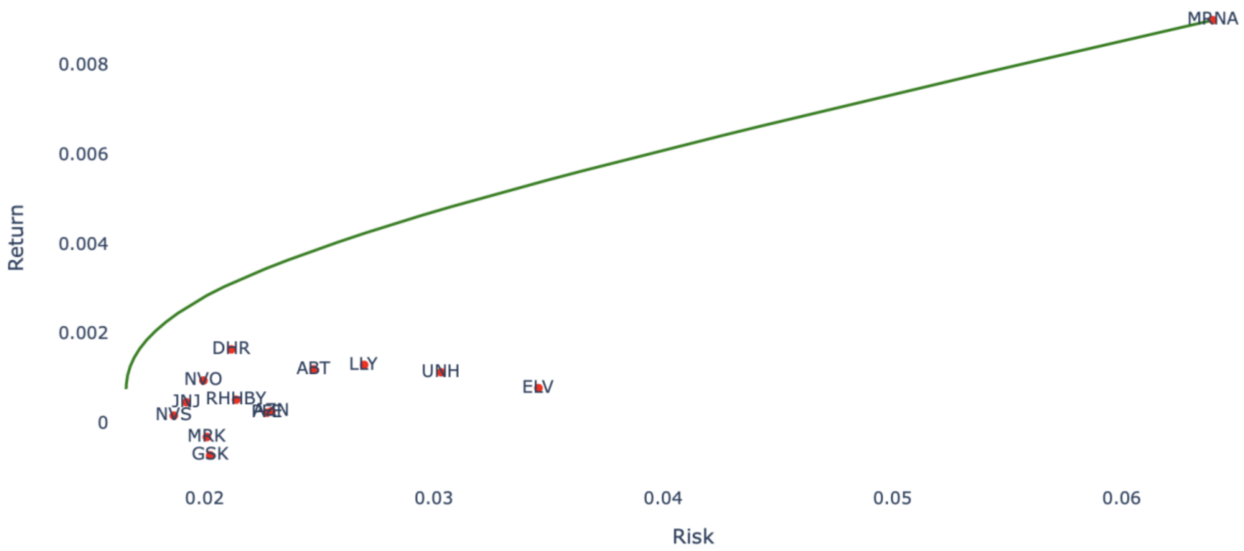
Discrete weight allocation: -

```
da_hrp = DiscreteAllocation(hrp_weights, latest_prices, total_portfolio_value=500000)

allocation, leftover = da_hrp.greedy_portfolio()
print("Discrete allocation (HRP):", allocation)
print("Funds remaining (HRP): ${:.2f}".format(leftover))
```

```
Discrete allocation (HRP): {'MRK': 860, 'LLY': 266, 'NVS': 442, 'JNJ': 264, 'NVO': 569, 'DHR': 179, 'GSK': 1048, 'P
FE': 1010, 'AZN': 727, 'RHHBY': 724, 'ELV': 86, 'ABT': 228, 'UNH': 56, 'MRNA': 107}
Funds remaining (HRP): $3.17
```

We can see that the inflated 225 percent we attained using mean-variance optimization is substantially lower than the predicted yearly return of 17.4 percent we have. In addition, we see a decreased Sharpe ratio of 0.53. Since HRP is less susceptible to outliers than mean-variance optimization, this conclusion is far more logical and more likely to hold up over time. We can conclude that we should build a portfolio of 107 shares of MRNA, 179 shares of DHR, 94 shares of UNH, 860 shares of MRK, 442 shares of NVS, 266 shares of LLY, 264 shares of JNJ, 179 shares of DHR, 569 shares of NVO, 1048 shares of GSK, 743 shares of AZN, 1010 shares of PFE, 727 shares of RHHBY, 228 shares of ABT. So the HRP method gives more diversified portfolio and good for long term investment and the stocks are less volatile as well.

Risk vs Return Curve: -



Risk(VaR) and Conditional Value at Risk(CVaR) at 99% confidence level

Value at risk (VaR) is a metric that attempts to assess the degree of financial risk experienced by a company or portfolio over a certain period of time. VaR may be calculated at different degrees of confidence and offers an estimate of the maximum loss from a specific position or portfolio over time.

There are two main ways to calculate VaR:

1.) Using Monte Carlo simulation

2.) Using the variance-covariance method

VaR, then, is a measurement of market risk. It is a measurement of the largest loss that, given a certain confidence interval and a time frame, is possible. Financial organizations can calculate the amount of capital reserves they need to cover losses using VaR. VaR also aids in determining if assets of higher-than-acceptable risk need to be lowered.

VaR99 for each Trading Day: -

```
VaR_99

array([-0.0221954 , -0.02888664, -0.02713725, -0.01575105, -0.0136109 ,
       -0.02574967, -0.04862932, -0.01066511, -0.01548178, -0.02592871,
       -0.02245076, -0.01851765, -0.03082661, -0.02370359, -0.0367622 ,
       -0.06695708, -0.03907071, -0.07266763, -0.0343036 , -0.0470217 ,
       -0.01350589, -0.00292052, -0.05964598, -0.0229779 , -0.11103259,
       -0.01456015, -0.06911183, -0.08995366, -0.02568234, -0.02889111,
       -0.02451771, -0.01044449, -0.03668465, -0.02100874, -0.09197387,
       -0.19125191, -0.12268739, -0.0868849 , -0.04319189, -0.0114875 ,
       -0.07729151, -0.02432132, -0.05092817, -0.04615793, -0.13996677,
       -0.05158588, -0.10349285, -0.12103209, -0.03348396, -0.30113786,
        0.02331892, -0.16694683, -0.11898861, -0.11155669, -0.11140025,
       -0.04372848, -0.05509884,  0.01517171, -0.08260392, -0.00161575,
       -0.02493379, -0.07467234, -0.04603147, -0.05900525, -0.03726976,
       -0.0679428 , -0.02658186, -0.03484896, -0.05054107,  0.01181231,
       -0.05419018, -0.02789435, -0.07649779, -0.06745438, -0.06795349,
       -0.01358283, -0.05810555, -0.02102332, -0.04091351, -0.04357026,
       -0.05378678, -0.0284171 , -0.05640033, -0.02927306, -0.02846522,
       -0.03177663, -0.07638131, -0.05611386, -0.05489119, -0.05565432,
       -0.04258532, -0.05160356, -0.02606789, -0.10892859, -0.0801097 ,
       -0.02651254, -0.06652889, -0.02263256, -0.12240582, -0.07451186,
       -0.01181235, -0.04571755, -0.05589939, -0.03106802, -0.02310427,
       -0.03846722, -0.04033899, -0.03515457, -0.04214652, -0.02492795,
       -0.09590906, -0.0186641 , -0.04982725, -0.07625089, -0.03205988,
       -0.03157685, -0.019672  , -0.02402663, -0.02628564, -0.06359492,
       -0.03335344, -0.03502774, -0.01114108, -0.0183523 , -0.03329939,
       -0.03672545, -0.0106354 , -0.03627934, -0.01517804, -0.05017845,
       -0.03181882, -0.07891283, -0.00914016, -0.03660728, -0.02326647,
       -0.07230679, -0.09763225, -0.03193749, -0.03265606, -0.08507845,
       -0.03619653, -0.04790005, -0.02187832, -0.04253055, -0.0476447 ,
       -0.0540627 , -0.0231773 , -0.01968132, -0.03333442, -0.02816228,
       -0.02027046, -0.02613658, -0.03812622,  0.00281948, -0.03071311,
       -0.02817434, -0.01031966, -0.03100957, -0.01505435, -0.01726348,
       -0.02691684, -0.03011565, -0.01702471, -0.0414536 , -0.06129844,
       -0.01651873, -0.03135057, -0.04340168, -0.00092003, -0.04928182,
       -0.034248  , -0.09974355, -0.01865924, -0.04439797, -0.01476688,
       -0.03306961, -0.02885992, -0.02419128, -0.01934658, -0.0165337 ,
       -0.03888645, -0.03357549, -0.03318277, -0.03605353, -0.02205066,
       -0.02208897, -0.02288036, -0.02478068, -0.02645998, -0.02701743,
       -0.01061509, -0.02917154, -0.02413019, -0.01670913, -0.00947745,
       -0.02135734, -0.04397395, -0.03449064, -0.03994662, -0.02139767,
       -0.03826213, -0.01435881, -0.04364398, -0.02134804, -0.01897896,
       -0.03825948, -0.06159144, -0.07000342, -0.06397267, -0.049531  ,
       -0.00403947, -0.02165901, -0.03921867, -0.03585689, -0.0326954 ,
       -0.09348491, -0.04602918, -0.04709668, -0.0596785 , -0.00684082,
       -0.07048104, -0.04772301, -0.04529928, -0.0381281 , -0.03512678,
       -0.03788008, -0.03406602, -0.0664497 , -0.07919781, -0.11114956,
       -0.0544449 , -0.0159083 , -0.06630901, -0.024242  , -0.03079529,
       -0.0334688 , -0.0693963 , -0.02312607, -0.01792446, -0.06848365,
```

Since all the values at risk are in negative, there is a higher probability of the portfolio returning more than invested amount.both the approaches return negative values (indicating more probability of higher returns), there is still a difference in the figures of values which implies that the distribution is not normal.

Var and CVaR for each stock: -

```
def historicalVaR(Health_portfolio_returns, alpha=1):
    """
    Read in a pandas dataframe of returns / a pandas series of returns.
    Output the percentile of the distribution at the given alpha confidence level.
    """
    if isinstance(Health_portfolio_returns, pd.Series):
        return np.percentile(Health_portfolio_returns, alpha)

    # A passed user-defined-function will be passed a Series for evaluation.

    elif isinstance(Health_portfolio_returns, pd.DataFrame):
        return Health_portfolio_returns.aggregate(historicalVaR, alpha=alpha)

    else:
        raise TypeError("Expected returns to be dataframe or series")


def historicalCVaR(Health_portfolio_returns, alpha=1):
    """
    Read in a pandas dataframe of returns / a pandas series of returns
    Output the CVaR for dataframe / series
    """
    if isinstance(Health_portfolio_returns, pd.Series):
        belowVaR = Health_portfolio_returns <= historicalVaR(Health_portfolio_returns, alpha=alpha)
        return Health_portfolio_returns[belowVaR].mean()

    # A passed user-defined-function will be passed a Series for evaluation.

    elif isinstance(Health_portfolio_returns, pd.DataFrame):
        return Health_portfolio_returns.aggregate(historicalCVaR, alpha=alpha)

    else:
        raise TypeError("Expected returns to be dataframe or series")
```

```
hist_var_99
```

```
ABT      -0.068116
AZN      -0.065341
DHR      -0.059271
ELV      -0.089358
GSK      -0.060808
JNJ      -0.057483
LLY      -0.070783
MRK      -0.057600
MRNA     -0.130128
NVO      -0.050583
NVS      -0.059367
PFE      -0.070724
RHHBY    -0.065860
UNH      -0.088479
dtype: float64
```

```
historicalCVaR(Health_portfolio_returns, alpha=1)
```

```
ABT      -0.087094
AZN      -0.090004
DHR      -0.080866
ELV      -0.131103
GSK      -0.088151
JNJ      -0.067022
LLY      -0.083996
MRK      -0.073022
MRNA     -0.158414
NVO      -0.067212
NVS      -0.082852
PFE      -0.073924
RHHBY    -0.078829
UNH      -0.127504
dtype: float64
```

Forward Testing 2021

```
rp_weights1 = ([0.023511898836872068,0.07050913770115579,0.08232929734691172,0.08122708233699079,0.13219320684750077,
               0.08323598156188067,0.06355975804407532,0.04957027087149978,0.0800234946443608,0.06580449716060635,
               0.07430202655109787,0.07770389612300004,0.03261413910779822,0.08341531286624768])
ew_portfolio_return_2021 = Health_portfolio_2021.dot(hrp_weights1)
```

```
new_portfolio_return_2021
```

```
Date
2020-01-02    114.967645
2020-01-03    113.862702
2020-01-06    114.379780
2020-01-07    114.099131
2020-01-08    115.376192
                 ...
2020-12-23    134.931139
2020-12-24    134.861244
2020-12-28    135.142296
2020-12-29    136.103125
2020-12-30    135.663714
Length: 252, dtype: float64
```

```
hrp.portfolio_performance(verbose=True)
print(dict(hrp_weights))
```

```
Expected annual return: 17.4%
Annual volatility: 28.8%
Sharpe Ratio: 0.53
{'ABT': 0.049555789146461665, 'AZN': 0.07301545109813028, 'DHR': 0.07885228010540335, 'ELV': 0.053989668059532266,
'GSK': 0.07767268602777815, 'JNJ': 0.08232453274207821, 'LLY': 0.08872374904903048, 'MRK': 0.1323892087232456, 'MRN
A': 0.023918583283296924, 'NVO': 0.07999376339460092, 'NVS': 0.0833487382371751, 'PFE': 0.0742599871887004, 'RHHBY'
: 0.06350891689831022, 'UNH': 0.038446646046256444}
```

Using the HRP model it almost mainted the consisitency for 2020 and 2021 in terms of annual return percent. So my conclusion was right for log term hold HRP model is appropriate.

VaR99 Hypothesis Testing Using Monte Carlo Simulation

You may use the following procedures to figure out a portfolio's VaR:

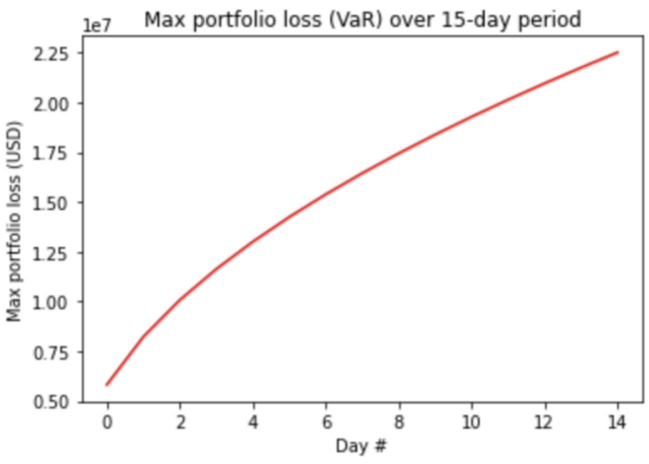
Calculate the portfolio's stocks' periodic returns. Make a covariance matrix using the returns as a basis. Do the portfolio's mean and standard deviation calculations (weighted based on investment levels of each stock in portfolio) With the given confidence interval, standard deviation, and mean, determine the inverse of the normal cumulative distribution (PPF). Determine the portfolio's value at risk (VaR) by deducting the initial investment from the calculation in step (4)

```
#Finally, we can calculate the VaR at our confidence interval
var_1d1 = initial_investment - cutoff1
var_1d1

#output
#22347.7792230231
```

```
5814704.822333044
```

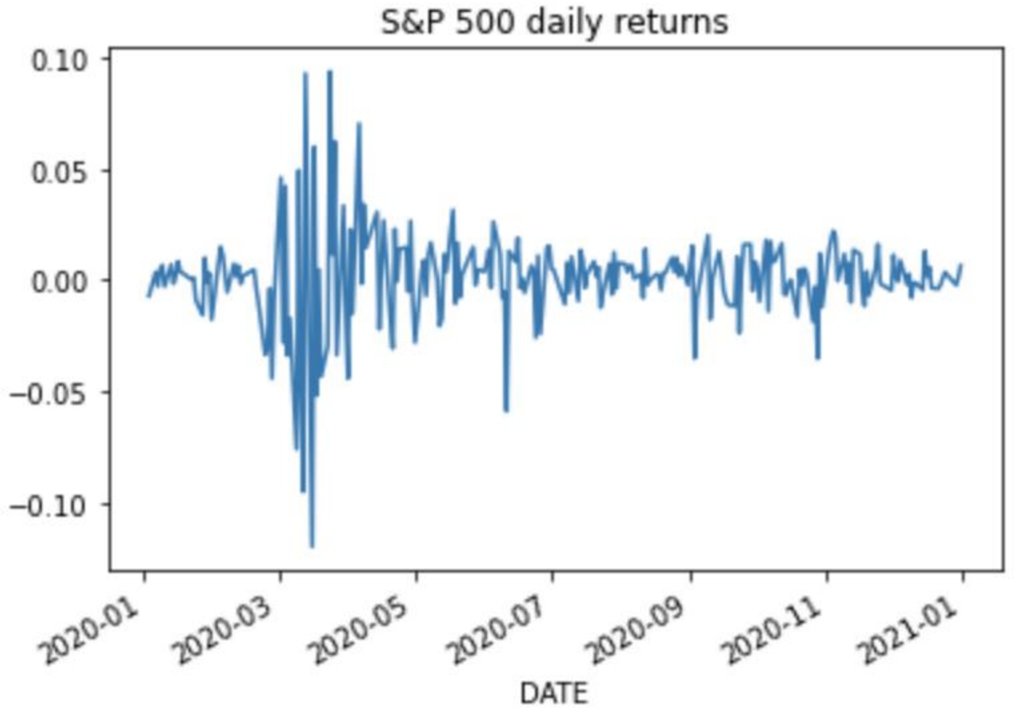
```
# Calculate n Day VaR
var_array = []
num_days = int(15)
for x in range(1, num_days+1):
    var_array.append(np.round(var_1d1 * np.sqrt(x),2))
    print(str(x) + " day VaR @ 99% confidence: " + str(np.round(var_1d1 * np.sqrt(x),2)))

# Build plot
plt.xlabel("Day #")
plt.ylabel("Max portfolio loss (USD)")
plt.title("Max portfolio loss (VaR) over 15-day period")
plt.plot(var_array, "r")
```

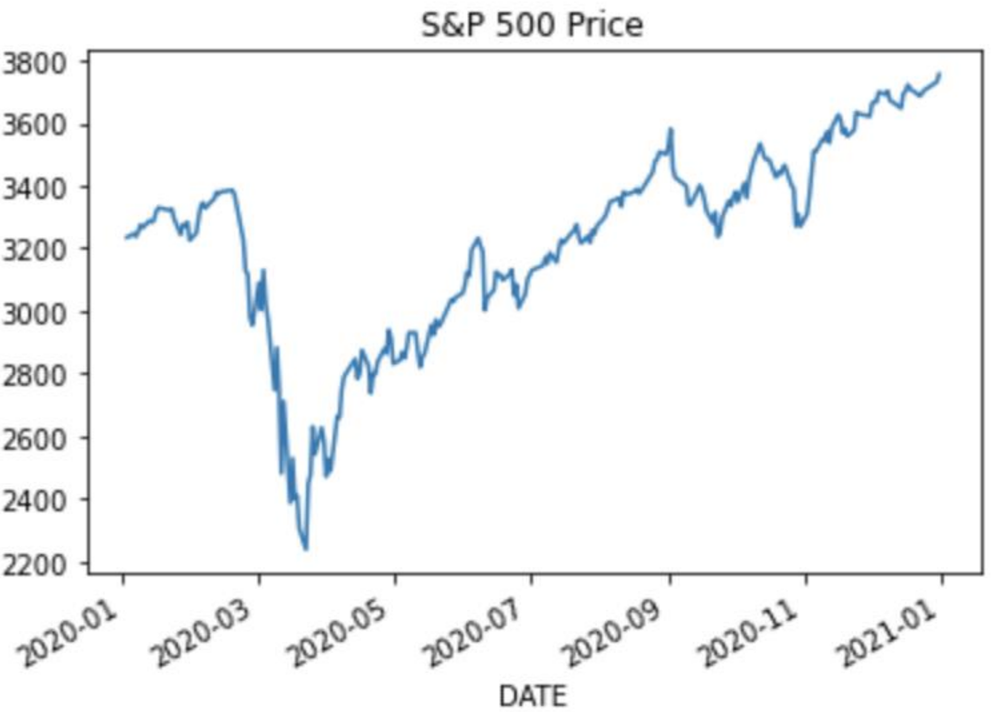
```
1 day VaR @ 99% confidence: 5814704.82
2 day VaR @ 99% confidence: 8223234.42
3 day VaR @ 99% confidence: 10071364.18
4 day VaR @ 99% confidence: 11629409.64
5 day VaR @ 99% confidence: 13002075.25
6 day VaR @ 99% confidence: 14243059.82
7 day VaR @ 99% confidence: 15384262.91
8 day VaR @ 99% confidence: 16446468.84
9 day VaR @ 99% confidence: 17444114.47
10 day VaR @ 99% confidence: 18387711.16
11 day VaR @ 99% confidence: 19285194.16
12 day VaR @ 99% confidence: 20142728.37
13 day VaR @ 99% confidence: 20965216.39
14 day VaR @ 99% confidence: 21756633.25
15 day VaR @ 99% confidence: 22520254.94
```



Comparison with S&P 500

Visualization of S&P500 Daily Return: -





Annual Return of S&P 500

```
hrp.portfolio_performance(verbose=True)
print(dict(hrp_weights2))
```

```
Expected annual return: 21.4%
Annual volatility: 34.9%
Sharpe Ratio: 0.56
{'daily_return': 3.5281105819073844e-06, 'sp500': 0.9999964718894181}
```

RISK FACTORS

The risk factors may be divided into numerous categories, including, Risk Elements Related to the Healthcare sector and stocks chosen. Those risks can be: -

1.) Risk Elements Affecting the Healthcare

- Operational
- Clinical & Patient Safety
- Strategic
- Financial
- Human Capital
- Legal & Regulatory
- Technological
- Environmental- and Infrastructure-Based Hazards.

Other risks I saw in the business report on MRNA include supply chain problems and significant investment in the technology as they attempt to develop a cancer vaccine utilizing it. Since mRNA is less effective and has more negative side effects, it might result in enormous loss and failure. The clinical and administrative systems, procedures, and reports used in risk management in the healthcare industry include those for risk detection, monitoring, assessment, mitigation, and prevention. Healthcare companies that use risk management proactively and methodically protect patient safety as well as their assets, market share, accreditation, levels of reimbursement, brand value, and reputation in the community.

Conclusions

Compared to other industries, the healthcare industry offers one of the lowest risk environments for investment. I want to anticipate stocks using these two methodologies for other industries like energy and IT before comparing them to the healthcare industry to reach a decision. As can be seen, the inflated 225 percent that we obtained using mean variance optimization is far lower than the anticipated yearly return of 17.9 percent that we had. A lower Sharpe ratio of 0.55 is also seen. This result is far more rational and more likely to persist over time since HRP is less prone to outliers than mean variance optimization. Our analysis leads us to the conclusion that we should create a portfolio consisting of 112 shares of MRNA, 183 shares of DHR, 94 shares of UNH, 846 shares of MRK, 441 shares of NVS, 247 shares of LLY, 262 shares of JNJ, 183 shares of DHR, 572 shares of NVO, 1055 shares of GSK, 743 shares of AZN, 957 shares of PFE, 725 shares of RHHBY,

References

1. *An Introduction to Portfolio Optimization in Python*. (n.d.). Built In. Retrieved December 18, 2022, from https://builtin.com/data-science/portfolio-optimization-python

2. Treece, D. D. (2022, December 1). *Best Healthcare Stocks of 2022*. Forbes Advisor. Retrieved December 18, 2022, from https://www.forbes.com/advisor/investing/best-healthcare-stocks/

3. *Calculating Value at Risk (VaR) of a stock portfolio using Python - InterviewQs*. (n.d.). Calculating Value at Risk (VaR) of a Stock Portfolio Using Python - InterviewQs. Retrieved December 18, 2022, from https://www.interviewqs.com/blog/value-at-risk

4. *Value at Risk (VaR) Calculation in Excel and Python*. (2022, December 9). Quantitative Finance & Algo Trading Blog by QuantInsti. Retrieved December 18, 2022, from https://blog.quantinsti.com/calculating-value-at-risk-in-excel-python/

5. *How to calculate in a pandas dataframe each day's Value at risk in rolling window manner*. (2022, August 17). Stack Overflow. Retrieved December 18, 2022, from https://stackoverflow.com/questions/73384021/how-to-calculate-in-a-pandas-dataframe-each-days-value-at-risk-in-rolling-windo

6. *Backtesting a Forecasting Strategy for the S&P500 in Python with pandas | QuantStart*. (n.d.). Backtesting a Forecasting Strategy for the S&P500 in Python With Pandas | QuantStart. Retrieved December 18, 2022, from https://www.quantstart.com/articles/Backtesting-a-Forecasting-Strategy-for-the-SP500-in-Python-with-pandas/

7. Boller, K. (2018, April 16). *Python for Finance: Stock Portfolio Analyses*. Medium. Retrieved December 18, 2022, from https://towardsdatascience.com/python-for-finance-stock-portfolio-analyses-6da4c3e61054

8. *https://catalyst.nejm.org/doi/full/10.1056/CAT.18.0197*. (n.d.). Retrieved December 18, 2022, from https://catalyst.nejm.org/doi/full/10.1056/CAT.18.0197