Portfolio Construction Algorithm for US Large Cap Equities

Authors: *Adil Gazder, Chris Moreira, Danish Maknojia, Vishesh Gupta*

Date: 12/11/2024

Affiliation: Duke University - Fundamentals of Finance Business Models

Acknowledgments

We would like to express our gratitude to Rakeen Rouf and Professor David Ye for their invaluable contributions to this project. Their insightful guidance during the project ideation phase and their keen ability to identify key avenues for improvement were instrumental in enhancing the performance of our stock portfolio algorithm. Their expertise and support were essential to the success of this work, and we are deeply thankful for their time and commitment.

Abstract

This project focuses on developing an investment portfolio optimization algorithm that iteratively selects stocks with the goal of maximizing Total Shareholder Return (TSR), optimizing volatility through equity beta, and determining stock allocations based on the equity ratings of each security analyzed. The final portfolio consists of 80 large-cap equities representing 8 distinct industries. Designed to maintain volatility lower than that of the S&P 500, the portfolio also incorporates algorithmic constraints to cap individual stock and industry weightings, thereby preserving diversification.

Background & Problem Statement

Algorithmic portfolio management has gained significant traction among asset managers in recent years due to its numerous advantages, including the ability to execute asset rebalancing efficiently and at lower costs compared to traditional, human-intensive strategies. Unlike conventional approaches that are often swayed by market sentiment, algorithmic portfolios rely on data-driven processes and pre-established constraints. Moreover, the advent of generative AI has greatly simplified the construction, testing, debugging, and refinement of these algorithms. The availability of real-time market data and the capability for firms to trade based on near-instantaneous information further underscore the growing reliance on algorithmic systems to respond dynamically to various market conditions.

This project explores the process of creating an algorithm capable of constructing a large-cap equity portfolio. The algorithm operates on a set of equities that represent the leading industries typically found in U.S. large-cap ETFs. The selection of stocks and industries draws inspiration from industry-specific ETFs managed by the U.S.'s largest asset managers, including Vanguard, BlackRock, and State Street.

Equity Variables

For each stock included in the analysis, Total Shareholder Return (TSR) serves as the key performance metric. TSR is a comprehensive indicator that accounts for both capital appreciation and dividend yield, thus providing a holistic view of a company's performance over a given period. The inclusion of both price appreciation and dividend yield is particularly important, as certain industries

exhibit a stronger tendency to distribute dividends than others. For example, in 2024, the Vanguard Utilities ETF (VPU) provided a dividend yield of approximately 3.2%, which is five times higher than that of the Vanguard Information Technology ETF (VGT). Conversely, the technology sector typically delivers returns through price appreciation rather than dividend payouts. Nevertheless, individual companies within any given sector may show distinct preferences for either dividends or reinvesting earnings. Therefore, TSR was chosen as the optimal metric to capture both the capital growth and income-generating components of stock performance, aligning with the goal of maximizing overall returns while balancing the risk associated with market volatility.

The second key variable is the company's beta, a measure of a stock's volatility relative to the broader market. Beta values above 1 indicate that a stock's price fluctuates more dramatically than the market, while stocks with betas below 1 experience less volatility. The algorithm developed for this project is designed to minimize the portfolio's overall beta, a common strategy for risk-averse investors seeking stable returns.

The final variable consists of equity ratings assigned to each of the 80 stocks selected for inclusion in the portfolio. These ratings, ranging from A to F (with A being the highest), are based on a variety of financial indicators, including cash flows, debt-to-market capitalization ratio, and price-to-earnings ratio, among others. These ratings, generated by human analysts rather than algorithms, provide an assessment of a stock's relative value and potential for future growth.The ideal portfolio leverages these 3 variables and minimizes for betas, optimizes for TSR and has higher allocation towards companies with higher equity ratings. See below figure 1.1 of this criterion, where the red dots represent what would be deemed as optimal portfolios based on TSR, beta and equity ratings alone. This article will further discuss how a portfolio should not only optimize for these metrics but for industry and single stock asset allocations.

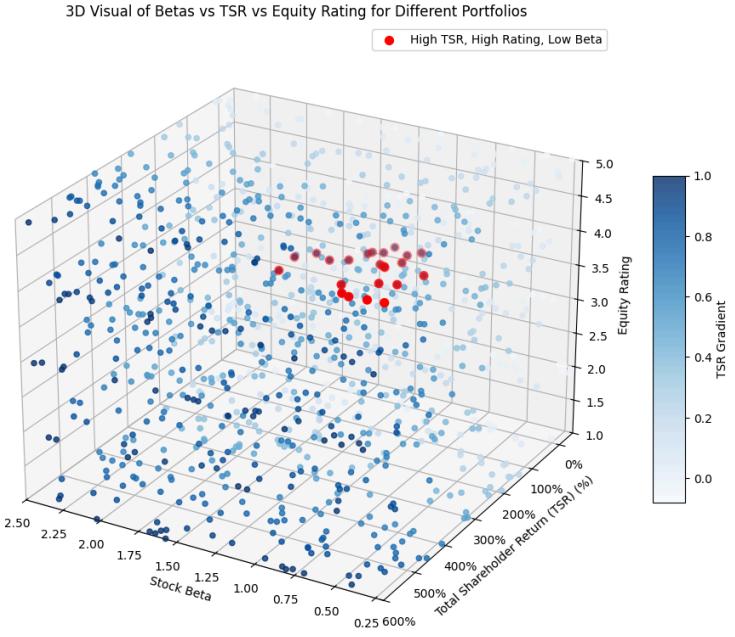


Figure 1.1 - Equity Rating VS TSR VS Beta 3-dimensional visual

Industry Diversification in Portfolios

The backbone of any well-constructed ETF is its ability to diversify risk across multiple industries. Certain sectors are traditionally less sensitive to market fluctuations, making them valuable components for balancing overall portfolio risk. For example, Vanguard's Utility ETF is notorious for its low correlation with other sector-specific ETFs, such as those representing technology, energy, and financials. This suggests that utility stocks can act as a hedge against volatility in more cyclical sectors. Similarly, industries such as consumer staples and healthcare are traditionally seen as defensive sectors that perform well during economic downturns, given their provision of essential goods and services that consumers continue to purchase regardless of market conditions. Below is an illustration(Figure 1.2) of how a consumer staples ETF exhibits less volatility than the market, yet modest returns. Meanwhile, a Growth ETF(VUG) is highlighted by higher returns while concomitantly demonstrating higher volatility patterns in relation to VDC and the S&P 500 market index.

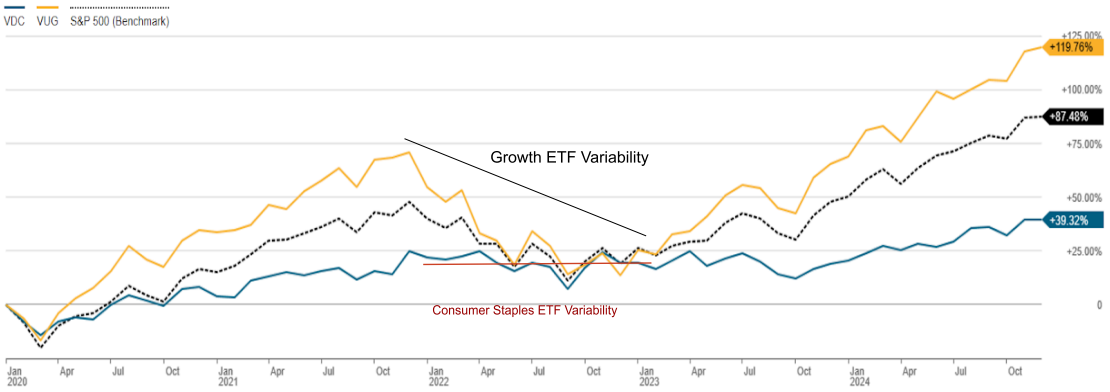


Figure 1.2 - Consumer Staples Vs Growth Vs S&P 500. Chart Produced using Charles Schwab Research Tools https://client.schwab.com/app/research/#/tools/etfs/etfcompare *(December 2024)*

However, a potential drawback of sectors like utilities, healthcare, and consumer staples is their relative underperformance during bull markets, where growth in industries like technology and industrials tends to outpace that of defensive sectors. To mitigate this, the algorithm in this project ensures that the portfolio includes a blend of stocks from both cyclical and defensive industries, enabling it to thrive in a variety of market environments. By applying constraints to the weightings of both individual stocks and industries, the algorithm aims to create a portfolio that not only performs well during periods of market growth but also offers resilience during times of higher volatility. Figure 1.3 highlights how the correlation coefficients vary by different market sectors, where we analyze industry-specific ETFs from Vanguard in relation to one another; These coefficients further underscore the advantages obtained in industry diversification.

Name	Ticker	VHT	VDC	VIS	VPU	VNQ	VDE	VGT	VFH	VOO
Vanguard Health Care ETF	VHT	1.00	0.66	0.73	0.41	0.63	0.43	0.66	0.64	0.80
Vanguard Consumer Staples ETF	VDC	0.66	1.00	0.70	0.64	0.68	0.42	0.58	0.59	0.75
Vanguard Industrials ETF	VIS	0.73	0.70	1.00	0.43	0.68	0.66	0.77	0.89	0.92
Vanguard Utilities ETF	VPU	0.41	0.64	0.43	1.00	0.69	0.21	0.34	0.35	0.49
Vanguard Real Estate ETF	VNQ	0.63	0.68	0.68	0.69	1.00	0.40	0.59	0.62	0.73
Vanguard Energy ETF	VDE	0.43	0.42	0.66	0.21	0.40	1.00	0.43	0.70	0.62
Vanguard Information Technology ETF	VGT	0.66	0.58	0.77	0.34	0.59	0.43	1.00	0.66	0.91
Vanguard Financials ETF	VFH	0.64	0.59	0.89	0.35	0.62	0.70	0.66	1.00	0.84
Vanguard S&P 500 ETF	VOO	0.80	0.75	0.92	0.49	0.73	0.62	0.91	0.84	1.00

Asset correlations for time period 03/01/2013 - 11/30/2024 based on monthly returns

Figure 1.3 - Industry & Market Index Correlations among Industry-Specific Vanguard ETFs & S&P500. Produced by Portfolio Visualizer, https://www.portfoliovisualizer.com/asset-correlations (December 2024)

Factors

Our quantitative research began with a thorough fundamental analysis of a selection of 80 companies focused on identifying future growth in revenues, avenues for increasing revenue, and risk factors that can hinder growth. In an effort to increase diversification and reduce sector-associated risks our algorithm will have a minimum exposure of 4.5% to a sector and a max of 16% exposure.

Using data collected from macrotrends.net and the Bloomberg terminal. The first step taken was an evaluation of the company's revenue projections and historical quarterly revenues to understand growth and strength of current business centers, diversity of revenue streams, and ability to scale the business. For example, identifying which product or service generates the bulk of revenue and assessing its market sustainability is crucial. Complementary streams should be analyzed for their contribution to overall revenue and their scalability. Comparing revenue distribution across segments highlights whether the company is overly reliant on a single income source or effectively diversified.

Next, a thorough examination of expenses helps with understanding the cost structure and profitability. This involves calculating gross and operating margins for each revenue stream. High-margin segments are often indicative of the company's competitive advantage, while lower-margin ones may signal areas for cost optimization. Operating expenses, including research and development, marketing, and administrative costs, should be compared to industry norms to evaluate operational efficiency.

Profitability is then assessed through key metrics such as net income, earnings before interest and taxes (EBIT), and profit margins. Tax impacts are considered to determine net income after tax. Earnings per share (EPS) is a vital metric that links net income to individual shares, providing insight into the value delivered to shareholders.

The final step is valuation. Price-to-earnings (P/E) ratios, forward P/E, and other metrics like price-to-book (P/B) or enterprise value-to-EBITDA (EV/EBITDA) are used to compare the stock against industry peers and historical performance. Forecasted metrics, such as future EPS and P/E ratios, help estimate the stock's potential price, guiding investment decisions. Calculating compounded annual return based on projected price growth provides a sense of the investment's attractiveness over time.

Initial Algorithm

In the landscape of investment management, having an optimal portfolio that balances our returns with risk is paramount to achieving success. The foundation of a portfolio optimization process begins with understanding and acquiring the necessary company data. Once we understood the above for the 80 companies across our selected sectors we wanted to focus our portfolio on, we cleansed the data to focus on the specific attributes we deemed critical for our investment decision which were:

- **Grade:** This was a categorical assessment ranging from A to F that evaluated the quality or performance of the company. The evaluation incorporated key financial metrics such as P/E ratio, P/B ratio, dividends and yields, and more. Beyond these metrics, each team member did a comprehensive quantitative analysis around their companies to help finalise the grade.
- **Beta Value:** This is a numerical indicator that represents the volatility or system risk of a company's stock relative to the broader market. The S&P500 usually has a beta of 1.0 and stocks with a beta value higher than that are more volatile than the S&P500.
- **Total Shareholder Return (TSR):** A percentage measure reflecting the return generated for shareholders, incorporating both capital gains and dividends.

Since our grades were defined as A-F to translate this qualitative assessment into a quantitative metric our algorithm first worked on providing a specific weight to each grade. For instance, grade "A" corresponded to a higher weight i.e. 90% while an "F" grade was associated with a lower weight i.e. 40%. Through this process, we ensure that the higher-grade companies extract more influence in our algorithm decisions while working on building an optimal portfolio.

To further ensure the comparability and proportional representation of the companies selected by us we decided to normalize the TSR values for our companies across their industry. This normalization process adjusted the TSR values so that they sum to 1 within their respective industries, thereby scaling individual TSR contributions relative to their sector peers.

Once we had our data ready we worked on the scoring mechanism which was a pivotal pillar in our algorithm. This Grade Score combines our weighted grade with the company's TSR. The Grade score was calculated using the following formula:

$$\text{Grade Score} = ((1 - y) * \text{TSR}) + (y * \text{Grade Weight})$$

Here, y is a randomly generated weight parameter that ranges from between 0.65 to 0.8, This variability introduces stochasticity into the scoring process, allowing the algorithm to explore a diverse array of scoring scenarios while always ensuring that the Grade Weight predominantly influences the final Grade Score.

Recognising the importance of diversification in a portfolio our algorithm we decided to have a few constraints within our code:

- **Maximum Industry Allocation Limit**: No single industry can constitute more than 16% of the total portfolio. To enforce this constraint the algorithm evaluated the cumulative weight of each industry within the portfolio. If an industry's allocation exceeds the stipulated maximum, it is capped to 16%. The normalization process was then reapplied to adjust the weights of the affected companies, maintaining the overall portfolio balance.
- **Portfolio Beta:** The algorithm must ensure that our final investment strategy to obtain our optimized portfolio does not cross a beta value of 1.0 thereby ensuring that the portfolio does not exhibit excessive market volatility. In our algorithm, the portfolio beta is calculated as the weighted average of individual company betas, which reflects the portfolio's sensitivity to market movements.

Algorithm Framework

1. **Iteration Process:** We conduct a predefined number of iterations each representing a unique weight allocation and grade score computation thereby giving us unique answers on each run
2. **Random Weight Parameter Generation:** In each iteration a new random variable y was generated introducing variability in the calculation
3. **Grade Computation:** Grade scores were calculated for each company reflecting the balance between TSR and Grade Weight
4. **Weight Normalisation and Allocation:** Investment weights are recalibrated based on the new Grade Scores, followed by normalization to adhere to industry and stock allocation constraints.
5. **Portfolio Beta Calculation**
6. **Optimal Portfolio Selection:** If the computed portfolio beta meets the target criteria set i.e. below 1.0, the algorithm identifies this portfolio as the optimal solution and terminates the simulation.

On completing the iterations, the algorithm produces an ideal portfolio configuration that best aligns with the defined risk parameters and constraints. While the current implementation provides a solid foundation for portfolio optimization, we are aware that this is only an initial step toward developing a more refined and resilient algorithm. This foundational approach highlights the potential to achieve effective portfolio results while acknowledging room for further improvement.

Improved Algorithm

With our initial framework in place, we set out to improve the model characteristics to ensure it delivers optimal performance and risk mitigation through diversification and minimal

over-indexing. In line with the principle of Modern Portfolio Theory (MPT), we try to optimize for high returns for a given level of risk. We worked on achieving this by refining our algorithm as detailed below:

- **Share weightage restrictions**: To ensure no over-exposure to a particular stock and adequate diversification to address market volatility, we chose to limit the maximum weightage of any stock to not more than 4.5% of the total portfolio. We implemented this by iterating through each industry and implementing an upper cap for each stock with a weightage of more than 4.5% from the initial algorithm. We then normalized and redistributed the weightage amongst the rest of the stocks in the same industry. This way, we ensured minimal idiosyncratic risk for a particular stock and diversified this risk across the entire sector.

- **Sectoral weightage restrictions:** To ensure we are able to mitigate sector-specific risks, we also ensured that as a result of our optimization algorithm, the weightage of each sector does not exceed more than 16% of the entire portfolio. This was especially crucial to deal with volatility and balancing economic cycles.

- **Objective function optimization:** The objection function used as the basis of our portfolio was defined as the overall 'Grade Score' (represented by 'y' in the equation above) of the algorithm. We initially determined the value of 'y' in this function by running a pseudorandom algorithm to generate a value between 0.65 and 0.8. We realized that with this there may be more optimal values that exist between 0.65 to 0.8 which the algorithm may not be considering as is the nature of the pseudorandom function. To overcome this limitation, we iterated over 15000 values between 0.65 and 0.8, ran the model for each of these scenarios, and generated the portfolio with the lowest beta as a deterministic output from our algorithm.

Sector *(SPDR ETF Ticker)*	**Weightage** *(As % of total portfolio)*
Consumer Staples (XLP)	11%
Energy (XLE)	13%
Finance (XLF)	15%
Healthcare (XLV)	14%
Industrials (XLI)	13%
Real Estates (XLRE)	9%
Technology (XLK)	15%
Utilities (XLU)	10%

Table: Sectoral weight allocation in our portfolio

As a result of these improvements, our resulting portfolio showed improved returns and reduced overall volatility (measured by Beta value) of 0.89 (~5% reduction in Beta value). Further results of the model are detailed below.

Model Results

To verify our portfolio's performance, we wanted to look at the returns it delivers as compared to individual sectoral indices (using Sector SPDR ETFs) and along with the S&P 500 ETF (Ticker: SPY). This would give us a good idea on how our portfolio would perform as compared to these benchmarks.

To implement this, we tracked returns across the past 30 days (indexed on 4th December 2024) and calculated returns as a percentage of total returns from the day of investment (11th November 2024) for recency. We used Yahoo Finance's package to get the OHLC price for each stock in our portfolio and indexed against the closing price for each day. With this, we then took the weighted average of each stock and determined the model's overall performance.

- **Sectoral performance:** To initially compare against the performance of individual sectors, we tracked the performance of each of our sectors under consideration ie (Finance, Technology, Energy, Utilities, Consumer Staples, Healthcare, Industries, and Real Estate) over the same time period (11/05/2024 - 12/04/2024). During this period, the return on investment (unweighted) averaged 5.57% with XLF (Finance) showing the highest return of 9.01% and XLV (Healthcare) showing the lowest return of 0.03%. The SPY's performance during this time period has also been plotted for reference. Compared to these, our portfolio gave a return of 10.08% outperforming all 8 out of the 8 industries under consideration. We attribute this to the diversification in our portfolio which mitigates sector-specific risk.

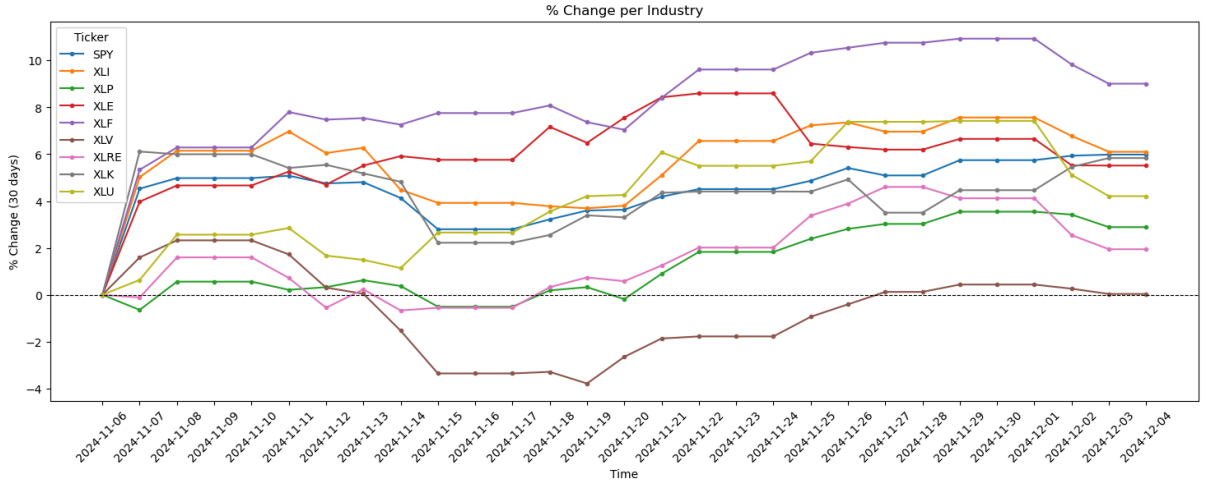


Figure 2.1: Sectoral overview (Total Returns) for the past 30 days

- **SPY performance:** Similar to the sectoral performance, we measured the returns of our portfolio compared to the SPDR S&P 500 ETF (considered a benchmark for the entire US Stock Market). During the same time period (11/05/2024 - 12/04/2024), the SPY gave returns of 5.98% as compared to our portfolio which gave returns of

10.08%. We attribute this to the iterative weightage process we followed to determine the weights for the objective function.

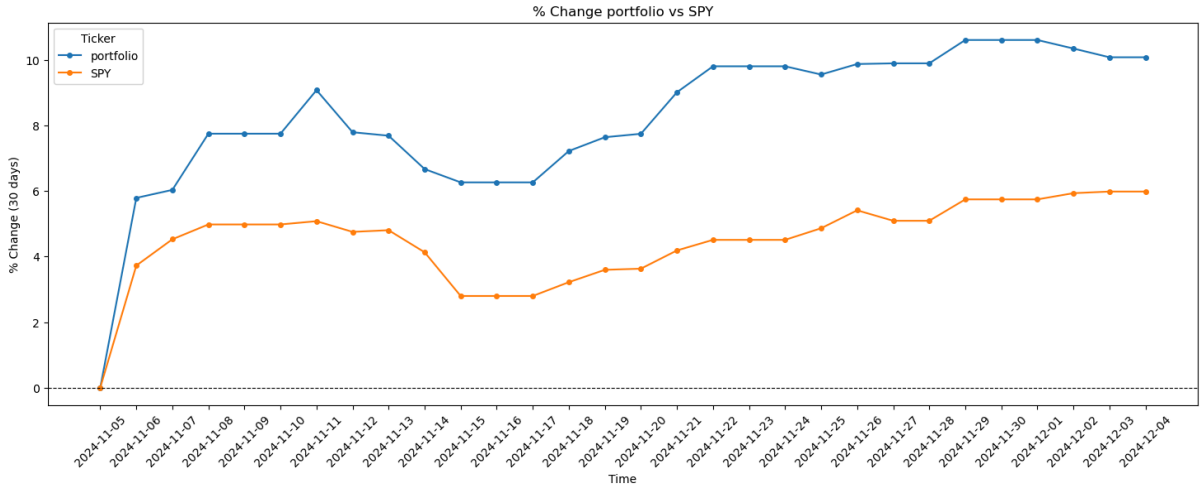
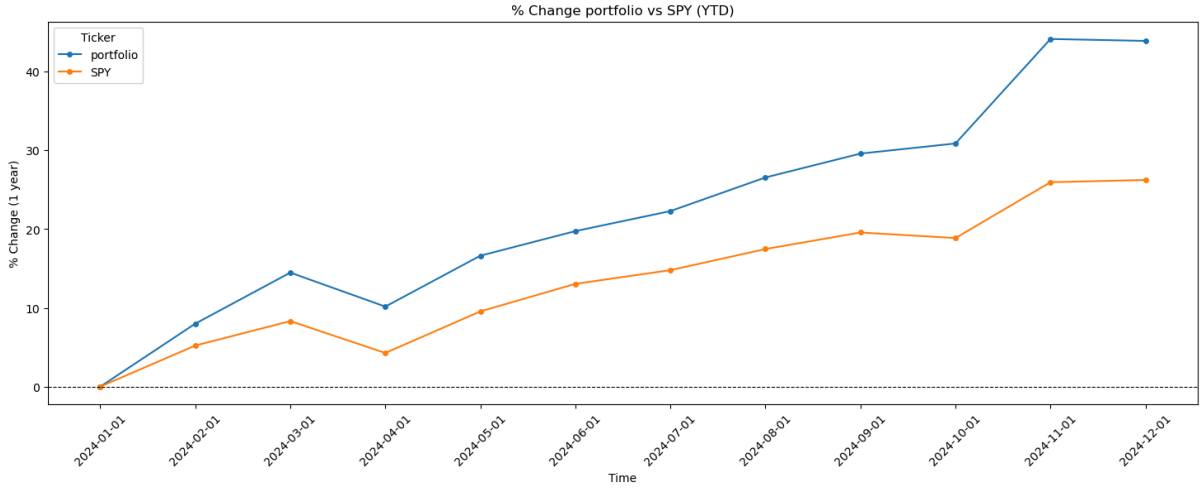


Figure 2.2: Portfolio vs SPY (Total Returns) for the past 30 days

- **YTD performance:** We also compared the performance of the portfolio to the SPY over the YTD 2024 time period and saw that our portfolio (44% overall returns) out-performed the SPY (26% overall returns) with a mean monthly return (calculated MoM) as 4% and a standard deviation of 4.6%



Future Enhancements

Looking ahead, we would aim to fine-tune this portfolio to ensure it can deliver alpha returns in all-weather market conditions ensuring our portfolio returns are immune to economic and seasonal volatility. We would try to achieve this by implementing a dynamic weightage allocation model that would be able to handle seasonal market changes, ensuring that under any economic condition, with minimal risk and higher than average returns, is able to design portfolios to outperform benchmark indices and ETFs. To enhance the portfolio's accuracy, we would also incorporate sentiment analysis to gather data from public data forums for fundamental analysis and sectoral grading.

Data Source

The data for grading stocks was sourced from the Bloomberg Terminal and Macrotrends.net. Bloomberg provided comprehensive and reliable financial data, while Macrotrends offered historical metrics and ratios. However, there were quality issues, including inconsistencies in Macrotrends, specifically the time intervals measured for metrics such as TSR (1yr vs 5yr). To address these, data from Macrotrends was double-checked with the Bloomberg Terminal which functions as the gold standard for financial data. For backtesting, data was obtained using the Yahoo Finance Python package, which provided historical stock prices and related financial metrics. Challenges included missing price data for tickers. Cleaning efforts focused on ensuring we had the correct ticker symbols for each stock and making sure the time series matched across the portfolio. These steps ensured the data was reliable and suitable for analysis.

Conclusion

With our knowledge of the financial markets, the various underlying risk factors, and the importance of diversification, we designed an algorithm which would be able to deliver alpha returns (returns above benchmark returns) with minimal risk (measured as the Beta value). To accomplish this, we started with our stock pool selection, selecting 80 companies across 8 different sectors and performing in-depth fundamental analysis on these stocks, and assigning a grade indicating the strength of these fundamentals. We then combined this parameter along with the Total Shareholder Return (TSR) value for each stock to determine our objective function which would assign weights to each of these factors in the resulting portfolio. We iterated over weight-values by restricting the range to give more weightage to the fundamental grade score to ensure that our portfolio does not just rely on technical factors to determine its performance. As a result, we observe that our portfolio was able to beat both sectoral indexes and the SPY over a 30 day time period. Due to adequate industry diversification and minimal over-indexing on certain stocks, we were able to successfully identify risk across the portfolio, with an overall beta value of 0.89.

Duke University
IDS 789 (Fall 2024)
Fundamentals of Finance Business Models for students in Math and Data Science

COURSE PROJECT

Non - Technical Project: *The skills, roles and responsibilities needed as Quantitative Research Intern with Two Sigma [Summer 2025]*

Adil Keku Gazder
adil.gazder@duke.edu

Chris Moreira
chris.moreira@duke.edu

Danish Maknojia
d.maknojia@duke.edu

Vishesh Gupta
vg157@duke.edu

COMPANY: TWO SIGMA
POSITION: Quantitative Researcher - Internship [2025 Summer]

About Two Sigma

Two Sigma is a financial science company that was founded in 2001 by David Diegel and John Overdeck. The company is a hedge fund that focuses on combining advanced technology and data science with human inquisitiveness to solve tough challenges in finance. Two Sigma aims to produce an alpha for its clients through varying solutions in investment spread across a wide spectrum such as securities, real estate, private equity, portfolio analytics, and more. As a company Two Sigma believes that the companies of the future will have three main pillars that are firstly availability of information, secondly advancements in science of computing, and lastly, the inventiveness of founders. With these principles at its core, Two Sigma continues to pioneer innovation in the financial sector, driving transformative growth and shaping the future of investment.

Company Strengths

- Quantitative investing strategies used by Two Sigma to analyze data and identify trends. These are then used to make investment decisions, by analyzing historic market data, alternative data, fundamental analysis and microeconomic conditions. They use multiple strategies to deliver alpha returns (returns above average market returns) and diversify their risk across asset classes.
- They use Advanced Data Science and statistical techniques to create predictive models, forecast price movements, portfolio returns and manage risk profiles. They additionally use both traditional and non traditional data sources to make informed investing decisions. They used rule based algorithmic trading strategies which execute trades on fractional time periods, eliminating bias and emotion from the decision making process and further enhancing the speed and accuracy of trading.
- Two Sigma employs multiple teams of scientists and researchers and invests heavily in its Research and Development (R&D) areas to enhance their proprietary trading algorithms and investment thesis.
- On top of their traditional hedge fund offerings, Two Sigma also does private investing for high and ultra high net worth individuals (HNIs and UHNIs) along with private equity and venture capital investing through its arm, Two Sigma Ventures.

About the Quantitative Research Position

The Quantitative Research Position spans over 10 weeks in the summer and takes place at their New York City office. Interns are made to partner with an assigned mentor and work on a single project during the course of your time here, which will culminate in a final presentation at the conclusion of the program. Interns will be alongside engineers and scientists from whom they can learn, collaborate and ideate with. This position allows incoming interns to tackle tough problems besides individuals that will challenge their ideas. Interns would use scientific methods to develop sophisticated investment models and shape our insights into how the markets will behave. They would apply quantitative techniques like machine learning to a vast array of datasets and lastly create and test complex investment ideas and partner with our engineers to test your theories.

In addition to gaining valuable practical experience, Two Sigma ensures interns stay engaged with the academic community by offering reading circles around the latest research papers, opportunities to attend academic seminars featuring leaders from top universities and more

To be qualified for this position candidates must be pursuing a degree in either a technical or quantitative discipline such as statistics, mathematics, physics, electrical engineering, or computer science with approximately one year remaining in your programs. All levels are welcome, from bachelor's to doctorate. While one does not need a background in finance, a candidate should demonstrate intermediate skills in at least one programming language such as C, C++, Java, or Python. Lastly, they are looking for independent thinkers who can creatively approach data analysis and communicate complex ideas clearly.

Company Strategy & Position Impact

Two Sigma Strategy: Two Sigma benefits from various investing verticals. Their most notorious ones being the access to their strategy investment funds, venture capital investing, real estate and private equity. Two Sigma integrates some of the most cutting edge technology and data science models to assess undervalued companies and provide their investors with unique options to maximise returns while combining assets that hedge market risks. The firm's competitive advantage is mainly driven by its new technology and how it leverages Artificial Intelligence, Machine Learning Models and Alternative to research patterns that explain securities pricing and form portfolios with a wide range of asset types.

A Quantitative Researcher (Quant Researcher) at Two Sigma plays a big role in developing and improving the firm's investment strategies by performing the following:

- They develop quantitative and statistical models that build the backbone of Two Sigma's trading strategy to analyze data and recognize price patterns. They primarily focus on data sourcing, cleaning, risk and price modelling, and forecasting.
- Developing novel algorithms to stay ahead of the competition by staying up to date and to rapidly adapt to changing market conditions. Use cases may include using Convolutional Neural Networks, Reinforcement Learning, and Natural Language Processing to build models for trade execution, portfolio rebalancing, and more. They additionally focus on enhancing and optimizing these algorithms by using live data and backtesting to ensure accuracy and improve performance in real-world trading environments
- They also must continue to balance risk and optimize portfolios to get the return they expect by diversifying investments across asset classes. They use Monte Carlo simulations to understand various case scenarios and how the models and strategies behave during adverse market events.

Position Fit

As a group we believe we 4 are a good fit for this job as being a student enrolled in the Master's in Interdisciplinary Data Science (MIDS) program, especially with a focus in Quantitative Finance as we are working on creating a strong technical foundation where we

are gaining advanced knowledge in statistics, machine learning, and data analysis that are skills directly aligning with the requirements of this internship. The program's focus on analytical and computation techniques allows students to tackle the complexly vast datasets and sophisticated investment models that they will face at Two Sigma. The program ensures the development of programming skills particularly in python for all students irrespective of prior experience. In our program we are taught how to apply these skills in real world problems thereby meeting Two Sigma's technical expectation in quantitative modelling and data analysis. We also believe with us being in the Quantitative Finance concentration at Duke University it will further strengthen our expertise in financial markets and quantitative techniques which would make us all well-prepared to develop and test complex investment ideas, which is an integral aspect of this role at Two Sigma.

Networking

We could leverage alumni connections from our school/university, especially with those who have pursued similar quant roles or are in similar firms in the quant landscape. Additionally, we could reach out and talk with professors, particularly those in applied mathematics, computer science, or finance, who may have industry contacts or know alumni working in quant roles. They can help with making introductions or suggest networking events.

We could also attend quantitative finance, data science, and AI-focused conferences which bring together professionals from top quant firms and hedge funds, offering networking opportunities during talks and workshops. Participating in industry-relevant hackathons like Kaggle competitions allows us to enhance our personal portfolio and build in the public sphere. Additionally, career fairs give us the opportunity to introduce ourselves to the company, engage with and learn more about the firm and role, and gather relevant contact information to then reach out over email/Linkedin.

Lastly, we could build robust profiles and technical portfolios on platforms like Linkedin and Github where we can highlight projects in quantitative research, data science, and machine learning. Contributing to open-source projects and possibly publishing research or even writing articles allows us to showcase our work and expertise.

Interview Process

Based on our research from online websites we were able to learn that the job application process goes as follows:

1. There will be an online coding assessment
2. 2 technical interviews
3. Virtual onsite interview where you meet with several hiring managers from different groups.

Preparing for an interview with Two Sigma for their Quantitative Research role requires a strong foundation in mathematics, programming, finance, and data analysis. Two Sigma is well-known for using mathematical models, machine learning, and data science to develop its trading strategies. Looking at Glassdoor reviews from students who have interviewed at Two Sigma, interviews are expected to have medium-hard leet code

questions, regression using python, statistics, stochastic calculus, and game theory questions. These questions are much harder than normal quantitative finance roles and expect the candidate to be well-rounded.

Preparation for this interview demands an organized strategy. Reviewing core mathematical concepts, including linear algebra, probability theory, and statistical inference, will help build the foundational knowledge needed for problem-solving. A good resource to brush up on these topics is the “Green Book”. In addition, practicing programming challenges on platforms like LeetCode and refining coding skills with large datasets are critical steps to demonstrate computational efficiency and accuracy. It would also be good to know key financial models, such as the Black-Scholes model and portfolio optimization techniques, as these are often integral to the work of a quantitative researcher at a firm like Two Sigma. Furthermore, reading the company's research papers and staying current with trends in the quantitative finance space can provide valuable insights into how Two Sigma approaches its innovative strategies.

Beyond technical skills, behavioural interview questions will likely emphasise problem-solving, creativity, and teamwork. According to Glassdoor reviews, you should prepare to discuss how one approaches complex challenges and develops efficient solutions, particularly in environments with incomplete data. Two Sigma values collaboration and intellectual curiosity, so it will be essential to demonstrate not only individual technical strength but also the ability to work well on a team. Additionally, team culture become important once you reach the final rounds of the interview process.

For preparation one could take inspiration from the Quantitative research interview demo here.

Growth at Two Sigma

Two Sigma is a highly competitive firm to enter and the growth phase for a prospective employer initiates as they begin interviewing for a given role. Incoming employees have to master the arts of solving complex problems, communicate their findings effectively and be self-aware of themselves and team members. The two sigma interview requires you to master all behavioural questions, problem solving and technical assessments in addition to questions that express solid understanding about the financial domain.

Normally most entry Levels roles are for Associates who then get to be promoted for more senior roles, and subsequently into managerial positions. According to company reviews however, Two Sigma doesn’t necessarily favour internal candidates for senior managerial positions- instead, when an opening emerges, the company still proceeds to run a rigorous process to find the right candidate for that position. In addition, while a well recognized firm, two sigma is still considered relatively small when compared to larger financial institutions which oftentimes results in not as many leadership openings being available right away. Despite this criticism from company reviews, most current employees still agree the company deviates from the archetypal traditionalist culture seen at most banks. Instead, the company is known to be a financial institution with a culture of problem solvers.

For employers that move across from entry level to managerial roles, they normally have internship experience while pursuing their education at a top financial institution. At two sigma new hires would work as a quantitative researcher initially, move into a VP role, SVP, and managing director. A common attribute from most individuals who work at two sigma is that they possess a very strong technical acumen paired with a background in engineering, computer science or mathematics. The firm is known for being structured by top talent and some of the most intelligent professionals in the industry. Oftentimes, individuals that are hired at two sigma don't always necessarily come from a finance background as the firm opts to select employees who worked on their own startups or in big technology companies.

Two sigma offers employees a distinctive way to grow by having individuals work with a reasonable amount of alternative data to drive business decisions. For instance, Two sigma is known to connect transportation data and observe how high traffic areas might be attributed to the overall business performance of a given company in the brick and mortar sector. Another way by which the company uses alternative data is by evaluating drone footage for given areas that explore commodities. This unconventional, yet unique approach provides them an edge when determining which assets to select in their portfolios. To quantitative researchers, this means they get exposed to unique projects and while putting their best skills in practice while working.

Conclusion

We believe that MIDS positions us well for a career as a Quantitative Researcher. With exposure to the world of finance through courses such as Fundamentals of Finance and Algorithmic Trading, we get to work on real-world data and projects to help us hone our technical skills and finance knowledge. We break down each step and skill above needed for Quantitative Research, including the interview process, desired technical skillset, roles and responsibilities and most of all the impact we can create at firms such as Two Sigma amongst others. Two Sigma leverages data science and advanced algorithmic trading strategies to solve some of the biggest challenges in the world of Finance. As quantitative research interns, we have the opportunity to be at the front seat and contribute to their growth.