[Markowitz Portfolio Optimizations on Varying Asset Classes]
[Data Analytics and Sports Performance: The Balance Between Numbers and Tradition]
A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science
By
Cameron Greene
November 3, 2023
On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.
ADVISORS
Prof. Pedro Augusto P. Francisco, Department of Engineering and Society

[INTRO]
In this new age of information, perfection is now seen as a goal that can finally be achieved in every sector, whether it be in finance or in sports. This realization is due to the widespread availability of data. Data, when expertly analyzed can lead to better decision making, better performance, and it can even predict the future. My Capstone project, “Markowitz Portfolio Optimizations on Varying Asset Classes”, uses financial data to maximize investment returns for a given level of risk. And this is mirrored by my STS research, “Data Analytics and Sports Performance: The Balance Between Numbers and Tradition”, which explores the cultural and management shifts in the sports industry due to the increase in data analytics.

The technical project devises an effective investment strategy within the financial markets by focusing on the optimization of risk and return. The implementation applies Modern Portfolio Theory (MPT) to achieve a balance between risk and potential gains. The developed Python script applies these financial principles to showcase the financial data in an efficient and accessible manner. The script represents the development and integration of longstanding financial strategies combined with the speed and accuracy of computer science. In the sports world a similar revolution is occurring. In the past few years the majority of the sporting culture has widely adopted data analytics to help improve performance on and off the field. My STS research explores the impact this change has had on the culture of sports, questioning if the rise of data is diminishing the human elements, such as the passion, tradition and culture, of sports.

Both the technical and STS topics are powered by the changes in data analytics and the growing desire to quantify everything.The connection between these topics lies in their growing reliance on employing data to make decisions. In finance, the constructed portfolio allows the user to input data to see the best combination of assets that maximize their returns. And in sports, a variety of data-heavy tools are used to make decisions that affect front office strategies and results on the field. Both use historical data, and statistical models to achieve their goals. And on the other hand, they pose challenges such as ethical dilemmas, privacy issues and managing the integration of machine learning into traditional practices.

The technical project seeks to secure financial well-being and the STS project aims to enhance athletic performance and recognize the implications of a data-driven sports world. These projects reflect the role that a computer scientist has on developing technology, but also how the technology they create influences society. The main idea binding these topics is the pursuit of excellence through the use of data analytics, and the aspiration to perfect human intuition with data-driven insights. These projects, although different, are essential to engineering and computer science, as they demonstrate how data can be utilized to advance our understanding of various aspects of life.

[Markowitz Portfolio Optimizations on Varying Asset Classes]

The modern financial market is a complex system where investors compete to maximize their returns. The core idea of their strategies often relies on constructing a portfolio that optimizes their returns for a given level of risk. The project I worked on is a representation of this technical solution, designed to implement the foundational ideals of Modern Portfolio Theory (MPT). MPT, was introduced by Nobel Prize winner Harry Markowitz in 1952, and revolutionized the way investors viewed portfolios. His work aimed to quantify the balance between risk and expected return through a diversified portfolio(Markowitz, 1952).

The technology behind the project is a combination of financial analysis and computational programming. It consists of a Python script that uses packages such as `yfinance` to retrieve financial data, `pandas` for data manipulation, `numpy` for numerical computations, and `matplotlib` for visualization. It is designed to automate the process of data retrieval, analysis, and visualization for investment portfolio optimization. The project's goal is to construct an efficient frontier, a concept from MPT. The efficient frontier represents a set of investment portfolios that are expected to provide the highest return for a given level of risk or the lowest risk for a given level of return. This is visualized as a curve on a graph where the x-axis represents portfolio risk (measured as standard deviation) and the y-axis represents expected return. This is achieved by downloading the historical stock prices, then using them to calculate the expected returns and volatility of each asset. This is then used to compute the covariance matrix which shows how the asset prices move in relation to each other. Finally, the weights for each risk level are optimized to produce the final frontier curve. The main feature of this project however, is its ability to apply constraints, such as setting a maximum or minimum weight for certain assets, which could represent investor preferences or regulatory requirements. Investors preferences can vary greatly depending on the level of risk they are willing to take on. This feature allows the user to see the optimized frontier for each portfolio class concurrently, so that they will be able to choose which portfolio is optimal for their risk level.

This project is a demonstration of how data science has become crucial in modern investment management. By leveraging historical data and mathematical models, investors can make more informed decisions that are supported by quantitative analysis rather than intuition alone. It simplifies complex financial concepts into actionable insights and the implications of

such technology are significant because it allows access to sophisticated investment strategies that were once only accessible to financial analysts and institutions. In conclusion, the code serves as a bridge between traditional financial theory and modern investment practice. It illustrates the impact of computer science in the financial sector and emphasizes the potential of technology and its use in finance.

[Data Analytics and Sports Performance: The Balance Between Numbers and Tradition]

Within the intersection of sports and data a critical question emerges, How is the rise of data analytics influencing the cultural landscape of sports, and potentially leading to a transition from traditional wisdom to relying purely on data? This inquiry delves into the essence of sports identity and it challenges the balance between objective analysis and the passion and human elements of sports. This research aims to explore the implications of a data-driven culture within the sports realm. It will investigate whether the reliance on analytics undermines the traditional wisdom and intuition that have long been celebrated in sports, and how this shift may potentially trade the raw emotion of the game for calculated objectivity.

Data analytics now powers modern sports, from boosting player performance to driving fan engagement and refining marketing strategies. Historical data and analytics guide teams in enhancing game tactics and even business operations such as fan behavior insights (Master's in Data Science, 2020). In addition, there has been a wide-spread adoption of injury-prevention technologies such as wearable sensors and simulation goggles, however, it also introduces ethical dilemmas related to data privacy, as the athletes' data is accessible to more people(Deloitte Insights, 2020). Astera (2023) notes a significant shift towards data-centered decision-making in sports, as seen with NBA teams employing machine learning algorithms to predict game outcomes to set their lineups, moving away from instinctual decisions. This evolution impacts not just performance but also the overall sports culture, influencing player morale, sponsorships, but most importantly it reduces the human interaction between the coach and players. In the NFL, analytics form the base of player evaluations and game strategies, demonstrating a broader shift across all of sports to data reliance and indicates a cultural redefinition of sports management and consumption. This technology is continuing to grow as the sports analytics sector, currently valued at 2.73 billion USD is expected to grow at a 23% CAGR until 2030, reflecting the industry's deepening investment in a data-driven game (Grandview Research, 2023). Such a deep-rooted change towards data reliance in sports culture, represents a significant turn from traditional methods. It not only signifies an evolution in sports performance but also hints at a potential shift in the sports culture that has long celebrated the unpredictability and emotions of sports.

The integration of data analytics into sports has been a game-changer, but not without its criticisms. Moruzzi (2014) in The Guardian warns that an over-reliance on statistics in football may strip the game of its essence, suggesting that the data lacks the context to capture the full

story. This sentiment is echoed across other sports, where despite the undeniable usefulness of AI, such as in the NFL's recruitment process, experts from Harvard Business Review (2023) contend that human judgment still has an irreplaceable role in strategy and execution. Wired (2016) also explores the NFL's embrace of data, with advanced tracking metrics like Sparta's force-plate, which can predict injuries and customize training. However, they also emphasize the significance of the human element amid the wealth of statistics. Similarly, Sleep and Thomas (2023) examine the impact of fantasy football on fan loyalties. Suggesting that the statistical performance of players is increasingly influencing fan behavior and support, as they start to shift towards supporting individual players as opposed to their teams and the emotional connections that traditionally accompany them.

The analysis of data analytics' impact on sports culture will be approached through a mixed-methods research design, using both qualitative and quantitative data. Qualitative insights will be gained from interviews with athletes, coaches, and analysts to understand their experiences and perceptions of data analytics in the decision-making process. Quantitative evidence will involve collecting performance statistics, injury rates, and fan engagement metrics before and after the rise of advanced analytics tools. This evidence will be interpreted through a by, assessing changes in performance, strategy, and fan behavior over time. Statistical tools will identify trends and correlations between the use of analytics and changes in sports culture. Qualitative data will be analyzed to extract common narratives and sentiments about the shift towards data-driven decision-making. The use of both data types will provide a comprehensive understanding of how analytics are redefining sports, by changing the balance between numerical insights and traditional wisdom and passion that define athletic competition.

[CONCLUSION]

In conclusion, the technical portion of my research consists of a python script that performs a portfolio optimization on variously weighted portfolios using Modern Portfolio Theory. This tool will offer investors a more refined approach to managing their assets at different risk levels to maximize returns. Meanwhile the STS research showcases the influence of data analytics in sports, a trend that is growing exponentially and changing everything in the sports realm. This investigation aims to evaluate the shifting balance between data-driven decision making and the human elements that are involved in sports. The expected results of this research will contribute towards developing data analytics that coexist with the traditional sports culture instead of overwhelm it. Together, these deliverables will enhance our understanding of the complex problems that may arise from the data revolution in both finance and sports. And in turn guide us towards solutions that honor bothe the power of data and value of human insight.

[Works Cited]

Big Data in Sports | Master's in Data Science. (2020, September 23). Master's in Data Science. https://www.mastersindatascience.org/resources/big-data-in-sports/

Bodie, Z., Kane, A., & Marcus, A. J. (2018). *Investments*. McGraw-Hill Education.

Brinson, G. P., Singer, B. D., & Beebower, G. L. (1991). Determinants of Portfolio Performance II: An Update. *The Financial Analysts Journal, 47*(3), 40-48. https://doi.org/10.2469/faj.v47.n3.40

Cocco, J. F., Gomes, F. J., & Maenhout, P. J. (2005). Consumption and Portfolio Choice over the Life Cycle. *The Review of Financial Studies, 18*(2), 491-533. https://doi.org/10.1093/rfs/hhi017

Dhar, V. (2013). Data science and prediction. *Communications of the ACM, 56*(12), 64-73. https://doi.org/10.1145/2500499

Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzmann, W. N. (2007). *Modern Portfolio Theory and Investment Analysis*. Wiley.

Fabozzi, F. J., Focardi, S. M., & Kolm, P. N. (2007). *Financial Modeling of the Equity Market: From CAPM to Cointegration*. Wiley.

Fabozzi, F. J., Gupta, F., & Markowitz, H. M. (2002). The Legacy of Modern Portfolio Theory. *The Journal of Investing, 11*(3), 7-22. https://doi.org/10.3905/joi.2002.319247

Givenchy, D. (2023, February 10). From Player Evaluation to Game Strategy: How Analytics is Changing the NFL - The Football Tech. https://thefootballtech.com/nfl/from-player-evaluation-to-game-strategy-how-analytics-is-changing-the-nfl/

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. (2020). Array programming with NumPy. *Nature, 585*(7825), 357-362. https://doi.org/10.1038/s41586-020-2649-2

Jarvis, D., & Westcott, K. (2020, December 7). The use of athlete data analytics. Deloitte Insights. https://www2.deloitte.com/xe/en/insights/industry/technology/technology-media-and-telecom-predictions/2021/athlete-data-analytics.html

Lazer, D., Pentland, A., Adamic, L., et al. (2009). Computational Social Science. *Science, 323*(5915), 721-723. https://doi.org/10.1126/science.1167742

Lindsey, J. (2016, January 13). The NFL Is Finally Tapping Into the Power of Data. WIRED. https://www.wired.com/2016/01/the-nfls-impending-data-revolution/

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance, 7*(1), 77-91. https://doi.org/10.2307/2975974

McKinney, W. (2010). Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*. https://conference.scipy.org/proceedings/scipy2010/mckinney.html

Moruzzi, M. (2014, February 3). Does analyzing football through statistics miss the point of the game? The Guardian. https://www.theguardian.com/football/when-saturday-comes-blog/2014/feb/03/statistics-football-analysis-miss-point-game

Sharpe, W. F. (1966). Mutual Fund Performance. *Journal of Business, 39*(1), 119-138. https://doi.org/10.1086/294846

Shapiro, J., & Velte, T. (2023, June 16). Will AI Replace the Front Office in Pro Sports? Harvard Business Review. https://hbr.org/2023/06/will-ai-replace-the-front-office-in-pro-sports

Sleep, S., Thompson, S., & Thomas, M. (n.d.). The Impact of Fantasy Football on the NFL Fan: Exploring Differences between the Fantasy Football Participant and the Traditional Fan.