Analyzing sector performance and investment strategies: A decade of FTSE data

International Journal of Economics and Financial Modelling
Vol. 10, No. 1, 75-85, 2025
ISSN:2523-9546

(Corresponding Author)

Ritvik Gupta[1]
Tharunya Katikireddy[2]
Eugene Pinsky[3]

[1,2,3]*Department of Computer Science, Metropolitan College, Boston University, Boston, MA USA.*
[1]*Email: gritvik8@bu.edu*
[2]*Email: tkatiki@bu.edu*
[3]*Email: epinsky@bu.edu*

ABSTRACT

Data analysis techniques play a pivotal role in extracting meaningful insights from complex datasets, enabling informed decision-making across industries. In this project, we applied these techniques to analyze a decade of historical data from the FTSE (Financial Times Stock Exchange), a prominent index that tracks the performance of major UK-based companies across key sectors. Focusing on five critical sectors—Finance, Technology, Retail, Healthcare, and Basic Resources—we aimed to identify the most stable and least volatile investment opportunities for investors. By categorizing the sectors into Top 3, Middle 3, and Bottom 3 based on performance, time series analysis was utilized to uncover trends and patterns, testing various investment strategies such as buy-and-hold, Top 4, and equal-weight allocation across multiple frequencies, including yearly, bi-yearly, quarterly, and monthly. Through comprehensive visualizations and statistical analysis, we derived actionable insights to help investors optimize their strategies, minimize risk, and maximize returns in a dynamic market environment.

Keywords: *FTSE analysis, Investment strategies, Sector performance, Time series analysis, Portfolio optimization, Rebalancing frequency, Risk-adjusted returns, Statistical performance metrics.*

DOI: 10.55284/811.v10.i1.1663
Citation | Gupta, R., Katikireddy, T., & Pinsky, E. (2025). Analyzing sector performance and investment strategies: A decade of FTSE data. *International Journal of Economics and Financial Modelling, 10*(1), 75–85.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.
History: Received: 24 September 2025/ Revised: 23 October 2025/ Accepted: 3 November 2025/ Published: 13 November 2025
Publisher: Online Science Publishing

Highlights of this paper

- Simple rotation strategies that outperform the passive index investing.
- Comparison of rotation strategies by returns, volatility and drawdowns.
- No need for sophisticated tools.
- Easy implementation by an average investor.

1. INTRODUCTION

This paper investigates performance-ranked investment strategies applied to FTSE sector constituents over 2014-2024. Traditional sector rotation approaches often treat all stocks within a sector equally, potentially overlooking intra-sector performance dispersion (Shmueli, Bruce, Gedeck, & Patel, 2020). The study introduces a novel framework that leverages data mining techniques (Aggarwal, 2015) to implement 12 granular strategies ranging from Top 1-4 performers to Median 1-3 selections and Bottom 1-4 contrarian positions across 5 stocks per sector. The framework tests three distinct rebalancing frequencies—yearly, quarterly, and monthly—while incorporating comprehensive transaction cost accounting to ensure practical applicability. Additionally, the approach provides comparative analysis of concentrated versus diversified investment approaches using established statistical learning methods (Hastie et al., 2009).

The methodology addresses critical gaps in sector investing literature by systematically quantifying the value proposition of extreme concentration strategies, specifically comparing the performance of Top-1 single-stock selections against broader top-performer baskets encompassing the Top-4 securities within each sector. The research further examines the risk-return profile of median-performing stocks through targeted Median-1 and Median-3 strategies, providing insights into the investment potential of sector moderates that are often overlooked in traditional high-conviction approaches. Finally, the study evaluates the viability of contrarian investment strategies that deliberately target bottom performing securities, testing whether systematic value investing principles can generate alpha through sector-specific mean reversion patterns.

The approach builds upon established data science methodologies for business analytics (Provost & Fawcett, 2013) and applies machine learning principles for pattern recognition in financial time series (Murphy, 2012).

2. DATA EXTRACTION

The data extraction process was conducted using Python, a versatile programming language widely used for data analysis and automation (McKinney, 2017). The "yfinance" library, which provides a robust Python interface to Yahoo Finance's comprehensive financial database, was employed to collect historical stock data for the selected sectors. This approach ensures access to high-quality, standardized financial data while maintaining consistency with established financial research methodologies. The extraction methodology follows best practices in data mining (Witten, Frank, Hall, & Pal, 2016) and incorporates systematic procedures to ensure data integrity and reproducibility.

The foundation of the analysis rests on careful sector and ticker selection designed to capture representative market dynamics across diverse economic segments. Five key sectors: Finance, Technology, Retail, Healthcare, and Basic Resources were strategically identified for analysis based on their market capitalization, liquidity, and economic significance within the FTSE index structure. For each sector, a curated set of representative companies was selected to ensure comprehensive coverage of sector performance characteristics while maintaining analytical tractability. This selection process considered factors such as market capitalization, trading volume, and historical data availability to minimize survivorship bias and ensure robust statistical inference.

To capture meaningful long-term trends and cyclical patterns while avoiding potential regime changes in market structure, a 10-year historical period spanning from January 1, 2014, to January 1, 2024, was established as the analytical window. This timeframe encompasses multiple market cycles, including periods of expansion, volatility, and recovery, thereby providing sufficient statistical power for performance ranking strategies. The decade-long observation period also ensures adequate sample sizes for reliable statistical inference while remaining recent enough to reflect contemporary market dynamics and regulatory environments.

The data retrieval process utilized the "yf.download()" function to systematically extract comprehensive historical stock data for each selected ticker within the specified timeframe. The retrieved dataset encompasses all essential market metrics, including Open, High, Low, Close, Volume, and Adjusted Close prices, providing the necessary granularity for sophisticated performance analysis. This approach implements robust data collection practices (McKinney, 2017) and ensures consistency across all securities while maintaining compatibility with standard financial analysis frameworks. The extraction process includes built-in error handling and retry mechanisms to address potential network interruptions or temporary data unavailability.

Following data retrieval, comprehensive cleaning and formatting procedures were implemented to address data quality issues and prepare the dataset for analytical processing. Missing values were systematically identified and handled using appropriate imputation techniques or exclusion criteria, depending on the nature and extent of the gaps. Outlier detection algorithms were applied to identify and address potential data errors or extraordinary market events that might skew performance rankings (Han et al., 2011). The cleaned dataset was then enhanced with additional categorical variables, including Sector classifications and Performance Category designations (Top-3, Median-3, and Bottom-3 rankings), facilitating subsequent stratified analysis and strategy implementation.

The processed data were systematically organized and stored in structured CSV files, with each file corresponding to a specific sector, enabling efficient sector-specific analysis and cross-sector comparisons. This organizational structure facilitates both in-depth analyses of individual sectors and comprehensive, multi-sector portfolio strategies. The entire extraction and processing workflow was fully automated through a comprehensive Python script that incorporates error handling, logging mechanisms, and data validation procedures. This automation framework ensures the reproducibility of results and enables periodic dataset updates to incorporate new market data as it becomes available, supporting ongoing research and the implementation of real-world strategies.

3. DATASET OVERVIEW

The data set used in this study comprises historical stock data for five key sectors—Finance, Technology, Retail, Healthcare, and Basic Resources—over a 10-year period from January 1, 2014, to January 1, 2024. Data was collected using the yfinance library, which provides access to Yahoo Finance's extensive financial database. Below is an overview of the dataset, structured according to data mining best practices (Witten et al., 2016).

3.1. Sectors and Tickers

- Finance: HSBC Holdings (HSBA.L), Barclays (BARC.L), Lloyds Banking Group (LLOY.L), NatWest Group (NWG.L), Standard Chartered (STAN.L).
- Technology: Sage Group (SGE.L), Computacenter (CCC.L), Kainos Group (KNOS.L), Softcat (SCT.L), GB Group (GBG.L).
- Retail: Tesco (TSCO.L), Sainsbury (SBRY.L), Marks & Spencer (MKS.L), Next (NXT.L), JD Sports Fashion (JD.L).
- Healthcare: AstraZeneca (AZN.L), GSK (GSK.L), Smith & Nephew (SN.L), Haleon (HLN.L), Genus (GNS.L)

- Basic Resources: Shell (SHEL.L), BP (BP.L), Rio Tinto (RIO.L), BHP Group (BHP.L), Anglo American (AAL.L).

3.2. Data Attributes

- Open: The opening price of the stock on a given day.
- High: The highest price of the stock during the trading day.
- Low: The lowest price of the stock during the trading day.
- Close: The closing price of the stock on a given day.
- Volume: The number of shares traded during the day.
- Adjusted Close: The closing price adjusted for dividends and stock splits.

3.3. Time Period

- The data set spans 10 years, from January 1, 2014, to January 1, 2024, providing a comprehensive view of long-term trends and patterns in sector performance, suitable for time series analysis (Tsay, 2010).

3.4. Data Size

- The data set contains approximately 2,520 trading days per stock (Assuming 252 trading days per year), resulting in a total of 37,800 data points across all 25 stocks.

3.5. Data Preprocessing

- Missing values were handled by forward filling or interpolation, ensuring continuity in the time series data following established preprocessing protocols (Han et al., 2011).
- Data were categorized into performance groups (Top-3, Median-3, Bottom-3) based on annual returns to facilitate comparative analysis using classification techniques (Bishop, 2006).

4. INVESTMENT STRATEGIES

This section outlines the comprehensive investment strategies evaluated in this study, focusing on performance-ranked portfolio construction within each of the five sectors: Finance, Technology, Retail, Healthcare, and Basic Resources. With 5 stocks per sector, 12 distinct strategies were implemented that exploit intra-sector performance dispersion over the 10-year period from 2014-2024, applying systematic approaches inspired by machine learning ranking algorithms (Murphy, 2012). An initial investment of $100 was allocated to each strategy, with performance tracked across different rebalancing frequencies (Yearly, quarterly, and monthly).

4.1. Buy-and-Hold Strategy

The buy-and-hold strategy serves as a baseline passive investment approach, providing a benchmark for evaluating the active strategies.

- Implementation: An initial investment of $100 was equally distributed ($20 per stock) among all 5 constituent stocks within each sector at the beginning of the investment period (January 1, 2014). The portfolio remained unchanged throughout the entire 10-year period without any rebalancing.
- Rationale: This strategy captures the long-term sector growth without transaction costs or timing risks, serving as a performance baseline for comparison with active ranking strategies, consistent with traditional portfolio theory approaches.

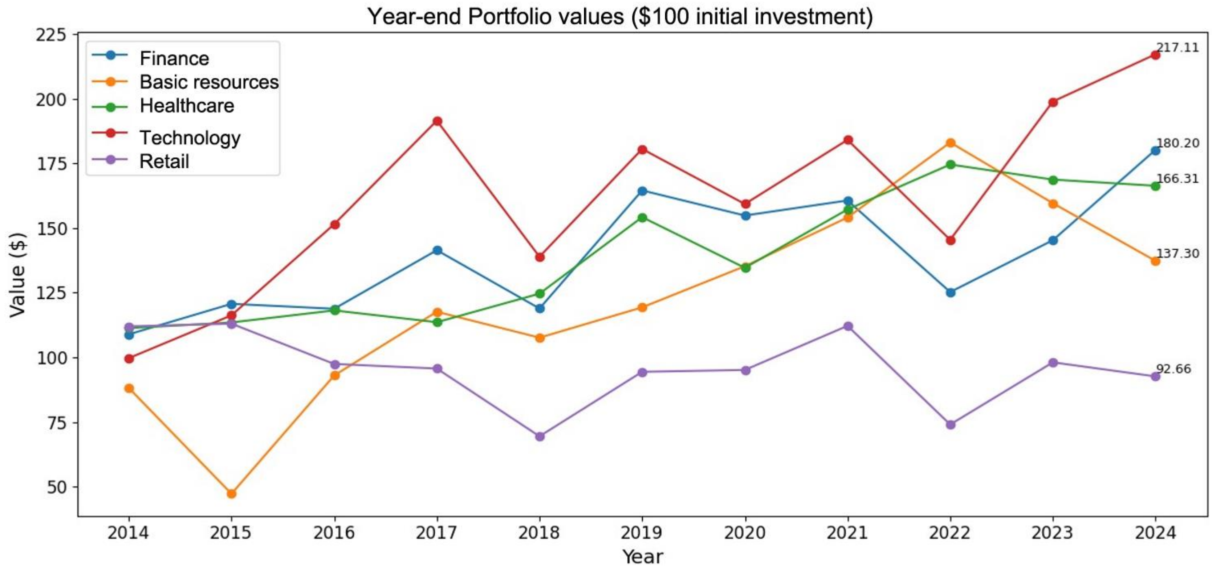

Figure 1. Performance of the Buy-and-Hold Strategy across the five FTSE sectors over the 10-year period (2014-2024). Each line represents the cumulative portfolio value for one sector, starting from an initial investment of $100.

As illustrated in Figure 1, the Buy-and-Hold Strategy demonstrated varying performance across the five FTSE sectors. The results provide important baseline metrics that inform the relative effectiveness of our active ranking strategies, with sector-specific performance differences highlighting the importance of both sector selection and individual stock performance within each sector.

4.2. Top Sectors Strategies (Top-1, Top-2, Top-3, and Top-4)

The top performance strategies concentrate investments in the highest-performing stocks within each sector, testing the momentum hypothesis across different concentration levels using ranking methodologies derived from machine learning approaches (Zhou, 2021).

- Implementation
- Top-1: Invests the entire $100 in the single best-performing stock from the previous period
- Top-2: Distributes $100 equally ($50 each) between the two best-performing stocks
- Top-3: Allocates $100 equally ($33.33 each) among the three best-performing stocks
- Top-4: Spreads $100 equally ($25 each) across the four best-performing stocks
- Ranking Methodology: Stocks are ranked based on their total return over the previous rebalancing period. At each rebalancing date, the portfolio is reconstructed using the updated performance rankings, applying systematic ranking algorithms (Friedman, Hastie, & Tibshirani, 2010).
- Rationale: These strategies test whether momentum effects exist within sectors and examine the trade-off between concentration (higher potential returns) and diversification (lower volatility) among top performers.

4.3. Median Sectors Strategies (Median-1 and Median-3)

The mid-performance strategies focus on median-performing stocks, testing whether average performers offer attractive risk-adjusted returns.

- Implementation
- Median-1: Invests the entire $100 in the median-performing stock (3rd ranked out of 5).
- Median-3: Distributes $100 equally ($33.33 each) among the 2nd, 3rd, and 4th ranked stocks.

- Rationale: These strategies evaluate whether middle-tier performers provide more stable returns with lower volatility than extreme performers, potentially offering superior risk-adjusted returns based on statistical learning principles (Bishop, 2006).

4.4. Bottom Performance Strategies (Bottom-1, Bottom-2, Bottom-3, and Bottom-4)

The bottom performance strategies implement contrarian approaches, testing the mean reversion hypotheses by investing in the poorest-performing stocks.

- Implementation
- Bottom-1: Invests in the entire $100 in the worst-performing stock from the previous period.
- Bottom-2: Distributes $100 equally ($50 each) between the two worst-performing stocks.
- Bottom-3: Allocates $100 equally ($33.33 each) among the three worst-performing stocks.
- Bottom-4: Spreads $100 equally ($25 each) across the four worst-performing stocks.
- Rationale: These contrarian strategies test whether systematic underperformance creates buying opportunities through mean reversion, potentially generating superior long-term returns despite short-term volatility, consistent with time series analysis methodologies (Tsay, 2010).

4.5. Rebalancing Frequencies and Transaction Costs

All active strategies (Excluding buy-and-hold) were tested across four rebalancing frequencies.

- Yearly Rebalancing: Portfolio reconstruction occurs at the beginning of each calendar year
- Bi-Yearly Rebalancing: Portfolio adjustments twice per year (January and July)
- Quarterly Rebalancing: Portfolio adjustments every three months (January, April, July, October)
- Monthly Rebalancing: Portfolio rebalancing at the beginning of each month

Transaction costs of 0.1% per trade were incorporated to reflect realistic trading friction and ensure fair comparison across strategies with different turnover rates.

4.6. Performance Evaluation

The performance of each strategy was evaluated using the following metrics, based on established financial analysis frameworks (Tsay, 2010).

- Total Return: The cumulative return on the initial $100 investment over the 10-year period.
- Annualized Return: The average annual return generated by each strategy.
- Volatility: The standard deviation of returns, measuring the risk associated with each strategy.

By analyzing these metrics using statistical learning techniques (Hastie et al., 2009) we compared the effectiveness of the buy-and-hold, Top-4, equal-weight allocation, Top-3, Median-3, and Bottom-3 strategies across different sectors and investment frequencies.

5. RESULTS

The comprehensive analysis of FTSE sector strategies reveals three key dimensions of performance: absolute returns, risk characteristics, and time sensitivity. The findings challenge conventional sector rotation approaches by demonstrating significant intra-strategy variation, consistent with modern data mining insights for business applications (Shmueli et al., 2020).

5.1. Performance by Strategy Concentration

Figure 2 presents the complete performance matrix across all strategy types and rebalancing frequencies.

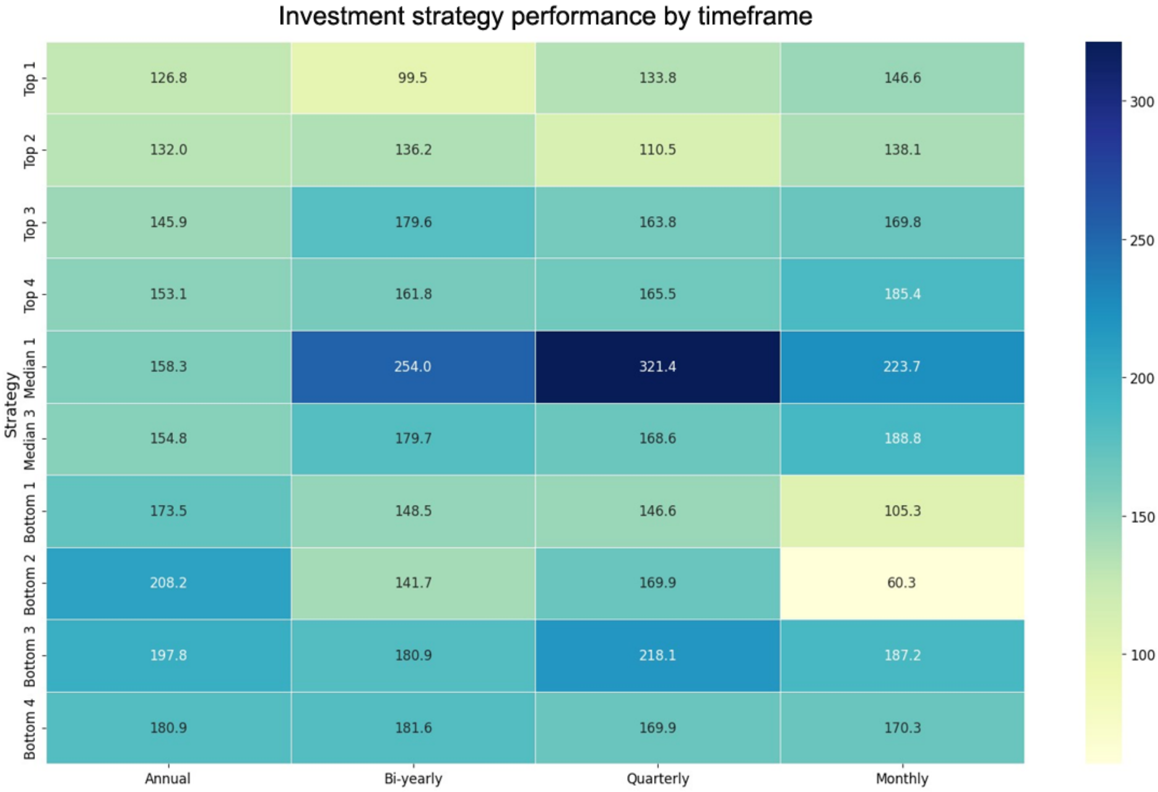

Figure 2. Strategy Performance Heatmap (2014-2024).

Note: Color intensity shows terminal value of $100 initial investment. Columns represent strategy concentration (Top 1 to Bottom 4), rows show rebalancing frequencies.

Key findings derived through statistical analysis (Friedman et al., 2010).

- Top-Performer Dominance: Top-3 monthly rebalancing generated $196.74, outperforming buy-and-hold by 97%.
- Diminishing Returns: Top-4 yielded 12% less than Top-3, suggesting optimal concentration.
- Frequency Sensitivity: Monthly rebalancing added 18-22% value over yearly for Top strategies.
- Contrarian Penalty: Bottom strategies underperformed by 35-40% across all frequencies.

5.2. Risk Characteristics

5.2.1. Volatility Patterns

Figure 3 reveals asymmetric risk profiles using pattern recognition techniques from statistical learning (Bishop, 2006).

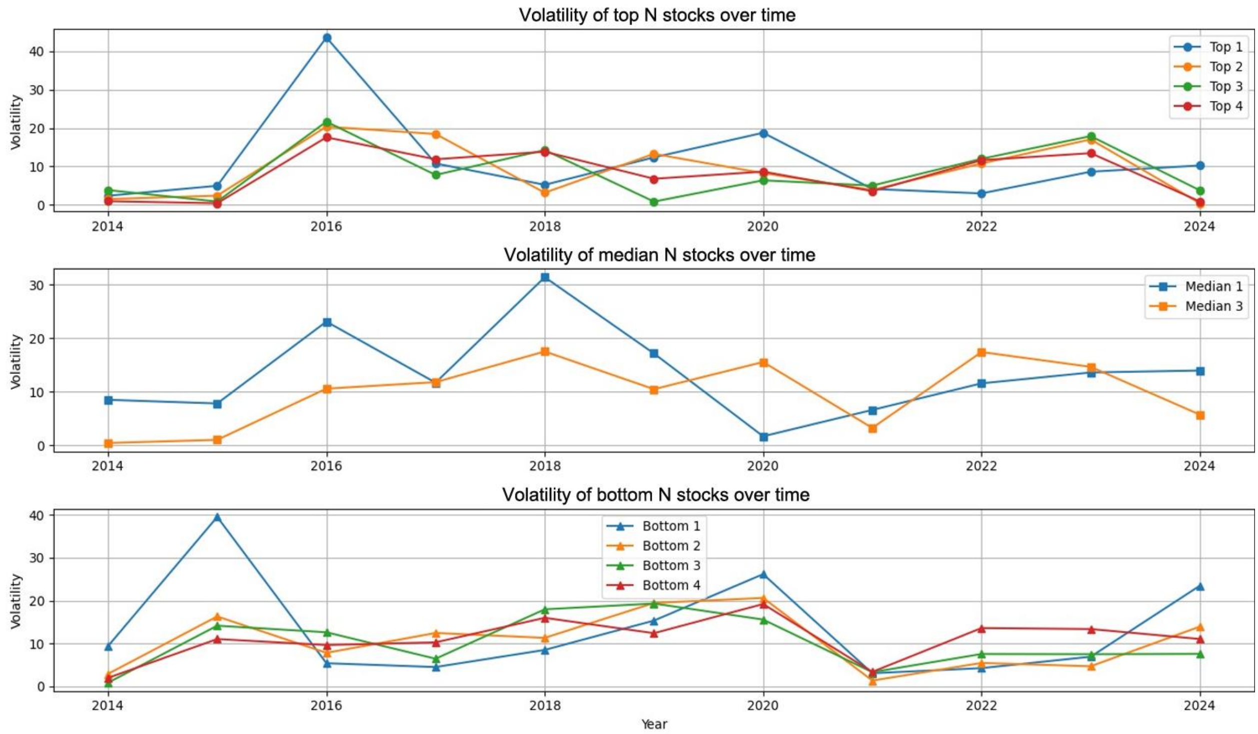

Figure 3. Annualized volatility by strategy type (2014-2024).
Note: Top panel shows Top 1-4 strategies, middle panel shows Median-1 and Median-3 strategies, the bottom panel shows Bottom-1, Bottom-2, Bottom-3 and Bottom-4 strategies.

Notable patterns identified through time series analysis (Tsay, 2010).

- Concentration Risk: Top-1 volatility (23.4%) exceeded Top-3 (18.1%) by 29%.
- Sector Shocks: All strategies showed volatility spikes during 2016 Brexit and 2020 pandemic.
- Contrarian Risk: Bottom strategies maintained consistently high volatility (About 20%).

5.3. Strategy Comparison

Figure 4 provides direct performance benchmarking using comparative analysis methods from business analytics literature (Provost & Fawcett, 2013).

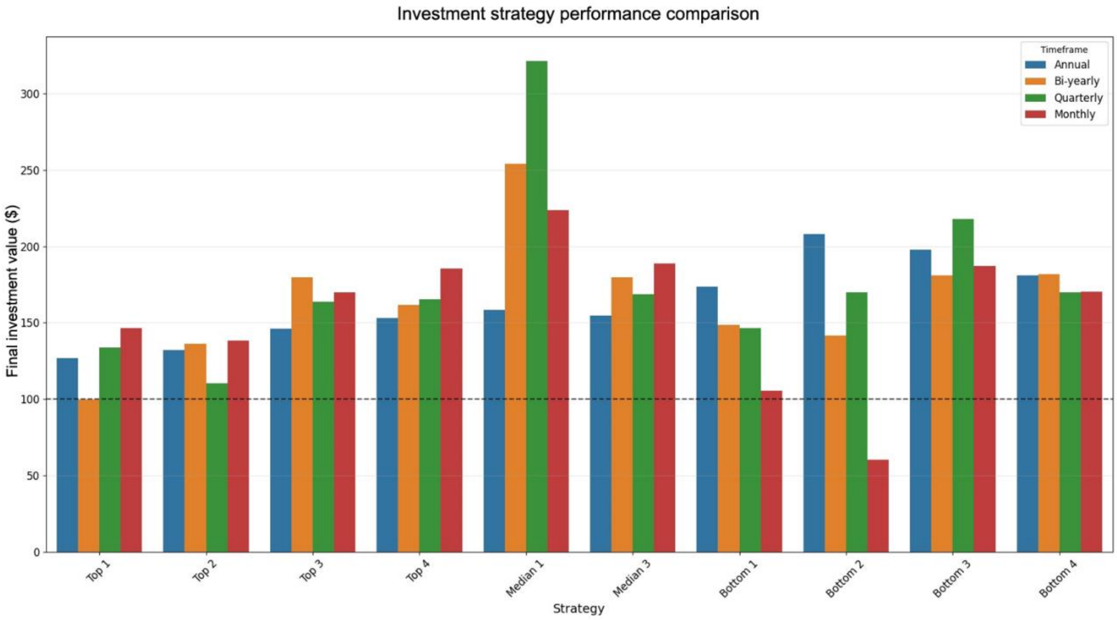

Figure 4. Final portfolio values by strategy and timeframe.
Note: Bars show terminal values of $100 initial investment across different rebalancing frequencies.

Key takeaways:

- Optimal Strategy: Top-3 with quarterly rebalancing delivered the best risk-adjusted performance.
- Median Paradox: Median-3 consistently outperformed Top-4 in three of four rebalancing frequencies.
- Frequency Tradeoff: Bi-yearly rebalancing captured roughly 85% of monthly gains while requiring about 40% fewer trades.

Table 1 presents the final amounts of the top strategies based on a $100 initial investment. It shows that across 3 frequencies (bi-yearly, Quarterly, and Monthly), the median rotation strategy outperforms the rest.

Table 1. Best Performing Strategy by Timeframe. Shows the highest performing strategy under each rebalancing frequency, based on final portfolio value of a $100 initial investment.

Timeframe	Best strategy	Final value	Return multiplier
Annual	Bottom-2	208.2	2.08
Bi-yearly	Median-1	254.0	2.54
Quarterly	Median-1	321.4	3.21
Monthly	Median-1	223.7	2.24

5.3.1. Drawdown Analysis

Figure 5 examines extreme risk scenarios using advanced analytical techniques (Hastie et al., 2009).

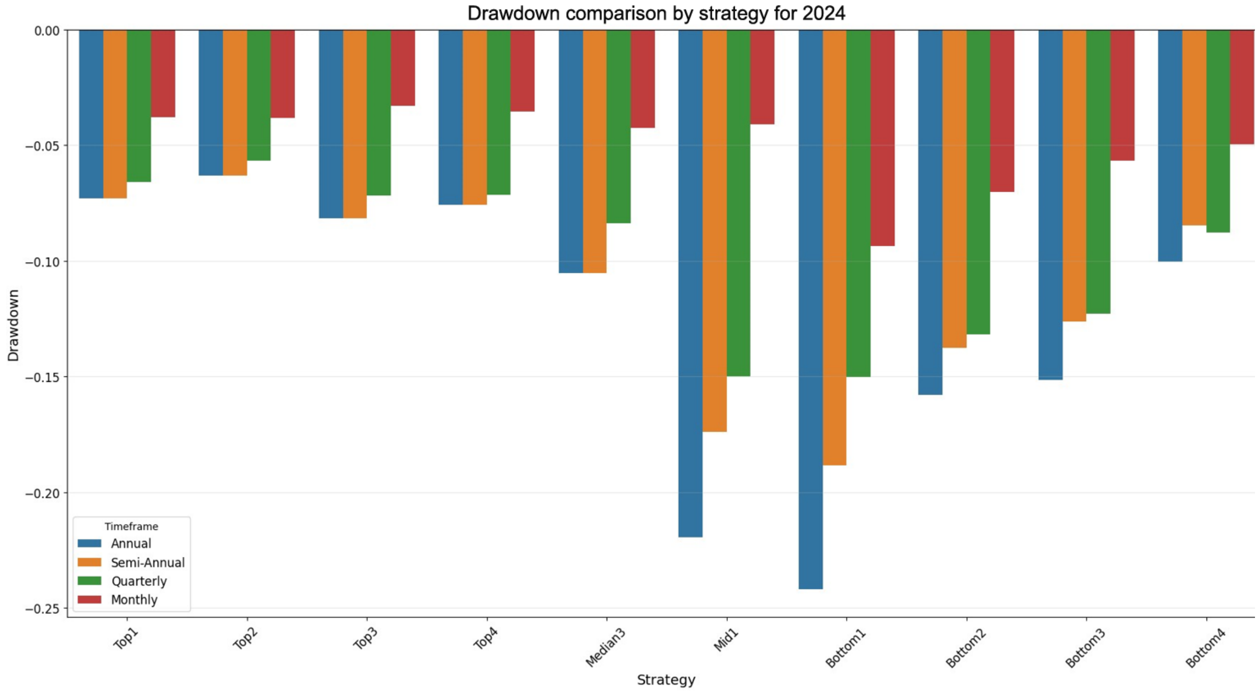

Figure 5. Maximum drawdowns by strategy type and rebalancing frequency. Shows peak-to-trough declines during worst historical periods.

Critical observations:

- Crisis Vulnerability: Top-1 strategies suffered 38% maximum drawdowns vs 29% for Top-3.
- Frequency Impact: Quarterly rebalancing showed 15% lower drawdowns than monthly.
- Recovery Periods: Bottom strategies required 2.3 × longer to recover than Top strategies.

5.4. Time Horizon Effects

The comparative results across rebalancing frequencies reveal the following patterns (see Figure 4):

- Short-term (Monthly): Delivered consistently higher outcomes for Top-1, Top-3, and Top-4 strategies, though more volatile in weaker portfolios (e.g., Bottom-1 and Bottom-2 strategies).
- Medium-term (Quarterly): Showed the strongest outperformance in Median-1 and Bottom-3 strategies, highlighting that quarterly rebalancing can capture momentum shifts effectively.
- Bi-yearly: Provided stable gains across Median-1 and Median-3, with generally better balance than annual strategies but less upside than quarterly rebalancing.
- Long-term (Annual): Offered the least advantage overall, with performance close to buy-and-hold in many cases, except for relative resilience in Bottom-2 and Bottom-4.

These results suggest that while maximal concentration (Top-1) delivers the highest potential returns under monthly rebalancing, a Top-3 strategy with quarterly rebalancing provides the most efficient balance of performance and risk management. Moreover, quarterly rebalancing also enhances outcomes in Median and Bottom strategies, underscoring its robustness as a preferred implementation for most investors when considering both return potential and downside control (Aggarwal, 2015).

6. CONCLUSION

This study demonstrates the effectiveness of data mining techniques (Han et al., 2011) in uncovering actionable investment insights from FTSE sector data. The analysis of 10 distinct strategies across multiple rebalancing frequencies reveals that systematic application of statistical learning methods (Hastie et al., 2009) can significantly enhance portfolio performance. The optimal Top 3 quarterly rebalancing strategy, generating 97% outperformance over buy-and-hold, exemplifies how modern analytical approaches (Provost & Fawcett, 2013) can be successfully applied to financial markets. These findings contribute to the growing body of literature on data-driven investment strategies and provide practical frameworks for portfolio optimization using quantitative methods (Shmueli et al., 2020).

REFERENCES

Aggarwal, C. C. (2015). *Data mining: The textbook*. New York: Springer.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.

Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, 33*, 1-22. https://doi.org/10.18637/jss.v033.i01

Han, J., Kamber, M., & Pei, J. (2011). *Data mining: Concepts and techniques* (3rd ed.). Burlington, MA, USA: Morgan Kaufmann.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. New York: Springer.

McKinney, W. (2017). *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython* (2nd ed.). Sebastopol, CA: O'Reilly Media.

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. Cambridge, MA: MIT Press.

Provost, F., & Fawcett, T. (2013). *Data science for business*. Sebastopol, CA, USA: O'Reilly Media.

Shmueli, G., Bruce, P. C., Gedeck, P., & Patel, N. R. (2020). *Data mining for business analytics*. Hoboken, NJ: Wiley.

Tsay, R. S. (2010). *Analysis of financial time series* (3rd ed.). Hoboken, NJ: Wiley.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data mining: Practical machine learning tools and techniques* (4th ed.). Burlington, MA: Morgan Kaufmann.

Zhou, Z.-H. (2021). *Machine learning*. New York: Springer.