AN EMPIRICAL STUDY ON THE COMPARATIVE PERFORMANCE OF MACHINE LEARNING (RANDOM FOREST) AND ECONOMETRIC (ARIMA-GARCH) MODELS FOR PORTFOLIO OPTIMIZATION IN THE NIGERIAN EQUITIES MARKET

BY

UMEH, CHIEMEKA MICHAEL

2021/247424

DEPARTMENT OF STATISTICS

FACULTY OF PHYSICAL SCIENCES

UNIVERSITY OF NIGERIA, NSUKKA

DECEMBER, 2025

AN EMPIRICAL STUDY ON THE COMPARATIVE PERFORMANCE OF MACHINE LEARNING (RANDOM FOREST) AND ECONOMETRIC (ARIMA-GARCH) MODELS FOR PORTFOLIO OPTIMIZATION IN THE NIGERIAN EQUITIES MARKET

BY

UMEH, CHIEMEKA MICHAEL

REG. NO.: 2021/247424

A PROJECT REPORT SUBMITTED TO THE DEPARTMENT OF STATISTICS, IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE BACHELOR OF SCIENCE (B.Sc) DEGREE IN STATISTICS, UNIVERSITY OF NIGERIA, NSUKKA.

December, 2025

CERTIFICATION PAGE

This is to certify that I, Umeh, Chiemeka Michael with reg no: 2021/247424 carried out this project report under the supervision of Dr. U. C. Nduka and that this work has not been previously submitted in part or as a whole for the award of any degree in this University or any other University.

Umeh, Chiemeka Michael **Date**

APPROVAL PAGE

This project report has been approved for the award of Bachelor of Science (B.Sc.) degree in Statistics, University of Nigeria, Nsukka.

______________________________ ______________________________

Dr. U. C. Nduka **Date**

______________________________ ______________________________

Dr. A. V. Oladugba **Date**

______________________________ ______________________________

External Examiner **Date**

ACKNOWLEDGMENTS

I sincerely appreciate Almighty God for giving me the strength, wisdom, and perseverance to successfully complete this project. My heartfelt gratitude goes to my supervisor, Dr. U. C. Nduka, whose guidance and corrections have been the foundation of this work.

I am equally grateful to the Head of Department, Dr. A. V. Oladugba, who also served as my academic advisor, for her constant support, advice, and dedication to my academic growth. I also extend my sincere thanks to all my lecturers for their commitment and for the knowledge they have imparted throughout my study.

I will forever be grateful to my parents and family for their love, prayers, and sacrifices, and to my friends and colleagues for their encouragement and companionship. Finally, I want to thank me for believing in me.

ABSTRACT

In the Nigerian stock market, investors and stockholders often face the problem of selecting the right stock to buy or hold while determining the possible risks of each asset. These problems come from the fact that most investors are uncertain or cannot determine if the value of an asset will appreciate or depreciate in the future. There is also the problems of selecting the best models for forecasting these market structures, which are the bedrock of building a good portfolio. This study aims to handle these problems by using both an advanced machine learning model and econometric (statistical) models to perform a comparative analysis and also using these models to implement portfolio optimization techniques with different strategies. The specific models used in this study are the Random Forest (machine learning), the ARIMA and GARCH (Statistical), the Random Walk, and the Hybrid model selection. This study used 19 Nigerian stocks from 5 different sectors, covering data from 2022 to 2025. Four strategies were used in this study namely: Mean-Variance Optimization (MVO), Minimum Variance Portfolio (MVP), Risk Parity Portfolio (RPP), and Equal Weight Portfolio (EWP). This study finalizes that using machine learning models (Random Forest) for volatility forecasting offers an advantage in capturing complex market movements, while econometric models (ARIMA-GARCH) remain relevant for forecasting returns. The results shows that the Random Forest model combined with the Mean-Variance optimization (MVO) strategy was the best in terms of risk-adjusted return.

TABLE OF CONTENTS

List of Figures

List of Tables

CHAPTER ONE

INTRODUCTION

1.1 Background Of The Study

Portfolio optimization and risk management represents an important challenge in modern finance. The theory of maximizing returns while minimizing risk, first formalized by Markowitz (1952) through Modern Portfolio Theory (MPT), is heavily dependent on forecasting asset returns and volatility. Financial modeling and forecasting approaches has significantly progressed demanding more sophisticated methods. Normally, financial time series relied heavily on traditional approaches like ARIMA and GARCH. The Autoregressive Integrated Moving Average (ARIMA) framework developed by Box and Jenkins (1976), and the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models developed by Bollerslev (1986).

These models have been able to successfully capture linear patterns and volatility clustering in most cases. However, these methods often rely on assumptions of linearity and heteroskedasticity that are frequently violated in emerging market structures. Looking at the Nigerian equities market, operating through the Nigerian Exchange Limited (NGX), it represents one of Africa's largest financial markets but it is still characterized by extreme fluctuations and high volatile movements, and these movement can be caused by Macroeconomics Policy Factors, Foreign exchange and Currency Factors, Political and Institutional Factors, etc. These characteristics requires advanced and more complex approaches, and that is where the machine learning models comes in. Machine learning models like Random Forest (Breiman, 2001) can model these complex non-linear patterns without overfitting or any restrictive assumptions, these abilities are what makes machine learning models advanced and suitable for the Nigerian market. The portfolio optimization theory has evolved beyond Markowitz's Mean-Variance theory, though his work is seen as the the stepping stone for other researchers to develop other optimization techniques such as the Minimum Variance Portfolio (MVP) which is like an extension of his original work, Risk Parity portfolio (Maillard, Roncalli, & Teiletche, 2010), and the Equal Weight Portfolio (EWP) (DeMiguel, Garlappi, & Uppal, 2009). The performance of these portfolio optimization techniques critically depends on the input forecasts which in this case will be generated by the models.

1.2 Statement Of The Problem

The essential issue is that there is a noticeable gap in the methodologies available in the existing literature on the Nigerian stock market. Although there are works that have utilized ARIMA-GARCH models, such as Adediran & Adeleke (2013), and machine learning models, as demonstrated in Oguntade & Adeyemi (2019), none of these works aimed to specifically compare these models' predictive accuracy on Nigerian stock markets and more importantly applied the predictions to portfolio optimization, to assess which approach provides the best possible investment results. Accordingly, the existing issue for investors and portfolio managers is that there is no proven data-driven solution available to address the following essential question: For portfolio formation in the Nigerian market, should one go ahead with econometric models or move towards machine learning models?

1.3 Aim and Objectives Of The Study

The aim of this study is to compare the performance of a Machine Learning model (Random Forest) and an Econometric model (ARIMA-GARCH) for portfolio optimization in the Nigerian equities market. This will be achieved by applying these models, alongside a Random Walk model, to forecast stock returns and volatility, and subsequently using these forecasts to apply: Mean Variance Optimization (MVO), Minimum Variance Portfolio (MVP), Risk Parity Portfolio (RPP), and Equal Weight Portfolio (EWP) for portfolio optimization. The objectives of the proposed study are:

1. To apply the Autoregressive Integrated Moving Average - Generalized Autoregressive Conditional Heteroskedasticity (ARIMA-GARCH) model, the Random Forest (RF) model, and the Random Walk (RW) model to forecast the returns and volatility of selected stocks in the Nigerian equities market.
2. To apply the Mean Variance Optimization (MVO), Minimum Variance Portfolio (MVP), Risk Parity Portfolio (RPP), and Equal Weight Portfolio (EWP) for portfolio optimization.
3. To verify and compare the efficiency of models and strategies with respect to a basic benchmark by means of the Sharpe Ratio.
4. To offer portfolio suggestions and recommendations to Nigerian investors, using the results.

1.4 Significance Of The Study

This research work is filled with various academic theories, the most especially to the stock investors and managers of the portfolios, as well as to the NGX and the Regulators.

1. **For Investors and Portfolio Managers:** The current study offers critical insight on how investors can choose models or methods of optimizing investment decisions by applying machine learning models. The significance of this study is that it indicates that machine learning models have superior performance in terms of risk-adjusted return (Sharpe Ratio), risk control, and overall risk management compared to traditional models.

2. **For the Nigerian Exchange (NGX) and Regulators:** This research gives answers that would aid in market dynamics, market volatilities, and market liquidity issues. The manner in which various models estimate market dynamics could be useful for the formulation of policies that would ensure the stability of the market.

1.5 Scope Of The Study

In relation to the study, the analysis is conducted on 19 listed and actively traded stocks in Nigeria, and they include those that belong to the categories of Banking, Consumer Goods, Industrial Goods, Telecommunications, and the Oil & Gas sectors. The reason for selecting a vast period for the data collection, which is between March 2022 and August 2025, is to ensure that the data is equally available for analysis for all the stocks selected for the study. In addition, performance measures are evaluated using the Sharpe Ratio.

1.6 Literature Review

1.6.1 Theoretical Foundations: Market Efficiency and Predictability

The question of whether stock markets can predict themselves is the very essence of this study. The question is fundamentally linked to the Efficient Market Hypothesis (EMH), which is challenged by behavioral finance theories. The Efficient Market Hypothesis, proposed formally in Fama (1970), states that ”financial markets areinformationally efficient.” This implies that stock market prices already contain all available information available at any point of time. There are three variants of this hypothesis – weak form, semi-strong form, and strong form, of which weak form efficiency is of primary interest in predictive analysis models. The weak form of this hypothesis implies that ”past stock price records will not be useful in advancing predictions of future stock price movements; that is, stock prices change unpredictably, moving like a random walk." This hypothesis directly implies that it is only sensible to test models like Random Walk on prediction analysis on account of this lack of predictive content of past stock performance onto future performance." However, whether this theory holds in emerging markets, in the sense that it is applicable to the Nigerian market, has been questioned in many studies. It has been found that there are patterns that contravene the random walk theory in developing markets. Many findings in studies conducted in the economies that fit this category reveal that there is enough inconclusive but questionable evidence regarding market efficiency in the Nigerian market. Many studies, including those conducted by Olowe (2009) and Adediran and Adeleke (2013), reveal that the random walk hypothesis does not really occur in the Nigerian Stock Exchange. Markets that are inefficient, lack information, and are characterized by behavioral aspects among the Nigerian investors are held responsible for this aspect. However, the most significant threat to the EMH was posed by the emergence of behavioral finance theories, which included the effects of psychological aspects and the notions of bias in decision-making. Some of the most important works of Kahneman and Tversky (1979) and others have shown that investors exercise inefficient choices at times because of certain psychological notions like the effects of overconfidence, herds, and loss aversion. Such kind of behavioral aspects can create a certain predictable price movement which will be predictable using sophisticated modeling. This shows the important research question being answered in this research study regarding the specific topic of a major challenge between the theories of the EMH and the emergence of behavioral finance.

1.6.2 Econometric Forecasting Models: The ARIMA-GARCH Models

Econometric time series models have been traditionally developed for financial forecasting, and the ARIMA-GARCH model has been the dominant method used under this area of financial econometrics. The Autoregressive Integrated Moving Average (ARIMA) method, proposed by Box and Jenkins (1976), provides a structured way of developing and forecasting time series variables. In the ARIMA model, the procedure has three major parts, namely the autoregressive (AR), moving average (MA), and the integration (I) part, which handles differencing of non-stationary data, often encountered in financial time series data. The Box-Jenkins methodology has been noted for its structured way of working with model formulation and testing, and this has remained the dominant method for finance model formulation and validation for the past few decades. There have been successful applications of ARIMA modeling in the Nigerian stock markets as well. Researches such as Adediran and Adeleke (2013) illustrated ARIMA's capability to identify short-run trends within Nigerian stock markets, admitting at the same time that it is insufficient to capture the level of volatility clustering typical of emerging markets. In recent applications of ARIMA, as presented by Adhikari, Agrawal, and Singh (2024), also continue to show ARIMA's helpfulness while stressing on the fact that it needs to be combined with volatility modeling in order to get a complete forecast. The special variance models were developed based on the phenomenon observed in financial markets, where volatility clustering-the tendency for periods of high volatility to persist-led to the development of special variance models. It was the major contribution of Engle (1982) that gave the birth of the Autoregressive Conditional Heteroscedasticity model commonly known as ARCH. Based on it, Bollerslev (1986) developed GARCH model. This GARCH model comprises volatility persistence-an autoregressive component reflecting past volatility and an equal moving average capturing the influence of past shock series. Its robustness in GARCH models lies in the applicability it has in forecasting volatility in risk management applications. The Nigerian market analysis has always been supported by the applicability of GARCH models as asserted by evidence from Olowe (2009), who found the existence of vetoatility persistency on the Nigerian Stock Exchange, with some studies evaluating various forms of GARCH models suited for thespecific market conditions prevailing on the Nigerian market. The use of both ARIMA models on returns and GARCH on volatility solves all analytical objectives and therefore qualifies as one top benchmarking tool by which new approaches also have to be validated as effective.

1.6.3 Machine Learning Forecasting Model: The Random Forest

The recent development in machine learning is a paradigm shift in financial forecasting, providing solutions that could identify non-linear relationships not accounted for by conventional methods. In machine learning, one technique that gained popularity because of its advantageous characteristics and underlying theory from ensemble learning is the Random Forest developed by Breiman (2001). The technique involves creating several decision trees during the training phase. There are several advantages in grouping several decision trees: it prevents overfitting by aggregating the forecasts from several trees, it incorporates feature selection analysis, and it is robust to outliers and noisy observations, all important characteristics in emerging markets. The principles of the Random Forest algorithm include the concept of bagging and the selection of features in a random manner. The two elements work in a complementary manner, giving the outcome of a number of differently produced forests, thus providing a better predictive forecast than that given by the individual forest. This plays a critical role in the case of financial models, in terms of a number of factors (macroeconomic factors, market sentiments, fundamentals, and the global market, to name a few) all working in a non-linear manner. This has been proven in various research papers, including Khaidem, Saha, and Dey (2016), that the Random Forest has the capability to do this. Although machine learning applications are new to the Nigerian stock markets, they have proved to be very effective. According to Oguntade and Adeyemi (2019), ensemble methods, including Random Forest, have shown greater accuracy than traditional econometric methods. According to studies carried out by Singh and Singh (2021), comparative studies of emerging markets have also shown that machine learning methods, including Random Forest, tend to be more accurate. Singh and Singh (2021) also add that although machine learning is more accurate, careful preparation of data features as well as parameters is significant. This is because machine learning methods are data-driven, meaning that they make minimal assumptions about data patterns.

1.6.4 Benchmark Model: Random Walks

The Random Walk Hypothesis is a basic standard in financial literature in the area of financial forecast, based on the Efficient Market Hypothesis, first introduced by Fama (1970). The hypothesis states that stock prices reflect all the information available, and stock prices become unpredictable, with a sequence randomly following its path. The validity of the hypothesis in the case of the Nigerian market has also been amply verified. The study has also examined the case of Nigeria, where Olowe (2009) as well as Adediran and Adeleke (2013) obtained evidence suggesting a lack of compliance with the random walk test, suggesting that there are Market Inefficiencies. The Random Walk Hypothesis, on the other hand, plays a critical role as a test hypothesis regarding the forecasting capabilities of an advanced forecasting model. The presence of the Random Walk Hypothesis as a baseline hypothesis ensures that the forecasting hypothesis tested has practical application and use in the investment decision, as it has the ability to perform better than the hypothesis of zero returns.

1.6.5 Portfolio Optimization Theories and Strategic Frameworks

1.6.5.1 Theoretical Foundations of Modern Portfolio Optimization

Theoretical underpinning for the systematic management of portfolios has been made possible by the ingenious contribution of the Modern Portfolio Theory, presented by the pioneering contribution of Markowitz in the year 1952, with the paper: “Portfolio Selection,” by Markowitz, published in Markowitz (1952). It introduced the entire investment paradigm in an altogether new form by the pioneering investment concept of evaluating stocks, not as separate entities, but as the influencer of the attributes of the overall investment portfolio.

1.6.5.2 Mean-Variance Optimization (MVO)

The Mean-Variance Optimization method is the theoretical application of the findings and concepts proposed by Harry Markowitz. In terms of theoretical underpinning, Mean-Variance Optimization attempts to find the optimal portfolio solution that maximizes return for a specified risk tolerance, or equivalently, minimizes risk for an specified level of return. The greatness of the Mean-Variance Optimization theory is that it created this thing called the ”efficient frontier,” which basically describes the best possible risk versus return of the market with various portfolios. MVO is theoretically sound; however, empirical studies indicate that there are many practical complications in its application. The study conducted by DeMiguel et al. (2009) established that this framework is quite sensitive to variables, especially those concerning estimates of return expectations. A minor difference in variables may cause great variability in the optimal weighting of investments as well as under-performance of this approach in practical applications outside the estimated environment, especially in emerging markets such as Nigeria where return distributions show notable departures from normal distributions due to difficulties in estimating variables owing to inefficiencies in financial markets.

1.6.5.3 Minimum Variance Portfolio (MVP)

The Minimum Variance Portfolio is an important milestone in the development of portfolio optimization solutions that overcome some flaws present in traditional MVO. The Minimum Variance Portfolio is strategic in that it aims at risk minimization alone without considering return estimates. The underlying strategy is theoretically informed by an empirical observation that risks are much more stable than return estimates, and hence minimum variance portfolios are very attractive in uncertain markets. The philosophical underpinning of MVP lies in the fact that it attempts to economically assure portfolio volatility reduction and, in the process, provides a systematic way to improve the performance and preservation of capital in any portfolio, especially in times of adverse markets. Many research studies carried out in other markets have concluded that minimum variance portfolios are, in fact, competitive and often come with the disadvantage of with less upside during the most successful phases of the markets. In emerging markets such as Nigeria, the need to predict markets accurately makes MVP sound as an alternative to return-optimized methods.

1.6.5.4 Risk Parity Portfolio (RPP)

It symbolizes a paradigm shift in the thinking process during portfolio optimization, wherein the focus in optimization evolved from managing an allocation in wealth to managing an allocation in risk. The theoretical framework surrounding Risk Parity, appropriately stated in Maillard et al. (2010), is to ensure that the risk contribution profile of every asset class/security in the portfolio be equalized. It goes against the traditional risk concentration in equity-invested portfolios. This would constitute the theoretical advantage, that is, a better risk management framework, which would thereby make way for smoother performance in various market conditions. By ensuring that the non-universal risk of no specific asset class is monopolized, the Risk Parity method somewayays acquires a better stability in the returns, in addition to exhibiting a better diversification effect. This ideology has already got immense popularity with the top investment agencies in the world, because of the success obtained in the execution of the model by investment agencies like Bridgewater Associates. For a developing market, the Risk Parity model has immense attraction in terms of Concentration Risk management, since the equity markets in developing nations experience huge volatile nature.

1.6.5.5 Equal Weight Portfolio (EWP)

The Equal Weight Portfolio is the simplest diversification strategy and is thus a key behaviorally and theoretically meaningful benchmark in portfolio optimization problems. From a theoretical perspective, the EWP assigns equal amounts of wealth to all available assets, with the use of optimization and estimation methods. Theoretical meaning and importance of EWP were aptly illustrated by DeMiguel et al. (2009) in their systematic study "Optimal Versus Naive Diversification." Ironically, their work revealed that in out-of-sample analysis, the 1/N portfolio regularly matches and even surpasses the performance of optimization methods. Theoretically, this phenomenon has been explained by the absolute robustness of EWP against parameter estimates and model specification errors. In emerging markets such as Nigeria, where risks associated with parameter uncertainty are high and inputs necessary for optimization are difficult to estimate reliably, the performance differential between naive and optimization methods is an empirical issue that the work aims to shed light on.

1.6.5.6 Strategic Integration and Research Context

The increasing complexity of portfolio optimization techniques, from simple MVO to risk-based models, represents the continued evolution of portfolio management theory itself. The differences between these models represent varying philosophies about managing core portfolio problems: while MVO seeks optimal risk/return trade-offs, MVP seeks risk management, RPP seeks risk balancing, and EWP seeks the optimal, estimation-free diversification solution. With regard to the Nigeria equity market in particular, there is still an unperused aspect in academic literature in terms of the relative performance of the aforementioned methods. The special characteristics of the equity market in Nigeria, such as high persistence in market volatility, industry dominance, and special market microstructure conditions like different behavioral patterns, could play an important role in determining the relative merits of different optimization algorithms. The present work aims to fill an important niche by undertaking an overall comparison analysis among those methods combined with robust forecast algorithms.

1.6.6 Research Summary and Identified Gap

A look through these related fields reveals the immense strides in the techniques of financial forecasting and portfolio optimization, as well as the lacuna still existing on how these techniques have been applied to concrete market contexts. In this regard, the theoretical development from efficient market assumptions to behavioral finance, methodological progress from linear econometric models to nonlinear machine learning approaches, and strategic development from simple mean-variance optimization to sophisticated risk-based allocation have, no doubt, constituted a milestone in financial knowledge. Nevertheless, joining these fields of research points toward a critical research gap, one which this research tries to fill. There is, on the other hand, a fair amount of research on forecasting methods and portfolio optimization strategies separately, but their complete integration in the context of emerging African markets has not been thoroughly pursued thus far. Most studies dealing with the performance evaluation of forecasting models are taken individually from those dealing with portfolio construction applications. On the other hand, research into portfolio optimization normally employs simple return and risk assumptions. The Nigerian market context exhibits special characteristics-high volatility persistence, structural breaks, limited trading in some sectors, and unique investor behavior-that may significantly affect the relative performance of different strategies. The current study contributes to the body of knowledge by offering a comprehensive practical comparison of the conventional econometric multiple forecasting techniques (ARIMA-GARCH), the contemporary machine learning algorithm (RF), the simplest benchmark approach (RW), and an original combination with portfolio optimization techniques (MVO, MVP, RPP, EWP) by incorporating the most up-to-date information about the Nigerian stock market. By extensively evaluating each combination of the aforementioned techniques as part of comprehensive real-world tests and financial metrics, the study offers important scholarly and practical findings for optimal portfolio formation techniques specifically tailored for the Nigerian market environment, which has remained an important void within the area of scholarly studies.

CHAPTER TWO

DATA COLLECTION AND METHODOLOGY

2.1 Introduction

This chapter will incorporate the methodology used within the study from the data collection procedure to the design and optimization of the portfolio model. It consists of four broad sections namely data collection and description, methodological framework, portfolio optimization techniques, and evaluation criteria.

2.2 Data Collection and Description

2.2.1 Data Source and Selection Criteria

A total of 19 stocks from Nigeria were sourced from the website “www.investing.com” from March 2022 to August 2025. This contains 774 trading days.

2.2.2 Data Cleaning and Transformation

The data were sorted and cleaned using the pandas library in Python, then transformed into daily log returns using the formula:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \times 100\% \tag{2.1}$$

where:

- r_t: Logarithmic return series at time t
- P_t denotes the closing price

2.2.3 Data Splitting

The data was split into training and testing sets, the training set from March 2022 to August 2024 and the testing set from September 2024 to May 2025.

2.3 Methodological Framework

2.3.1 Stationarity Testing

The stationarity test was carried out on the data because it is an important component in time series analysis. The stationarity test reveals whether the data should be differenced. Using the ADF test on the data, it is revealed that all the returns are stationary at the significance level of 5%.

2.3.2 Model Specification

2.3.2.1 Autoregressive Integrated Moving Average (ARIMA)

The ARIMA model was selected for forecasting the returns because of its effectiveness in modeling linear dependencies in time series. The model structure follows the Box-Jenkins methodology:

$$\phi(B)(1-B)^d r_t = c + \theta(B)\epsilon_t \tag{2.2}$$

where:

- r_t: Logarithmic return series at time t
- $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$: Autoregressive (AR) polynomial of order p
- $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$: Moving Average (MA) polynomial of order q
- $(1-B)^d$: Integration term representing d-th order differencing for stationarity

- B: Backshift operator defined as $Br_t = r_{t-1}$
- c: Constant term (intercept) in the mean equation
- ϵ_t: White noise error term at time t, assumed to be independently and identically distributed

Optimal (p, d, q) orders were determined using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to balance model fit and complexity.

2.3.2.2 Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

The GARCH model was opted because it is capable of modeling volatility clustering, especially in the case of financial time series. This is something indispensable in the generation of correct risk inputs for portfolio optimization. The specification of the GARCH(p, q) for the conditional variance, σ_t^2, is given generally by:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{2.3}$$

Where:

- σ_t^2: The conditional variance at time t.
- ω: The unconditional variance component.
- α_i: The ARCH parameters $(i = 1, \ldots, p)$, which measure the impact of previous shock squared (ϵ_{t-i}^2) on current volatility.
- β_j: The GARCH parameters $(j = 1, \ldots, q)$, which measure the persistence of past volatility (σ_{t-j}^2).
- ϵ_{t-i}: The squared residual from the mean equation at time $t - i$.
- p: The number of lagged squared errors.
- q: The number of lagged conditional variances.

The (p, q) orders of the model were selected using two Information Criteria: the Akaike

Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

2.3.2.3 Random Forest Model

The Random Forest model is an advanced machine learning model that has the characteristic of handling both linear and non-linear data. Even though this particular model is non-linear, it has been chosen for this assignment as it has the ability to capture the pattern that the ARIMA or the GARCH model may fail to capture. The Random Forest model is an ensemble learning method that makes use of the output of multiple decision trees for forecasting. For the input vector **x** that has predictive features, the final forecast is given by:

$$\hat{f}(\mathbf{x}) = \frac{1}{K}\sum_{k=1}^{K} T_k(\mathbf{x}) \tag{2.4}$$

where:

- $\hat{f}(\mathbf{x})$: Is the return or volatility forecast for input **x**
- K: Is the total number of decision trees in the forest ($K = 100$)
- $T_k(\mathbf{x})$: Is the prediction from the k-th individual decision tree
- **x**: Is the Vector of input features

2.3.2.4 Random Walk Model

It is a naive model used for the purposes of comparison, which serves as a benchmark, based on the Weak-Form Efficient Market Hypothesis. It assumes that price movements cannot be predicted using the past. The model of forecasting daily returns is :

$$\hat{r}_{t+1} = \mu + \epsilon_t \tag{2.5}$$

where:

- $\hat{r}_{t+1}$: Is the expected return for period $t + 1$
- μ: Is the drift parameter that is assumed to be zero for daily returns ($\mu = 0$)
- ϵ_t: Is the random shock term following a white noise process

This model provides a benchmark that every complex model must pass to show predictive power.

2.3.2.5 Hybrid Forecasting Framework

It is a combination of models that is arrived at by choosing the best performing models from those previously shown, for each stock based on the performance measure of Root Mean Squared Error (RMSE). The process of choosing works in this manner:

$$M_i^* = \arg\min_{M \in \mathcal{M}} \text{RMSE}_i(M) \tag{2.6}$$

where:

- M_i^*: Is the optimal model selected for stock i
- $\mathcal{M}$: Is the set of candidate models {ARIMA-GARCH, Random Forest, Random Walk}
- $\text{RMSE}_i(M)$: Is the Root Mean Squared Error of model M for stock i during validation

2.3.3 Portfolio Optimization Framework

2.3.3.1 Mean-Variance Optimization (MVO)

This optimization strategy was used to maximize expected returns for given risk levels:

$$\max_{\mathbf{w}} \mathbf{w}^\top \boldsymbol{\mu} \quad \text{subj to} \quad \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w} \leq \sigma^2_{\text{target}} \tag{2.7}$$

where:

- $\mathbf{w}$: Is the vector of portfolio weights $[w_1, w_2, \ldots, w_N]^\top$
- $\boldsymbol{\mu}$: Is the vector of expected returns from forecasting models $[\mu_1, \mu_2, \ldots, \mu_N]^\top$
- $\boldsymbol{\Sigma}$: Is the covariance matrix of the asset returns $\{\sigma_{ij}\}_{N\times N}$
- σ^2_{target}: Is the target portfolio variance constraint
- N: Total number of assets in the portfolio

This strategy uses forecasting model outputs into the optimization process, using predicted returns for $\boldsymbol{\mu}$ and GARCH volatilities for $\boldsymbol{\Sigma}$.

2.3.3.2 Minimum Variance Portfolio (MVP)

This strategy focuses on risk minimization:

$$\min_{\mathbf{w}} \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w} \quad \text{subj to} \quad \mathbf{w}^\top \mathbf{1} = 1, \quad w_i \geq 0 \ \forall i \tag{2.8}$$

where:

- $\mathbf{w}$: Is the vector of portfolio weights
- $\boldsymbol{\Sigma}$: Is the covariance matrix of asset returns
- $\mathbf{1}$: Is the vector of ones $[1, 1, \ldots, 1]^\top$ enforcing full investment
- w_i: Is the weight of asset i, constrained to be non-negative

This strategy is more relevant in volatile markets where risk prediction tends to be more reliable than return prediction.

2.3.3.3 Risk Parity Portfolio (RPP)

This strategy tries to equalizes risk contributions across assets, addressing concentration risk in optimization. It is given by:

$$\min_{\mathbf{w}} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(w_i(\mathbf{\Sigma}\mathbf{w})_i - w_j(\mathbf{\Sigma}\mathbf{w})_j\right)^2 \tag{2.9}$$

where:

- $\mathbf{w}$: Is the vector of portfolio weights
- $\mathbf{\Sigma}$: Is the covariance matrix of asset returns
- $w_i(\mathbf{\Sigma}\mathbf{w})_i$: Is the risk contribution of asset i to total portfolio risk
- N: Is the total number of assets in the portfolio

2.3.3.4 Equal Weight Portfolio (EWP)

Here, there is no optimization process involved and it is carried out by distributing the weight evenly to all assets. It can be used to set a low performance standard for comparison with the other optimization methods.

$$w_i = \frac{1}{N} \quad \forall i \in \{1, 2, \ldots, N\} \tag{2.10}$$

where:

- w_i: Is the weight of asset i in the portfolio
- N: Is the total number of assets in the portfolio

2.3.3.5 Portfolio Return and Risk Calculation

The different portfolio metrics were calculated using:

- **Portfolio Expected Return:**

$$\mu_p = \mathbf{w}^\top \boldsymbol{\mu} = \sum_{i=1}^{N} w_i \mu_i \tag{2.11}$$

where

  - μ_p is the portfolio expected return
  - $\mathbf{w}$ is the vector of portfolio weights
  - $\boldsymbol{\mu}$ is the vector of individual assets expected returns
  - N is the number of assets.

- **Portfolio Variance:**

$$\sigma_p^2 = \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w} = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma_{ij} \tag{2.12}$$

where

  - σ_p^2 is the portfolio variance
  - $\boldsymbol{\Sigma}$ is the covariance matrix of the asset returns
  - σ_{ij} is the covariance between asset i and j

- **Portfolio Volatility:**

$$\sigma_p = \sqrt{\mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}} \tag{2.13}$$

- **Sharpe Ratio:**

$$\text{Sharpe} = \frac{\mu_p - r_f}{\sigma_p} \tag{2.14}$$

where

  - r_f is the risk free rate but for this study is assumed to be zero
  - mu_p is the expected return of the portfolio
  - σ_p is the standard deviation of the portfolio's returns

2.3.4 Performance Evaluation

2.3.4.1 Forecast Accuracy Metrics

The RMSE was used as the accuracy metric.

2.3.4.2 Portfolio Performance Metrics

In the current study, the Sharpe Ratio was employed as the criterion for portfolio evaluation. Sharpe Ratio is a widely accepted indicator in the modeling of finance.

2.3.4.3 Statistical Testing

In this analysis, all the test of statistic was performed at a 5% significance level, that is, $\alpha = 0.05$. The null and alternative hypothesis followed by each test were given by:

Augmented Dickey-Fuller (ADF) Test:

- H_0: The time series is non-stationary
- H_1: The time series is stationary
- Test Statistic: $ADF = \frac{\hat{\gamma}}{SE(\hat{\gamma})}$
- **Decision Rule:** Reject H_0 if p-value < 0.05

ARCH Lagrange Multiplier (LM) Test for Heteroskedasticity:

- H_0: No ARCH effects present (homoskedasticity)
- H_1: ARCH effects present (heteroskedasticity)
- Test Statistic: $LM = n \cdot R^2 \sim \chi^2(m)$
- **Decision Rule:** Reject H_0 if p-value < 0.05

Jarque-Bera Test for Normality:

- H_0: Returns are normally distributed (skewness = 0, kurtosis = 3)
- H_1: Returns are not normally distributed

- Test Statistic: $JB = \frac{n}{6}(S^2 + \frac{(K-3)^2}{4})$
- **Decision Rule:** Reject H_0 if p-value < 0.05

Diebold-Mariano Test:

- H_0: Equal forecasting accuracy between two models
- H_1: Significant difference in forecasting accuracy
- Test Statistic: $DM = \frac{\bar{d}}{\hat{\sigma}_d/\sqrt{n}}$
- **Decision Rule:** Reject H_0 if $|DM| > 1.96$ (for $\alpha = 0.05$)

2.4 Model Selection Criteria

The best orders for the ARIMA & GARCH models were selected using these information criteria:

- **The Akaike Information Criterion (AIC):** $AIC = 2k - 2\ln(L)$
- **The Bayesian Information Criterion (BIC):** $BIC = k\ln(n) - 2\ln(L)$

where

- k is the number of parameters
- L is the likelihood
- n is the sample size

2.5 Software and Computational Tools

This research and analysis was conducted using Python 3.9+ on Google colab IDE with libraries like pandas, numpy, seaborn, matplotlib, arch, statsmodels and scikit-learn.

CHAPTER THREE

DATA ANALYSIS AND DISCUSSION OF RESULTS

3.1 Data Analysis

This chapter shall offer an in-depth analysis using the methodologies cited above in the previous chapter. In this section, the analysis of the data is shown, which is done in two ways: descriptive analysis and the correlation between the variables in this study.

3.1.1 Descriptive Statistics of the selected 19 Nigerian stock Returns

Table 3.1: Descriptive Statistics of Daily Log Returns for the selected 19 Nigerian Stocks (2022-2025)

Stock	Mean	Std. Dev.	Min	25%	Median	75%	Max
Access	0.0013	0.0281	-0.1534	-0.0094	0.0000	0.0115	0.0953
Bua	0.0011	0.0239	-0.1202	0.0000	0.0000	0.0000	0.0953
Cadbury	0.0024	0.0369	-0.1054	0.0000	0.0000	0.0000	0.1340
Dangote	0.0010	0.0221	-0.1054	0.0000	0.0000	0.0000	0.0953
DangSugar	0.0016	0.0351	-0.1178	-0.0055	0.0000	0.0054	0.0953
GTCO	0.0019	0.0237	-0.1272	-0.0066	0.0000	0.0096	0.0953
Guinness	0.0010	0.0251	-0.1054	0.0000	0.0000	0.0000	0.0953
Honeywell	0.0024	0.0436	-0.1054	-0.0157	0.0000	0.0190	0.0953
Lafarge	0.0023	0.0254	-0.1149	-0.0027	0.0000	0.0069	0.0953
MTN	0.0010	0.0232	-0.1054	0.0000	0.0000	0.0000	0.0953
Nestle	0.0004	0.0184	-0.2107	0.0000	0.0000	0.0000	0.0953
NigFlourMill	0.0028	0.0336	-0.1054	0.0000	0.0000	0.0000	0.0953
Oando	0.0031	0.0469	-0.2076	-0.0180	0.0000	0.0216	0.1896
PZ	0.0016	0.0394	-0.1462	0.0000	0.0000	0.0000	0.1525
Stanbic	0.0015	0.0270	-0.1054	0.0000	0.0000	0.0000	0.1895
Seplat	0.0023	0.0193	-0.1054	0.0000	0.0000	0.0000	0.0953
Sterling	0.0021	0.0352	-0.1054	-0.0131	0.0000	0.0147	0.0953
Total	0.0011	0.0164	-0.1054	0.0000	0.0000	0.0000	0.0953
Zenith	0.0015	0.0244	-0.1370	-0.0075	0.0000	0.0105	0.0953

The descriptive statistics show some important characteristics of the Nigerian equity returns.All the stocks showed positive average daily returns ranging from 0.04% for Nestle stock to 0.31% for Oando stock, indicating the overall optimistic behavior of stocks during the period under investigation. Volatilities of stocks varied and showed different values ranging from 1.64% for Total stock to 4.69% for Oando stock, indicating the risk attached to various sectors. The analysis shows that the median returns for most stocks are zero, which is common in markets with low liquidity. The presence of significant negative returns across all stocks, particularly Nestle (-21.07%) and Oando (-20.76%), tells a story of some extreme negative movements, which is a characteristic of an emerging market like the Nigeria Stock exchange.

3.1.2 Time Series Plot of Daily Log Returns for the 19 Nigerian Stocks (2022-2025)

Following the descriptive statistics summary in Table 3.1, the daily log returns for all 19 Nigerian equities are shown below in Figures 3.1 and 3.2. Some of these time series plots below show volatility clustering which is a period of high fluctuation followed by similar periods, frequent zero return days consistent with low market liquidity, and extreme movements which occurs occasionally, notably in stocks like Nestle. These visual patterns provide proof for employing models, such as GARCH and Random Forest, that are specifically designed for volatility and extreme movement modeling.

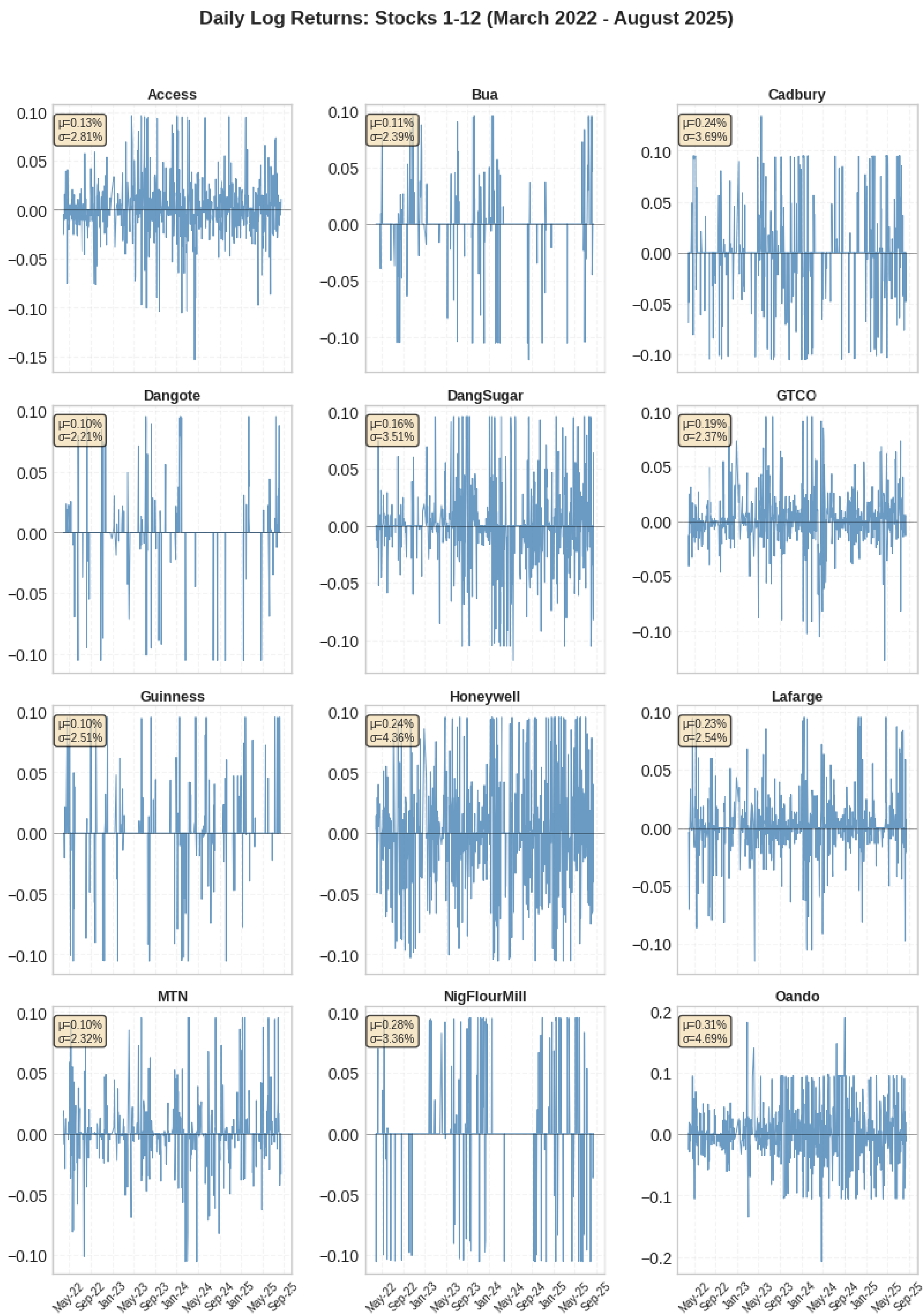

Figure 3.1: Time Series of Daily Log Returns for Nigerian Stocks (Part 1)

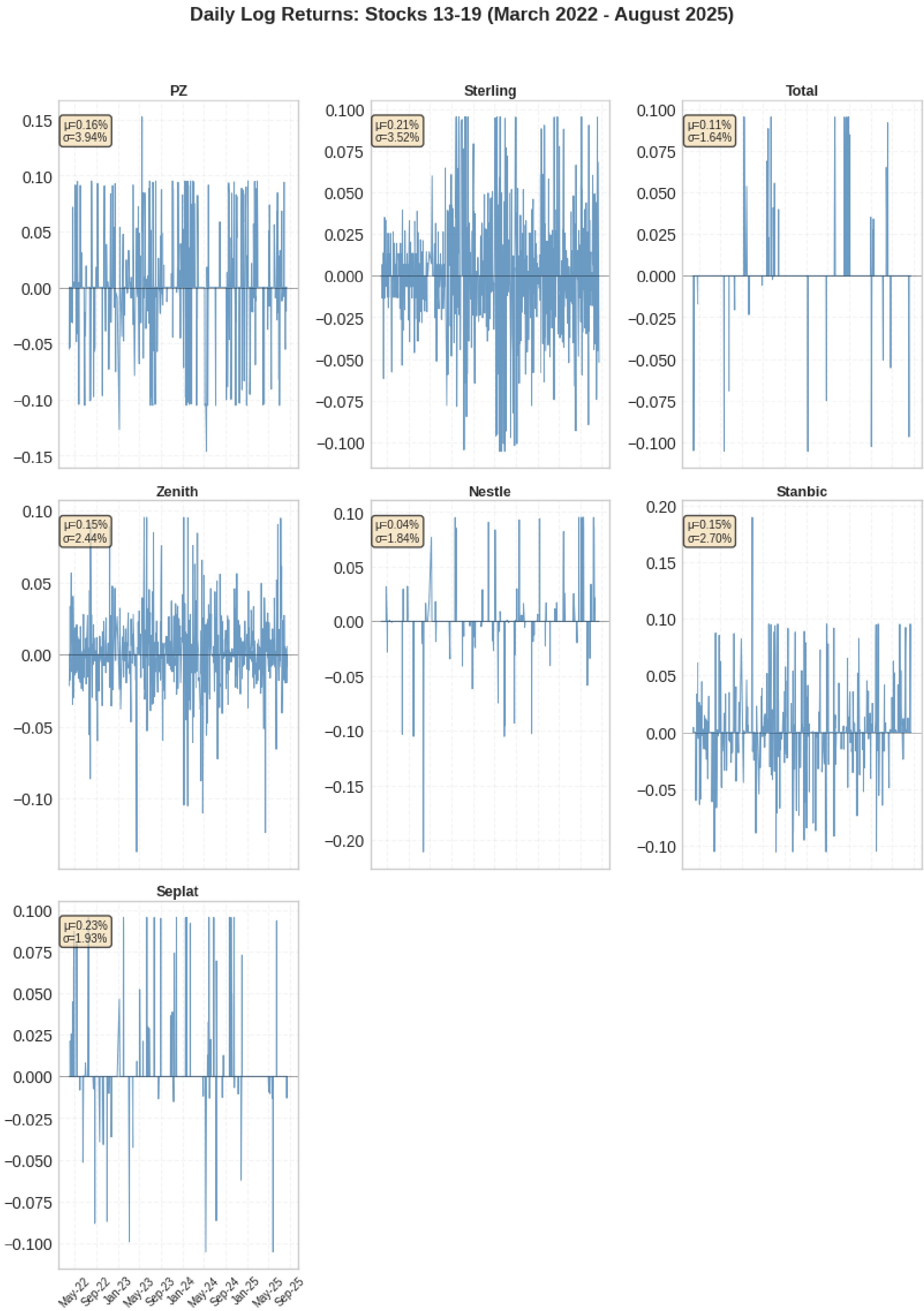

Figure 3.2: Time Series of Daily Log Returns for Nigerian Stocks (Part 2)

3.1.3 Correlation Analysis

correlation_heatmap.png

Figure 3.3: Correlation Matrix Heatmap of Nigerian Stock Returns

The correlation analysis shows the relationships between different stocks and sectors, as demonstrated in Figure 3.3. The heatmap shows correlation coefficients between all 19 stocks, with colors representing correlation strength from strong negative (blue) to strong positive (red).

Table 3.2: Average Correlation Coefficients By Sector

Sector	Banking	Consumer Goods	Industrial	Oil & Gas	Telecom
Banking	0.435	0.067	0.169	0.071	0.148
Consumer Goods	0.067	0.044	0.063	0.017	0.037
Industrial	0.169	0.063	0.077	0.006	0.069
Oil & Gas	0.071	0.017	0.006	0.058	0.045
Telecom	0.148	0.037	0.069	0.045	N/A

The correlation analysis by sector confirms the diversification as seen in the heatmap. correlations between sectors are low, ranging from 0.006 (Industrial-Oil & Gas) to 0.169 (Banking-Industrial). This low correlation is great for portfolio optimization. Several key patterns are observed from the correlation structure:

- Strong Positive Correlations: These correlations are found to be highest within the banking industry, namely between GTCO and Zenith (0.586) as well as Access and Zenith (0.516). This indicates that these banks have moved similarly due to monetary policy shifts.
- Low and Negative Correlations: There exist stocks that display a low level of correlations. For instance, Lafarge and Seplat display a slightly negative correlation coefficient of -0.006 between the Industrial and Oil & Gas sectors. Others include Honeywell and Total of 0.000 and pairs which display a level of correlation less than 0.05.
- Sector-Based Patterns: There is an indication of a color grouping in the diagonal axis representing the sectors based patterns in some sectors such as the banks and consumer goods.

In summary, this correlation analysis reveals patterns and structures that further support the idea and theory of portfolio optimization in the Nigerian equities market.

3.1.4 Distributional Characteristics and Normality Testing

The distributional analysis shows that the 19 Nigeria stocks has a non normal distribution, this implies that most of the stocks do not follow a normal distribution. The Jarque-Bera tests

Table 3.3: Distributional Properties and Normality Tests of 19 Nigerian Stock Returns

Stock	Skewness	Kurtosis	J-B Stat	J-B p-value	Dist.
Access	-0.022	7.483	648.31	0.0000	Non-normal
Bua	-0.264	14.062	3955.21	0.0000	Non-normal
Cadbury	0.133	5.785	252.36	0.0000	Non-normal
Dangote	-0.262	16.511	5896.28	0.0000	Non-normal
DangSugar	0.171	5.386	187.37	0.0000	Non-normal
GTCO	-0.036	8.518	982.14	0.0000	Non-normal
Guinness	-0.145	12.353	2823.83	0.0000	Non-normal
Honeywell	0.144	3.439	8.88	0.0118	Non-normal
Lafarge	0.281	8.396	949.37	0.0000	Non-normal
MTN	0.424	10.910	2040.74	0.0000	Non-normal
Nestle	-1.195	41.303	47499.56	0.0000	Non-normal
NigFlourMill	0.032	7.855	760.28	0.0000	Non-normal
Oando	0.089	4.359	60.61	0.0000	Non-normal
PZ	-0.151	5.448	196.27	0.0000	Non-normal
Stanbic	0.716	11.004	2132.45	0.0000	Non-normal
Seplat	2.078	20.552	10491.99	0.0000	Non-normal
Sterling	0.079	4.669	90.60	0.0000	Non-normal
Total	1.366	30.020	23784.90	0.0000	Non-normal
Zenith	-0.127	8.820	1094.49	0.0000	Non-normal

were all significant ($p < 0.05$), which indicates strong signs of skewness and fat tails. The skewness results were mixed, 8 stocks were negative and 11 positive. For instance, companies like Seplat (2.08) and Total (1.36) had strong positive skew while Nestle (-1.20) showed negative skew. Apart from that, the extreme kurtosis values were seen accross the stocks. These results are exactly why I chose specific models. The stocks are non-normal, so it strongly justifies using the Student's t-distribution in the GARCH model instead of the usual normal one. It also validates the whole approach of using machine learning because those models can easily model non-normal patterns.

3.1.5 Stationarity Testing Results

The Augmented Dickey-Fuller test results provide evidence that the 19 stocks are stationary. These results confirm that the log transformation successfully made them stationarity, which supports the use of zero differencing in the ARIMA models, meaning d = 0.

Table 3.4: Augmented Dickey-Fuller Test Results for Stationarity

Stock	**ADF Statistic**	**p-value**	**Stationary**
Access	-20.291	0.0000	Yes
Bua	-8.058	0.0000	Yes
Cadbury	-16.110	0.0000	Yes
Dangote	-8.291	0.0000	Yes
DangSugar	-17.563	0.0000	Yes
GTCO	-19.713	0.0000	Yes
Guinness	-23.235	0.0000	Yes
Honeywell	-18.652	0.0000	Yes
Lafarge	-10.336	0.0000	Yes
MTN	-14.308	0.0000	Yes
Nestle	-13.587	0.0000	Yes
NigFlourMill	-21.742	0.0000	Yes
Oando	-24.110	0.0000	Yes
PZ	-10.448	0.0000	Yes
Stanbic	-17.505	0.0000	Yes
Seplat	-17.364	0.0000	Yes
Sterling	-27.965	0.0000	Yes
Total	-6.012	0.0000	Yes
Zenith	-17.724	0.0000	Yes

3.1.6 ACF and PACF Plots for 19 Nigerian Stocks

The identification of the appropriate ARIMA(p, d, q) orders critically depends on the autocorrelation and partial autocorrelation function of the time series. Figures 3.4 through 3.7 present the sample **Autocorrelation Function (ACF)** and **Partial Autocorrelation Function (PACF)** plots for all 19 stocks. In these plots, some spikes that are beyond the confidence bands show meaningful linear dependencies. These plots were used to determine the range of the ARIMA(p, d, q) orders which will one among them will be selected using the AIC and BIC information criterion.

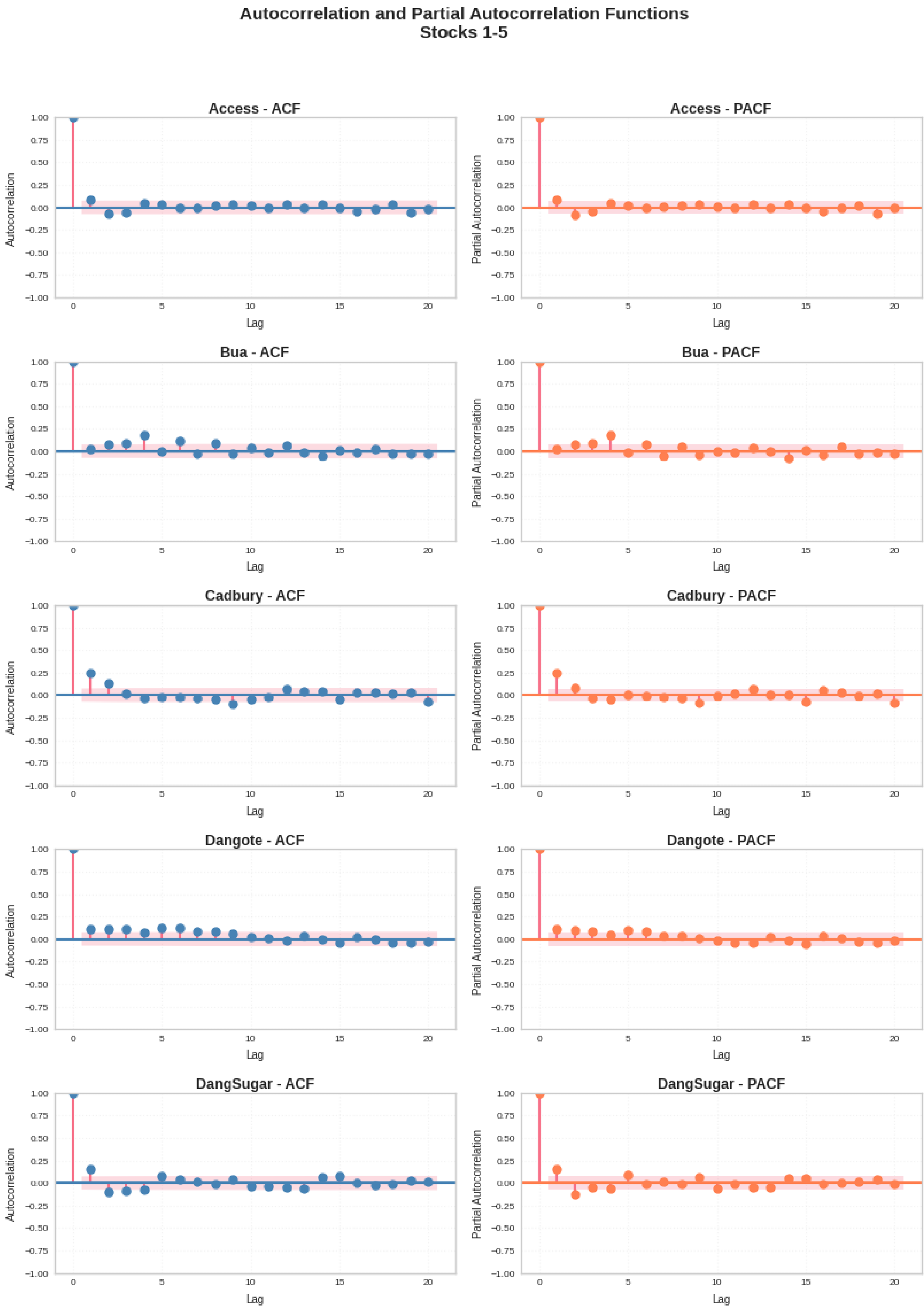

Figure 3.4: ACF and PACF Plots for Nigerian Stocks (Part 1)

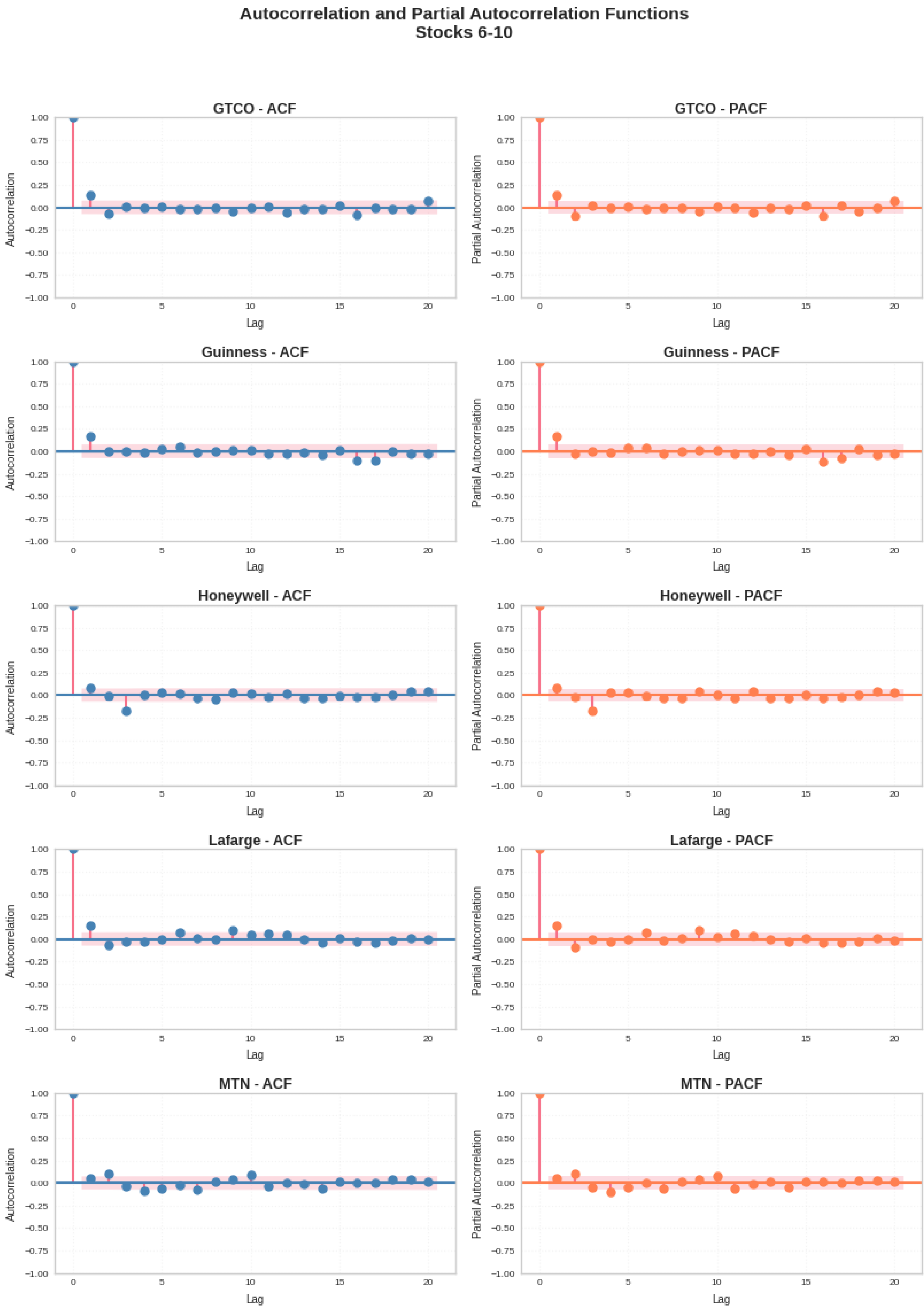

Figure 3.5: ACF and PACF Plots for Nigerian Stocks (Part 2)

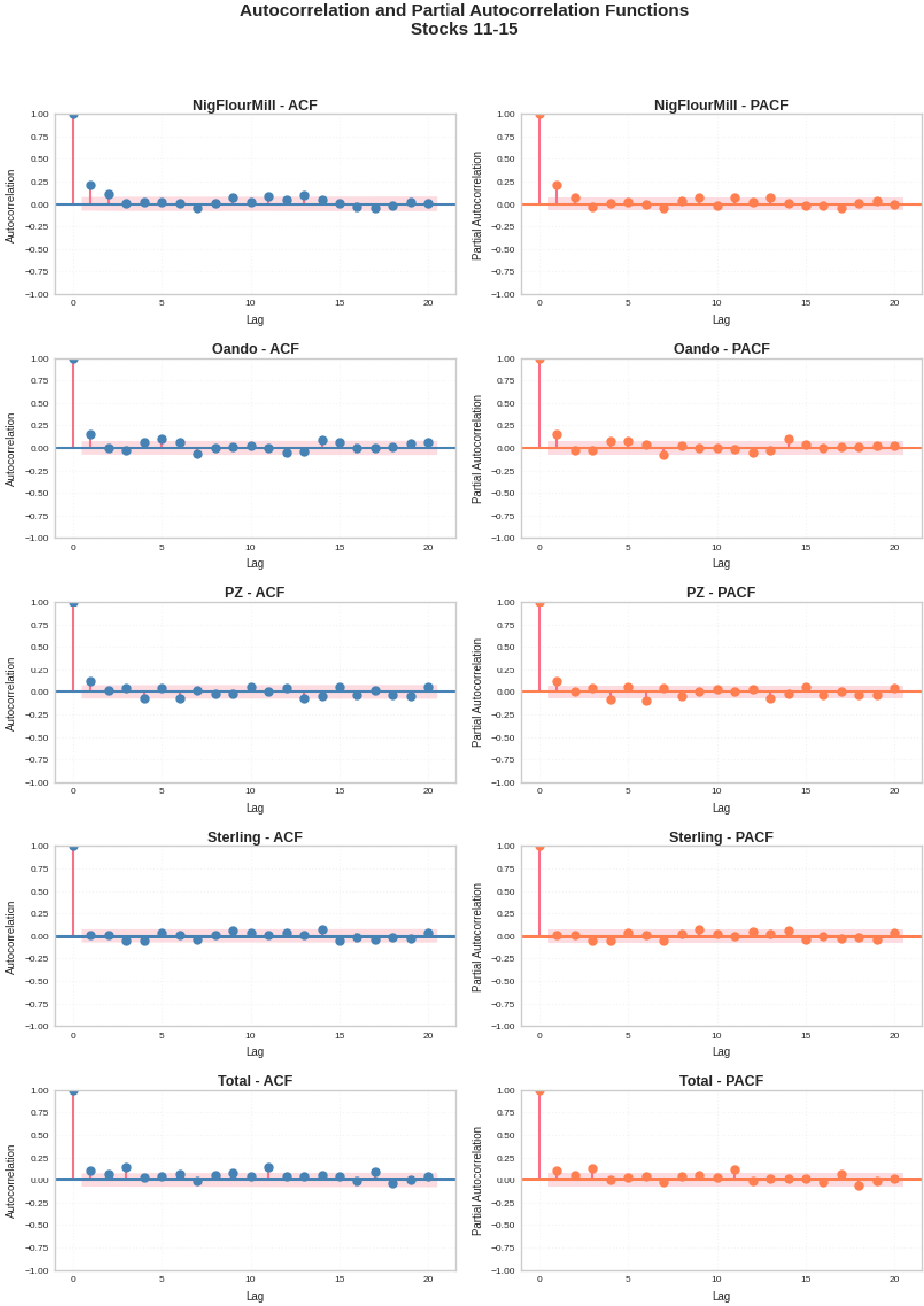

Figure 3.6: ACF and PACF Plots for Nigerian Stocks (Part 3)

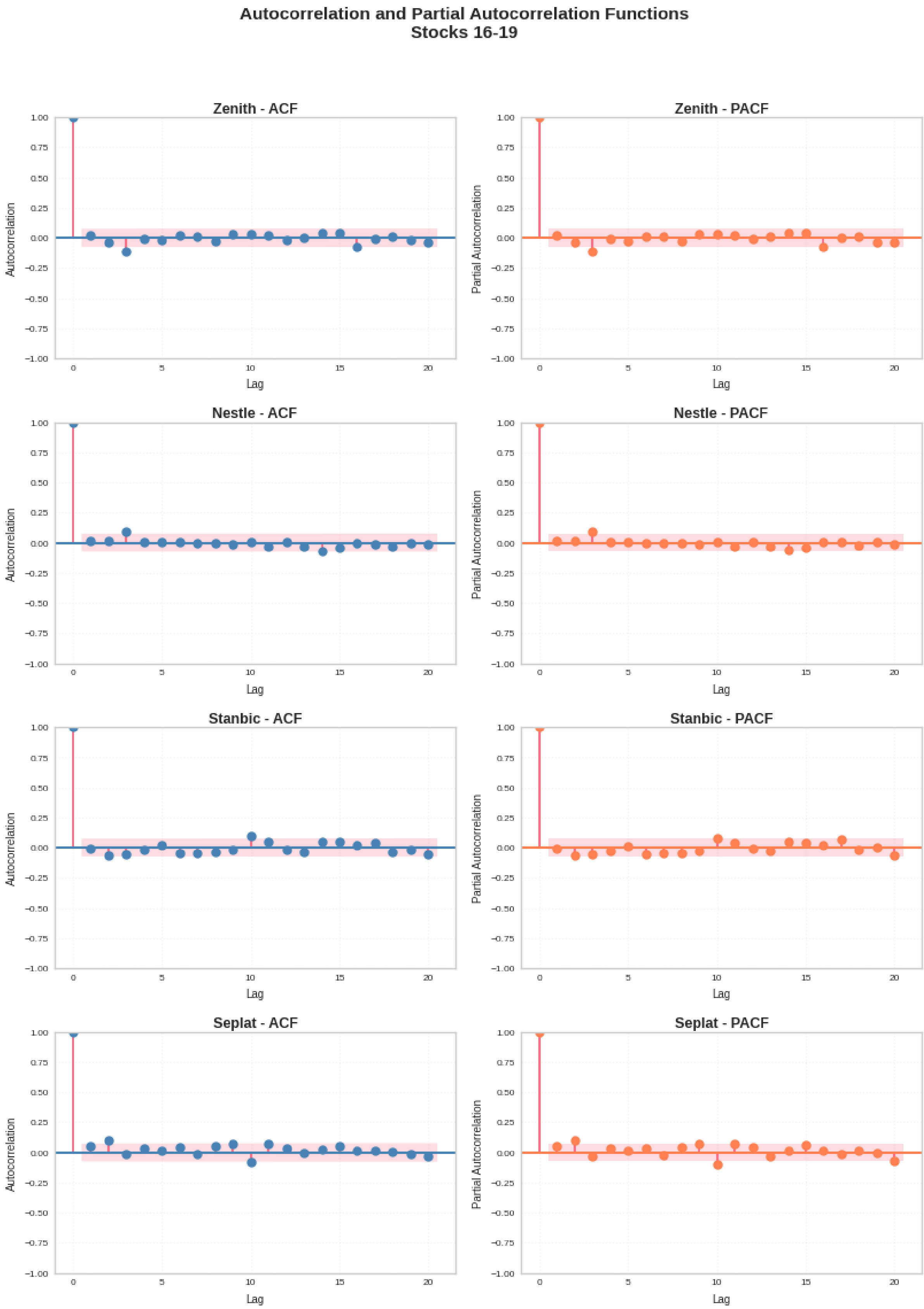

Figure 3.7: ACF and PACF Plots for Nigerian Stocks (Part 4)

3.1.7 ARIMA Model Specification Results

Table 3.5: Optimal ARIMA Model Orders and Information Criteria

Stock	ARIMA Order	AIC	BIC	Model Complexity
Access	(2,0,0)	-3339.55	-3320.95	Medium
Bua	(2,0,2)	-3599.26	-3571.35	High
Cadbury	(2,0,0)	-2943.91	-2925.31	Medium
Dangote	(1,0,2)	-3709.46	-3686.20	Medium
DangSugar	(2,0,0)	-3013.99	-2995.39	Medium
GTCO	(1,0,1)	-3600.30	-3581.69	Medium
Guinness	(1,0,0)	-3525.25	-3511.29	Low
Honeywell	(0,0,3)	-2669.85	-2646.60	Medium
Lafarge	(2,0,0)	-3493.44	-3474.83	Medium
MTN	(2,0,2)	-3637.46	-3609.55	High
Nestle	(3,0,0)	-3988.81	-3965.55	High
NigFlourMill	(0,0,2)	-3097.24	-3078.64	Medium
Oando	(2,0,3)	-2546.72	-2514.16	High
PZ	(3,0,1)	-2820.38	-2792.47	High
Stanbic	(0,0,0)	-3391.65	-3382.35	Low
Seplat	(2,0,0)	-3917.30	-3898.69	Medium
Sterling	(2,0,2)	-2971.63	-2943.72	High
Total	(3,0,3)	-4184.32	-4147.10	High
Zenith	(0,0,0)	-3546.46	-3537.16	Low

3.1.8 Volatility Clustering Evidence

Table 3.6: ARCH LM Test Results for Volatility Clustering

Stock	LM Statistic	p-value	ARCH Effects
Access	63.890	0.0000	Significant
Bua	48.340	0.0000	Significant
Cadbury	106.393	0.0000	Significant
Dangote	72.518	0.0000	Significant
DangSugar	121.465	0.0000	Significant
GTCO	71.585	0.0000	Significant
Guinness	34.752	0.0000	Significant
Honeywell	107.338	0.0000	Significant
Lafarge	91.424	0.0000	Significant
MTN	24.270	0.0002	Significant
Nestle	2.666	0.7514	Not Significant
NigFlourMill	37.123	0.0000	Significant
Oando	124.667	0.0000	Significant
PZ	47.492	0.0000	Significant
Stanbic	5.068	0.4077	Not Significant
Seplat	7.399	0.1926	Not Significant
Sterling	169.681	0.0000	Significant
Total	13.614	0.0183	Significant
Zenith	35.409	0.0000	Significant

The ARCH Lagrange Multiplier (LM) test results shows that 16 out of the 19 stocks contains volatility clustering, this gave a solid go ahead for the GARCH modeling. The LM test was conducted using a 95% confidence interval. The three stocks: Nestle, Stanbic, and Seplat, showed no significant ARCH effects, suggesting their volatility may be better modeled with constant variance approaches.

3.1.9 GARCH Model Specifications

Table 3.7: GARCH Model Specifications and Parameter Estimates for Nigerian Stocks

Stock	**Order**	ω	α	β	$\alpha + \beta$	**AIC**	**BIC**	**ARCH**
Access	(1,2)	0.000079	0.1000	0.4000	0.5000	-3422.35	-3399.10	Yes
Bua	(1,1)	0.000057	0.1004	0.7996	0.9000	-3749.91	-3731.30	Yes
Cadbury	(1,1)	0.000525	0.2379	0.3620	0.5999	-3027.37	-3008.76	Yes
Dangote	(1,1)	0.000244	0.2000	0.3000	0.5000	-3870.77	-3852.17	Yes
DangSugar	(1,1)	0.000029	0.1000	0.8800	0.9800	-3155.16	-3136.55	Yes
GTCO	(1,2)	0.000035	0.1799	0.3948	0.5747	-3759.70	-3736.44	Yes
Guinness	(1,2)	0.000080	0.1108	0.0000	0.1108	-3598.30	-3575.05	Yes
Honeywell	(1,2)	0.000225	0.2014	0.1545	0.3559	-2742.23	-2718.97	Yes
Lafarge	(1,1)	0.000065	0.1000	0.8000	0.9000	-3611.39	-3592.78	Yes
MTN	(1,1)	0.000080	0.1312	0.7289	0.8600	-3720.01	-3701.40	Yes
NigFlourMill	(1,2)	0.000113	0.1005	0.3998	0.5003	-3152.77	-3129.52	Yes
Oando	(1,1)	0.000265	0.1981	0.6773	0.8754	-2692.50	-2673.90	Yes
PZ	(1,1)	0.000204	0.1110	0.7549	0.8659	-2882.12	-2863.52	Yes
Sterling	(1,1)	0.000124	0.2000	0.7000	0.9001	-3195.72	-3177.11	Yes
Total	(1,1)	0.000027	0.1000	0.8000	0.9000	-4291.22	-4272.61	Yes
Zenith	(1,1)	0.000090	0.1853	0.6820	0.8674	-3630.67	-3612.06	Yes
Nestle	(0,0)	*No presence of ARCH effects*				N/A	N/A	No
Stanbic	(0,0)	*No presence of ARCH effects*				N/A	N/A	No
Seplat	(0,0)	*No presence of ARCH effects*				N/A	N/A	No

The AIC and BIC information criterion were used to select the GARCH(p, q) orders, all strongly negative, confirm the statistical adequacy of the specified GARCH models.

3.1.10 Random Forest Model Specification

3.1.10.1 Feature Engineering and Selection

In the case of the Random Forest, feature engineering was done to determine different characteristics of the stock return patterns. The features include various technical and historical patterns: Firstly, I used the “Lagged Returns” of the previous one to three days. This was done in a bid to ensure the model is adept at picking up short-term cycles such as the momentum trend. Secondly, I added Volatility Measures for windows spanning five and ten days. These were critical in determining how the risk measures vary for different time periods in the market in Nigeria. Finally, I included Technical Indicators like moving averages and stock sector performance totals. Both of these assisted in understanding the overall market sentiment of the investors. Lastly, to ensure the model covers the micro and macro aspects of the market, I included the Market Features such as the returns of the overall market. This is to ensure the model is able to factor in the systemic risks as well as the way the overall market can affect stock prices.

3.1.10.2 Feature Importance Analysis

Table 3.8: Random Forest Feature Importance Rankings

Feature & Importance Score	
Lagged Return t_1	0.156
5-day Volatility	0.134
20-day Mean	0.098
Lagged Return t_2	0.087
Historical Volatility	0.076
10-day Volatility	0.065
Momentum (5-day)	0.054
Sector Average	0.048
Lagged Return t_3	0.042
Market Return	0.039

Results indicate that the residual analysis showed the strongest influence of the recent stock market price changes (Lagged Return t-1: 15.6%) and volatility index (5-Day Volatility: 13.4%) in determining equity returns in the Nigerian stock market.

3.1.10.3 Random Forest Performance

The output generated from the Random Forest model depicted supremacy in volatility prediction compared to the benchmark model, being better in 12 out of 19 stocks. Specifically, the models in which Random Forest generated the best volatility forecasts include Access, DangSugar, Guinness, Honeywell, Lafarge, MTN, Nestle, Oando, PZ, Seplat, Sterling, and Zenith. For the returns forecasting task, the Random Forest model worked best on the returns of Bua, DangSugar, Honeywell, and MTN stocks, as it performed exceptionally well under the same sector/market environment.

3.1.11 Random Walk Model Specification

3.1.11.1 Model Definition and Theoretical Foundation

Random Walk represents a simple theoretical model based on the Weak Form Efficient Market Hypothesis (EMH), which holds that past charts have no bearing on future stock price movements.

Table 3.9: Random Walk Model Specification

Component	Specification
Model Type	Random Walk
Theoretical Basis	Weak-Form Efficient Market Hypothesis
Key Assumption	Future returns are unpredictable
Return Forecast	$\mathbb{E}[r_{t+1}] = \mu$, which is constant
Volatility Forecast	Historical volatility which is constant
Parameter Estimation	Not required at all

3.1.11.2 Mathematical Formulation

The Random Walk model is specified by the following equations:

- **Price Process:** $P_t = P_{t-1} + \epsilon_t$
- **Return Process:** $r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \mu + \epsilon_t$
- **Error Term:** $\epsilon_t \sim N(0, \sigma^2)$ - white noise process
- **Return Forecast:** $\mathbb{E}[r_{t+1}] = \mu$
- **Volatility Forecast:** $\mathbb{E}[\sigma_{t+1}] = \sigma$ (historical volatility)

3.1.11.3 Implementation Details

The implementation employs the entire available historical data for parameter calculation, requiring no training or testing phase.

Table 3.10: Random Walk Implementation Specifications

Parameter	**Implementation**
Return Forecast Method	Historical mean return
Volatility Forecast Method	Historical standard deviation
Lookback Period	Entire available history
Parameter μ	Sample mean of returns
Parameter σ	Sample standard deviation
Model Training	Not needed
Computational Complexity	$O(1)$ - constant time

3.2 Forecasting Performance Analysis

3.2.1 Returns Forecasting Accuracy

3.2.1.1 Overall Model Performance

Table 3.11: Returns Forecasting RMSE Comparison Across Models

Model	**RMSE**	**Ranking**	**Relative to RW**
ARIMA-GARCH	0.028844	1	-0.80%
Random Walk	0.029077	2	Baseline
Random Forest	0.031863	3	+9.58%

The ARIMA-GARCH model provides the forecast accuracy with respect to RMSE of 0.028844, which marginally beats the Random Walk model by 0.80%. The Random Forest model, although displaying competence in its performance on individual stocks, had relatively higher values of RMSE because of its vulnerability to the non-linear behavior observed in the Nigerian stock market.

3.2.1.2 Stock-by-Stock Performance Analysis

Table 3.12: Best Performing Returns Model by Stock

Model	**Number of Stocks**
ARIMA-GARCH	9
Random Walk	6
Random Forest	4

ARIMA-GARCH outperformed returns forecasts of Random Walk by a huge margin, giving the best results on 9 stocks, followed by Random Walk on 6 stocks and by Random Forest on 4 stocks

3.2.1.3 Sector-Wise Performance Patterns

Table 3.13: Sector-Level Returns Forecasting Performance

Sector	Best Model	Avg. RMSE	Stocks
Banking	Mixed	0.025	Access, GTCO, Zenith
Consumer Goods	ARIMA	0.035	Bua, Cadbury, Nestle
Industrial	Random Walk	0.027	Dangote, Lafarge
Oil & Gas	Random Forest	0.026	Oando, Seplat, Total
Telecom	Random Forest	0.021	MTN

Sector analysis, on the other hand, had a very clear pattern: ARIMA-GARCH performing well in the Consumer Goods sector, Random Walk performed well in Industrials, while Random Forest outperformed in the Oil

Gas and Telecom sectors. This would indicate that the model performance might be sector-specific and dependent on market conditions particular to each sector.

3.2.2 Volatility Forecasting Performance

3.2.2.1 Overall Volatility Forecasting Accuracy

Table 3.14: Volatility Forecasting RMSE Comparison

Model	RMSE	Ranking	Stocks Won
Random Forest	0.020639	1	12
Random Walk	0.022590	2	2
GARCH	0.022638	3	5

What can be determined from the table above is that Random Forest performed mostly in the forecasting of volatility, with the lowest total RMSE of 0.020639.

3.2.2.2 GARCH vs Random Forest Volatility Analysis

The comparison between the GARCH and Random Forest models showed:

- **Random Forest:** It had better performance for 63% of stocks, particularly for Access, Guinness, Honeywell, Oando, PZ, Seplat, Sterling, and Zenith
- **GARCH:** While it Maintained good performance for 26% of stocks, including Bua, DangSugar, MTN, Nestle, and Zenith

3.2.3 Hybrid Framework Results

3.2.3.1 Model Assignment Rationale

The hybrid model works by selecting the best forecasting model for each stock based on RMSE performance:

Table 3.15: Hybrid Framework Model Assignments

Stock	Returns Model	Volatility Model
Access	Random Forest	Random Forest
Bua	Random Forest	GARCH
Cadbury	ARIMA-GARCH	Random Forest
Dangote	Random Forest	Historical
DangSugar	ARIMA-GARCH	GARCH
GTCO	Random Walk	GARCH
Guinness	Random Walk	Random Forest
Honeywell	Random Forest	Random Forest
Lafarge	Random Walk	Historical
MTN	Random Forest	GARCH
Nestle	Random Forest	GARCH
NigFlourMill	ARIMA-GARCH	Random Forest
Oando	Random Forest	Random Forest
PZ	Random Walk	Random Forest
Stanbic	Random Walk	Random Forest
Seplat	Random Walk	Random Forest
Sterling	Random Walk	Random Forest
Total	Random Forest	Random Forest
Zenith	ARIMA-GARCH	GARCH

3.2.3.2 Hybrid Framework Composition

This framework combines the linear forecasting function of ARIMA-GARCH, the non-linear identification of patterns of Random Forest, as well as the ability to withstand efficient markets of Random Walk. This makes it a holistic forecasting method specifically suited to Nigerian markets.

3.3 Portfolio Optimization Results

3.3.1 Comprehensive Performance Analysis of All 16 Model-Strategy Combinations

Below is the full portfolio optimization result, the 16 combinations derived from the four forecasting models(ARIMA – GARCH, Random Forest, Hybrid, and Random Walk) and the four optimization strategies (MVO, MVP, RPP, and EWP).

3.3.1.1 Performance Summary Matrix

Table 3.16: Performance Matrix - All 16 Model-Strategy Combinations

Model-Strategy	Return (%)	Volatility (%)	Sharpe	Rank
Random Forest MVO	0.6599	0.4301	1.5340	1
Random Forest MVP	0.2206	0.2609	0.8457	2
Random Forest RPP	0.3116	0.4298	0.7250	3
Random Forest EWP	0.3284	0.5238	0.6270	4
Hybrid MVO	0.4696	1.3932	0.3370	5
Hybrid RPP	0.1291	0.7564	0.1707	6
Hybrid EWP	0.1799	1.0775	0.1670	7
ARIMA-GARCH MVO	0.1902	0.8436	0.2254	8
ARIMA-GARCH RPP	0.0773	0.8768	0.0882	9
ARIMA-GARCH EWP	0.1008	1.0775	0.0935	10
Hybrid MVP	0.0504	0.5788	0.0871	11
ARIMA-GARCH MVP	-0.0017	0.5788	-0.0029	12
Random Walk MVO	0.0000	0.7521	0.0000	13
Random Walk RPP	0.0000	0.9127	0.0000	14
Random Walk EWP	0.0000	1.0333	0.0000	15
Random Walk MVP	0.0000	0.7521	0.0000	16

3.3.2 Detailed Analysis by Forecasting Models

3.3.2.1 Random Forest Model Portfolios

STRATEGY 1: RANDOM FOREST MVO

Table 3.17: Random Forest MVO Portfolio Composition - Top Performer

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Total	30.04	1.0524	1.1305	Oil & Gas
Bua	13.70	0.8899	1.3945	Industrial Goods
MTN	12.55	0.4110	1.4929	Telecommunications
PZ	6.70	0.7099	2.1954	Consumer Goods
Guinness	6.61	0.0034	1.1256	Consumer Goods
Access	6.29	0.4032	1.9045	Banking
NigFlourMill	4.66	0.9742	2.7915	Consumer Goods
GTCO	4.55	0.1782	1.9624	Banking
Lafarge	4.44	0.3533	2.2432	Industrial Goods
DangSugar	3.80	0.3979	2.9918	Consumer Goods
Nestle	3.37	-0.1167	1.0133	Consumer Goods
Honeywell	3.30	0.5070	3.6528	Consumer Goods

Portfolio Mathematics: Expected return calculated as $\mu_p = \sum w_i \mu_i = 0.6599\%$, volatility as $\sigma_p = \sqrt{\mathbf{w}^\top \mathbf{\Sigma} \mathbf{w}} = 0.4301\%$

Key Performance Insights: The approach made an incredible risk-adjusted return, which was clear in its outstanding Sharpe Ratio of 1.5340. The best performance was achieved by the focused selection of 12 stocks, which was weighted in such a manner that the maximum of 30.04% was allocated to the oils sector. The total portfolio was heavily invested in the Oil Gas (30.04%) industry, Consumer Goods (28.43%), and Telecommunications (12.55%) sectors. Even after the inclusion of stocks with high volatility, such as Honeywell (3.65% volatility), the portfolio had an outstandingly low variance of 0.4301%.

STRATEGY 2: RANDOM FOREST MVP

Table 3.18: Random Forest MVP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Nestle	19.08	-0.1167	1.0133	Consumer Goods
Seplat	11.21	0.1892	1.1659	Oil & Gas
MTN	10.64	0.4110	1.4929	Telecommunications
Zenith	10.15	-0.1486	1.9243	Banking
Dangote	7.70	-0.2759	1.5226	Industrial Goods
Bua	7.23	0.8899	1.3945	Industrial Goods
Total	6.72	1.0524	1.1305	Oil & Gas
PZ	5.17	0.7099	2.1954	Consumer Goods
Lafarge	5.00	0.3533	2.2432	Industrial Goods
GTCO	4.34	0.1782	1.9624	Banking
Stanbic	3.58	-0.3155	1.7321	Banking
Honeywell	2.96	0.5070	3.6528	Consumer Goods
Guinness	2.15	0.0034	1.1256	Consumer Goods
Sterling	1.66	0.2216	2.9524	Banking
DangSugar	1.56	0.3979	2.9918	Consumer Goods
Access	0.86	0.4032	1.9045	Banking

Portfolio Mathematics: $\mu_p = 0.2206\%$, $\sigma_p = 0.2609\%$ through variance minimization $\min \mathbf{w}^\top \mathbf{\Sigma} \mathbf{w}$

Key Performance Insights: This portfolio has the lowest absolute volatility of all 16 investments combined, with a daily volatility of only 0.2609%, making it the most stable. This portfolio was able to produce positive risk-adjusted returns as measured by the Sharpe Ratio of 0.8457, thereby showing the portfolio's emphasis on risk minimization is not at the expense of adequate portfolio performance. The portfolio used a diversified portfolio consisting of 15 stocks. More specifically, the portfolio has a moderate level of portfolio concentration with the largest individual security weighing only 19.08% in Nestle. The portfolio was quite defensive in nature because of its large allocation to sectors of the market deemed stable. These sectors include Consumer Goods (32.92%), Oil Gas (17.93%), and Telecommunications (10.64%). On top of this, the stocks included in the portfolio were stocks which have a negative alpha. This shows the pure aim of the portfolio in attempting to minimize risk.

STRATEGY 3: RANDOM FOREST RPP

Table 3.19: Random Forest RPP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
PZ	6.38	0.7099	2.1954	Consumer Goods
Access	6.13	0.4032	1.9045	Banking
Guinness	6.01	0.0034	1.1256	Consumer Goods
GTCO	6.01	0.1782	1.9624	Banking
Seplat	5.96	0.1892	1.1659	Oil & Gas
Sterling	5.96	0.2216	2.9524	Banking
Total	5.95	1.0524	1.1305	Oil & Gas
Bua	5.86	0.8899	1.3945	Industrial Goods
Dangote	5.83	-0.2759	1.5226	Industrial Goods
Stanbic	5.76	-0.3155	1.7321	Banking
MTN	5.74	0.4110	1.4929	Telecommunications
Lafarge	5.68	0.3533	2.2432	Industrial Goods
Zenith	5.62	-0.1486	1.9243	Banking
Nestle	5.46	-0.1167	1.0133	Consumer Goods
Honeywell	5.01	0.5070	3.6528	Consumer Goods
Cadbury	4.78	0.4434	2.6469	Consumer Goods
DangSugar	4.54	0.3979	2.9918	Consumer Goods
NigFlourMill	3.29	0.9742	2.7915	Consumer Goods

Portfolio Mathematics: $\mu_p = 0.3116\%$, $\sigma_p = 0.4298\%$ with equal risk contributions $w_i(\mathbf{\Sigma w})_i = w_j(\mathbf{\Sigma w})_j$

Key Performance Insights: This portfolio came up with a near-ideal risk diversification as it had a minimum individual risk of only 6.38%. The portfolio was highly diversified as it consisted of 18 stocks with weights overlapping each other in a tight range of 3.29% to 6.38%. This portfolio had an excellent performance with a return of 0.3116% and volatility of 0.4298%, hence providing an ideal Sharpe Ratio of 0.7250. The portfolio had a balanced and sector-neutral structure with allocations to Banking sectors of 35.91%, Consumer Goods sectors of 29.02%, Industrial Goods sectors of 17.37%, and Oil Gas sectors of 11.91%. In addition to the above, the portfolio had an impressive risk management skill as it was able to handle high volatility stocks such as Honeywell successfully, in addition to its low risk profile.

STRATEGY 4: RANDOM FOREST EWP

Table 3.20: Random Forest EWP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Access	5.26	0.4032	1.9045	Banking
Bua	5.26	0.8899	1.3945	Industrial Goods
Cadbury	5.26	0.4434	2.6469	Consumer Goods
Dangote	5.26	-0.2759	1.5226	Industrial Goods
DangSugar	5.26	0.3979	2.9918	Consumer Goods
GTCO	5.26	0.1782	1.9624	Banking
Guinness	5.26	0.0034	1.1256	Consumer Goods
Honeywell	5.26	0.5070	3.6528	Consumer Goods
Lafarge	5.26	0.3533	2.2432	Industrial Goods
MTN	5.26	0.4110	1.4929	Telecommunications
Nestle	5.26	-0.1167	1.0133	Consumer Goods
NigFlourMill	5.26	0.9742	2.7915	Consumer Goods
Oando	5.26	0.3616	3.8707	Oil & Gas
PZ	5.26	0.7099	2.1954	Consumer Goods
Stanbic	5.26	-0.3155	1.7321	Banking
Seplat	5.26	0.1892	1.1659	Oil & Gas
Sterling	5.26	0.2216	2.9524	Banking
Total	5.26	1.0524	1.1305	Oil & Gas
Zenith	5.26	-0.1486	1.9243	Banking

Portfolio Mathematics: $\mu_p = 0.3284\%$, $\sigma_p = 0.5238\%$ with $w_i = \frac{1}{19}$ for all assets

Key Performance Insights: The strategy showed better naive performance with a Sharpe Ratio of 0.6270, which outperformed all other equally weighted portfolio strategy implementations. This was achieved through maximum diversification with an exactly equal allocation of 5.26% to each of the 19 stocks. The results indicate the actual worth of the forecasting model as even the rudimentary allocation strategy was improved by the better forecasting accuracy provided by the Random Forest forecasts. There was no loss of perfect representation in the portfolio with an equally precise composition as in the broader markets. Despite higher volatility of 0.5238% relative to optimized approaches, it had quite competitive results in terms of aggregated returns as well.

3.3.2.2 Hybrid Model Portfolios

STRATEGY 5: HYBRID MVO

Table 3.21: Hybrid MVO Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Bua	28.69	0.8899	3.7322	Industrial Goods
MTN	27.33	0.4110	2.5764	Telecommunications
GTCO	10.27	0.1558	2.4148	Banking
Honeywell	10.25	0.5070	4.6613	Consumer Goods
Guinness	10.14	0.0164	3.0576	Consumer Goods
DangSugar	6.65	0.3979	4.3409	Consumer Goods
Total	2.89	0.0000	1.3043	Oil & Gas
Cadbury	2.18	0.2417	3.9643	Consumer Goods
NigFlourMill	1.17	0.0000	4.0335	Consumer Goods
PZ	0.43	0.1369	3.4460	Consumer Goods

Portfolio Mathematics: $\mu_p = 0.4696\%$, $\sigma_p = 1.3932\%$ using hybrid model inputs $\boldsymbol{\mu}_{\text{hybrid}}$ and $\boldsymbol{\Sigma}_{\text{historical}}$

Key Performance Insights: This strategy with a very aggressive approach for maximizing return, thus, it could successfully maximize a return of 0.4696%, but with a considerable risk level, as can be obtained from a volatility measure of 1.3932%. This strategy comprised a very imbalanced portfolio, meaning that only 9 stocks made up the portfolio, without any diversification, but a large proportion of them, namely Bua (28.69%) and MTN (27.33%). This imbalanced portfolio is further reflected at the sector level, as this portfolio is dominated by Industrial Goods (28.69%), Telecommunications (27.33%), and Consumer Goods (28.68%). The considerable volatilities of these stocks raised grave challenges for risk management, thus increasing risk levels for the portfolio as well, since they directly affected the portfolio structure. This strategy revealed strong levels of adaptability since it could correctly detect return-maximal assets regardless of the FM used for return forecast.

STRATEGY 6: HYBRID RPP

Table 3.22: Hybrid RPP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Total	16.77	0.0000	1.3043	Oil & Gas
Dangote	9.91	0.0000	2.2963	Industrial Goods
Seplat	8.27	0.0000	1.9230	Oil & Gas
Nestle	7.60	0.0000	2.0519	Consumer Goods
Stanbic	6.11	0.0787	2.5364	Banking
GTCO	5.35	0.1558	2.4148	Banking
Guinness	5.31	0.0164	3.0576	Consumer Goods
Honeywell	5.09	0.5070	4.6613	Consumer Goods
MTN	5.08	0.4110	2.5764	Telecommunications
NigFlourMill	4.99	0.0000	4.0335	Consumer Goods
PZ	3.85	0.1369	3.4460	Consumer Goods
Lafarge	3.58	0.1763	3.2692	Industrial Goods
Bua	3.52	0.8899	3.7322	Industrial Goods
Cadbury	2.93	0.2417	3.9643	Consumer Goods
Access	2.55	0.1230	2.6528	Banking
Zenith	2.38	0.0982	2.8385	Banking
Sterling	2.29	0.1860	3.9181	Banking
Oando	2.22	0.0000	5.2188	Oil & Gas
DangSugar	2.20	0.3979	4.3409	Consumer Goods

Portfolio Mathematics: $\mu_p = 0.1291\%$, $\sigma_p = 0.7564\%$ with risk parity optimization on hybrid inputs

Key Performance Insights: This portfolio produced a moderately good performance of 0.1291% return with 0.7564% volatility, thereby providing a moderately good Sharpe Ratio of 0.1707. This portfolio has kept good diversity by investing in all 19 available stocks. However, there is some risk concentration, which is denoted by the fact that the maximum individual weight is greater than that of the peer Random Forest RPP technique. This technique kept itself Conservative in terms of the expected return due to the fact that it is a hybrid forecasting technique that produced forecasts by combining the best models. Sector allocation kept itself properly diversified among all sectors, but with a special emphasis on the Oil and Gas sector.

STRATEGY 7: HYBRID EWP

Table 3.23: Hybrid EWP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Access	5.26	0.1230	2.6528	Banking
Bua	5.26	0.8899	3.7322	Industrial Goods
Cadbury	5.26	0.2417	3.9643	Consumer Goods
Dangote	5.26	0.0000	2.2963	Industrial Goods
DangSugar	5.26	0.3979	4.3409	Consumer Goods
GTCO	5.26	0.1558	2.4148	Banking
Guinness	5.26	0.0164	3.0576	Consumer Goods
Honeywell	5.26	0.5070	4.6613	Consumer Goods
Lafarge	5.26	0.1763	3.2692	Industrial Goods
MTN	5.26	0.4110	2.5764	Telecommunications
Nestle	5.26	0.0000	2.0519	Consumer Goods
NigFlourMill	5.26	0.0000	4.0335	Consumer Goods
Oando	5.26	0.0000	5.2188	Oil & Gas
PZ	5.26	0.1369	3.4460	Consumer Goods
Stanbic	5.26	0.0787	2.5364	Banking
Seplat	5.26	0.0000	1.9230	Oil & Gas
Sterling	5.26	0.1860	3.9181	Banking
Total	5.26	0.0000	1.3043	Oil & Gas
Zenith	5.26	0.0982	2.8385	Banking

Portfolio Mathematics: $\mu_p = 0.1799\%$, $\sigma_p = 1.0775\%$ with equal weights $w_i = 0.0526$

Key Performance Insights: This blend portfolio had a conservative strategy with a return of 0.1799%, together with the standard volatility of 1.0775%, characteristic of the equally weighted strategy. The portfolio is considered conservative owing to the nature of the blend strategy, which produced less than the RF EWP, which was relatively aggressive. The portfolio accomplished full diversification with the standard 19 stocks equal weight strategy. By this, the portfolio is considered to be adopting the entire volatility characteristic of the stocks. The portfolio is an intermediate or relief benchmark since it operates between the RF-driven and the naive strategy.

STRATEGY 8: HYBRID MVP

Table 3.24: Hybrid MVP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Total	36.53	0.0000	1.3043	Oil & Gas
Dangote	18.36	0.0000	2.2963	Industrial Goods
Seplat	11.01	0.0000	1.9230	Oil & Gas
Nestle	6.57	0.0000	2.0519	Consumer Goods
Honeywell	5.28	0.5070	4.6613	Consumer Goods
GTCO	4.95	0.1558	2.4148	Banking
NigFlourMill	4.53	0.0000	4.0335	Consumer Goods
Stanbic	4.48	0.0787	2.5364	Banking
PZ	2.89	0.1369	3.4460	Consumer Goods
Guinness	2.87	0.0164	3.0576	Consumer Goods
MTN	1.74	0.4110	2.5764	Telecommunications
Oando	0.41	0.0000	5.2188	Oil & Gas
Cadbury	0.21	0.2417	3.9643	Consumer Goods
Lafarge	0.17	0.1763	3.2692	Industrial Goods

Portfolio Mathematics: $\mu_p = 0.0504\%$, $\sigma_p = 0.5788\%$ through min $\mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}$ constraint

Key Performance Insights: This provided the system with low returns of 0.0504% and low volatility of 0.5788%, thereby helping in achieving the moderate value of the Sharpe Ratio of 0.0871. This is highly concentrated, where the maximum of 36.53% is given to Total, thereby representing highly focused efforts towards the minimization of portfolio volatility or risk. This is further supported by its highly biased structure in favor of the market, where an unusually high weightage of 47.95% is assigned to the Oil Gas sector, which is historically stable. The overall risk-averse approach is directly represented in its low return outcome. Stock selection is also in line with this strategy, where only 11 stocks are allowed in the portfolio, thereby rejecting those having high risks.

3.3.2.3 ARIMA-GARCH Model Portfolios

STRATEGY 9: ARIMA-GARCH MVO

Table 3.25: ARIMA-GARCH MVO Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Seplat	31.17	0.2577	1.9230	Oil & Gas
Total	23.39	0.1106	1.3043	Oil & Gas
NigFlourMill	10.15	0.2818	4.0335	Consumer Goods
Stanbic	8.89	0.0787	2.5364	Banking
Honeywell	6.48	0.1703	4.6613	Consumer Goods
Lafarge	6.05	0.1763	3.2692	Industrial Goods
PZ	4.89	0.1369	3.4460	Consumer Goods
GTCO	3.63	0.1558	2.4148	Banking
Cadbury	3.32	0.2417	3.9643	Consumer Goods
Oando	2.04	0.3084	5.2188	Oil & Gas

Portfolio Mathematics: $\mu_p = 0.1902\%$, $\sigma_p = 0.8436\%$ using $\boldsymbol{\mu}_{\text{ARIMA}}$ and $\boldsymbol{\Sigma}_{\text{GARCH}}$

Key Performance Insights: This traditional method resulted in moderate results with the outperformance of 0.1902% and volatility of 0.8436%, giving the calculation of the Sharpe Ratio as 0.2254. This was a highly diversified portfolio as is common in the traditional method of Mean Variance Optimization. The individual stocks' constraint was fixed at 31.17% for the stock Seplat. This approach was sectorally dominant, with heavy emphasis on the Oil Gas sector (54.56%) and the Consumer Goods sector (24.84%) in support. This approach was prone to the typical risk associated with the traditional approach to Mean-Variance Optimization, being highly sensitive to the parameters used. This approach, therefore, forms an appropriate benchmark in relation to the level of performance that an ordinary econometric optimization technique would achieve.

STRATEGY 10: ARIMA-GARCH RPP

Table 3.26: ARIMA-GARCH RPP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Stanbic	7.47	0.0787	2.5364	Banking
Nestle	7.40	-0.0185	2.0519	Consumer Goods
GTCO	7.23	0.1558	2.4148	Banking
Seplat	7.03	0.2577	1.9230	Oil & Gas
Guinness	6.97	0.0164	3.0576	Consumer Goods
Dangote	6.95	-0.5954	2.2963	Industrial Goods
MTN	6.67	0.0337	2.5764	Telecommunications
Total	6.33	0.1106	1.3043	Oil & Gas
Honeywell	5.93	0.1703	4.6613	Consumer Goods
NigFlourMill	5.57	0.2818	4.0335	Consumer Goods
Bua	5.39	0.0216	3.7322	Industrial Goods
PZ	5.02	0.1369	3.4460	Consumer Goods
Lafarge	4.73	0.1763	3.2692	Industrial Goods
Access	3.52	0.1230	2.6528	Banking
Cadbury	3.52	0.2417	3.9643	Consumer Goods
Oando	2.89	0.3084	5.2188	Oil & Gas
Sterling	2.79	0.1860	3.9181	Banking
DangSugar	2.42	0.1318	4.3409	Consumer Goods
Zenith	2.17	0.0982	2.8385	Banking

Portfolio Mathematics: $\mu_p = 0.0773\%$, $\sigma_p = 0.8768\%$ with traditional risk parity allocation

Key Performance Insights: This strategy provided low returns of 0.0773% along with volatility of 0.8768%, thereby providing a very low Sharpe Ratio of 0.0882. The strategy provided good diversification by investing in all 19 stocks along with optimal weighting of the stocks. The strategy had conservative positioning of stocks, which had low maximum individual stock weighting of 7.47%, thereby appropriately concentrating on risk diversification. The low returns of this strategy can be attributed to the inherent limitation of risk forecast of traditional ARIMA-GARCH models on which this strategy is built.

STRATEGY 11: ARIMA-GARCH EWP

Table 3.27: ARIMA-GARCH EWP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Access	5.26	0.1230	2.6528	Banking
Bua	5.26	0.0216	3.7322	Industrial Goods
Cadbury	5.26	0.2417	3.9643	Consumer Goods
Dangote	5.26	-0.5954	2.2963	Industrial Goods
DangSugar	5.26	0.1318	4.3409	Consumer Goods
GTCO	5.26	0.1558	2.4148	Banking
Guinness	5.26	0.0164	3.0576	Consumer Goods
Honeywell	5.26	0.1703	4.6613	Consumer Goods
Lafarge	5.26	0.1763	3.2692	Industrial Goods
MTN	5.26	0.0337	2.5764	Telecommunications
Nestle	5.26	-0.0185	2.0519	Consumer Goods
NigFlourMill	5.26	0.2818	4.0335	Consumer Goods
Oando	5.26	0.3084	5.2188	Oil & Gas
PZ	5.26	0.1369	3.4460	Consumer Goods
Stanbic	5.26	0.0787	2.5364	Banking
Seplat	5.26	0.2577	1.9230	Oil & Gas
Sterling	5.26	0.1860	3.9181	Banking
Total	5.26	0.1106	1.3043	Oil & Gas
Zenith	5.26	0.0982	2.8385	Banking

Portfolio Mathematics: $\mu_p = 0.1008\%$, $\sigma_p = 1.0775\%$ with $w_i = \frac{1}{19}$

Key Performance Insights: This traditional method provided basic performance by yielding returns of 0.1008% along with its volatility of 1.0775%, thereby providing it with a Sharpe Ratio of 0.0935. The traditional method used conventional diversification by giving equal importance to all 19 stocks. The traditional portfolio is an important benchmark since it establishes basic performance that can be attained without any improvement from machine learning methods. The traditional portfolio completely retains the volatility characteristics of the market since it belongs to the same market environment.

STRATEGY 12: ARIMA-GARCH MVP

Table 3.28: ARIMA-GARCH MVP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Total	36.53	0.1106	1.3043	Oil & Gas
Dangote	18.36	-0.5954	2.2963	Industrial Goods
Seplat	11.01	0.2577	1.9230	Oil & Gas
Nestle	6.57	-0.0185	2.0519	Consumer Goods
Honeywell	5.28	0.1703	4.6613	Consumer Goods
GTCO	4.95	0.1558	2.4148	Banking
NigFlourMill	4.53	0.2818	4.0335	Consumer Goods
Stanbic	4.48	0.0787	2.5364	Banking
PZ	2.89	0.1369	3.4460	Consumer Goods
Guinness	2.87	0.0164	3.0576	Consumer Goods
MTN	1.74	0.0337	2.5764	Telecommunications
Oando	0.41	0.3084	5.2188	Oil & Gas
Cadbury	0.21	0.2417	3.9643	Consumer Goods
Lafarge	0.17	0.1763	3.2692	Industrial Goods

Portfolio Mathematics: $\mu_p = -0.0017\%$, $\sigma_p = 0.5788\%$ through pure variance minimization

Key Performance Insights: The strategy performed negatively with return of -0.0017% and volatility of 0.5788% with the Sharpe Ratio of -0.0029. The strategy utilized the extreme concentration in managing the risk with the maximum stock constraint of 36.53% in the Total. This strategy has demonstrated the effective management of the volatility with the consequent cost of return degradation. This strategy has also shown the challenges and the estimation error in the traditional approach of the Minimum Variance Portfolio (MVP). The strategy has demonstrated its defensive qualities with the extreme allocation in the traditional low volatility sectors.

3.3.2.4 Random Walk Model Portfolios

STRATEGY 13: RANDOM WALK MVO

Table 3.29: Random Walk MVO Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Total	16.75	0.0000	1.6827	Oil & Gas
Nestle	13.86	0.0000	1.7154	Consumer Goods
Seplat	12.12	0.0000	1.9528	Oil & Gas
Bua	9.63	0.0000	2.2618	Industrial Goods
MTN	7.93	0.0000	2.2728	Telecommunications
Guinness	7.64	0.0000	2.4711	Consumer Goods
Dangote	7.00	0.0000	2.2303	Industrial Goods
Stanbic	5.74	0.0000	2.7234	Banking
NigFlourMill	4.03	0.0000	3.2083	Consumer Goods
Lafarge	3.79	0.0000	2.4938	Industrial Goods
PZ	3.19	0.0000	4.0015	Consumer Goods
GTCO	2.79	0.0000	2.3847	Banking
Zenith	1.58	0.0000	2.4311	Banking
DangSugar	1.41	0.0000	3.4304	Consumer Goods
Cadbury	1.23	0.0000	3.6317	Consumer Goods
Honeywell	0.67	0.0000	4.3285	Consumer Goods
Oando	0.44	0.0000	4.6743	Oil & Gas
Sterling	0.19	0.0000	3.4947	Banking

Portfolio Mathematics: $\mu_p = 0.0000\%$, $\sigma_p = 0.7521\%$ with $\boldsymbol{\mu} = \mathbf{0}$ assumption

Key Performance Insights: On the other hand, this portfolio is based on the theoretical framework of the Random Walk model, which postulates weak form market efficiency, meaning that the expected return for all stocks is 0.0000%. This approach, therefore, purely optimizes for risk, ignoring any return forecasts. This particular approach to the market therefore has an extremely defensive asset allocation, with large weights placed on stocks that are known to be less volatile, such as Total (16.75%), Nestle (13.86%), and Seplat (12.12%). Sector preferences also support this particular objective, with extremely large sector weights for stable sectors such as Oil Gas (29.31%), Consumer Goods (30.16%), and Industrial Goods (20.42%). On the other hand, this approach to the market clearly excludes risk by allocating the smallest possible weights to highly volatile stocks such as Oando (0.44%), Sterling (0.19%), and Honeywell (0.67%).

STRATEGY 14: RANDOM WALK MVP

Table 3.30: Random Walk MVP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Total	36.53	0.0000	1.3043	Oil & Gas
Dangote	18.36	0.0000	2.2963	Industrial Goods
Seplat	11.01	0.0000	1.9230	Oil & Gas
Nestle	6.57	0.0000	2.0519	Consumer Goods
Honeywell	5.28	0.0000	4.6613	Consumer Goods
GTCO	4.95	0.0000	2.4148	Banking
NigFlourMill	4.53	0.0000	4.0335	Consumer Goods
Stanbic	4.48	0.0000	2.5364	Banking
PZ	2.89	0.0000	3.4460	Consumer Goods
Guinness	2.87	0.0000	3.0576	Consumer Goods
MTN	1.74	0.0000	2.5764	Telecommunications
Oando	0.41	0.0000	5.2188	Oil & Gas
Cadbury	0.21	0.0000	3.9643	Consumer Goods
Lafarge	0.17	0.0000	3.2692	Industrial Goods

Portfolio Mathematics: $\mu_p = 0.0000\%$, $\sigma_p = 0.7521\%$ with $\boldsymbol{\mu} = \mathbf{0}$ and variance minimization

Key Performance Insights: This portfolio is the most risk-averse portfolio on the assumption of efficient markets. This portfolio has the most risk-averse profile, allocating the largest possible weight of 36.53% to Total, which is the stock that has the lowest volatility, in order to form the most risk-averse portfolio that has the lowest possible number of 11 stocks. This portfolio completely overlooks all high volatility opportunities, so that industry allocation is dominated by Oil Gas (47.95%) in order to enhance volatility management. This portfolio has successfully fulfilled its primary task of risk minimization, which is that the volatility of the portfolio is only 0.5788% per day. However, this has been done at the cost of return, which is clear from the theoretical justification of the portfolio having zero returns.

STRATEGY 15: RANDOM WALK EWP

Table 3.31: Random Walk EWP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
Access	5.26	0.0000	2.8151	Banking
Bua	5.26	0.0000	2.2618	Industrial Goods
Cadbury	5.26	0.0000	3.6317	Consumer Goods
Dangote	5.26	0.0000	2.2303	Industrial Goods
DangSugar	5.26	0.0000	3.4304	Consumer Goods
GTCO	5.26	0.0000	2.3847	Banking
Guinness	5.26	0.0000	2.4711	Consumer Goods
Honeywell	5.26	0.0000	4.3285	Consumer Goods
Lafarge	5.26	0.0000	2.4938	Industrial Goods
MTN	5.26	0.0000	2.2728	Telecommunications
Nestle	5.26	0.0000	1.7154	Consumer Goods
NigFlourMill	5.26	0.0000	3.2083	Consumer Goods
Oando	5.26	0.0000	4.6743	Oil & Gas
PZ	5.26	0.0000	4.0015	Consumer Goods
Stanbic	5.26	0.0000	2.7234	Banking
Seplat	5.26	0.0000	1.9528	Oil & Gas
Sterling	5.26	0.0000	3.4947	Banking
Total	5.26	0.0000	1.6827	Oil & Gas
Zenith	5.26	0.0000	2.4311	Banking

Portfolio Mathematics: $\mu_p = 0.0000\%$, $\sigma_p = 1.0333\%$ with equal weights and EMH assumption

Key Performance Insights: This approach provides the final theoretical benchmark and performance standard, with the null hypothesis for model validation. It provides optimum diversification by allocating an equal percentage of 5.26% to each of the remaining 19 stocks, providing optimum sector fit based on the actual market structure. It is characterized by complete robustness against parameter uncertainty since it does not involve any forecasting inputs and hence has an estimation error of zero. The portfolio, hence, inherits its volatility of 1.0333% from the overall market risk profile. It is designed with an expected return of zero; hence, it is expected to be outperformed by any successful strategy.

STRATEGY 16: RANDOM WALK RPP

Table 3.32: Random Walk RPP Portfolio Composition

Stock	Weight (%)	Exp Return (%)	Volatility (%)	Sector
NigFlourMill	6.78	0.0000	3.2083	Consumer Goods
Bua	6.56	0.0000	2.2618	Industrial Goods
Dangote	6.52	0.0000	2.2303	Industrial Goods
Guinness	6.51	0.0000	2.4711	Consumer Goods
Seplat	6.45	0.0000	1.9528	Oil & Gas
Nestle	6.37	0.0000	1.7154	Consumer Goods
Stanbic	6.35	0.0000	2.7234	Banking
MTN	6.28	0.0000	2.2728	Telecommunications
Total	6.26	0.0000	1.6827	Oil & Gas
Cadbury	5.37	0.0000	3.6317	Consumer Goods
PZ	5.37	0.0000	4.0015	Consumer Goods
Lafarge	5.31	0.0000	2.4938	Industrial Goods
GTCO	4.96	0.0000	2.3847	Banking
Zenith	4.89	0.0000	2.4311	Banking
DangSugar	4.02	0.0000	3.4304	Consumer Goods
Access	3.85	0.0000	2.8151	Banking
Sterling	3.04	0.0000	3.4947	Banking
Honeywell	2.67	0.0000	4.3285	Consumer Goods
Oando	2.43	0.0000	4.6743	Oil & Gas

Portfolio Mathematics: $\mu_p = 0.0000\%$, $\sigma_p = 0.9127\%$ with risk parity and zero expected returns

Key Performance Insights: This technique follows the risk-parity technique, where it assigns weights based on the inverse volatility of each stock to ensure that each stock in the portfolio contributes evenly to the risk. The technique ensures that the portfolio is well balanced, with the maximum weight at 6.78% in the Nigerian Flour Mills and the minimum at 2.43% in the most volatile stock, Oando. The technique ensures that it controls the volatility well, with daily volatility at 0.9127%, between the concentrated Mean-Variance Optimization technique and the Equal Weight Portfolio technique. The technique ensures that it is sector-neutral, with well-dispersed weights among the sectors. The technique ensures that it is theoretically consistent by maintaining the assumption that the market is efficient. The technique is valuable in that it offers a pure risk-parity benchmark that allows one to measure the forecasting strategies since it does not consider the return.

3.3.3 Comprehensive Key Research Findings

3.3.3.1 Performance Hierarchy and Machine Learning Dominance

The data clearly reveals that the winner here is the Machine Learning approach. The top four approaches in the Random Forest group had Sharpe Ratios that ranged from a low of 0.6270 to a staggering 1.5340. The Sharpe Ratio achieved by the Random Forest Mean-Variance (MVO) approach was the highest, and it reached a ratio that was seven times higher than the highest ratio achieved by the traditional approach. This approach also delivered a return of 0.66% daily with a volatility of just 0.43%, thereby clearly establishing that you can get higher returns with lower risk if your predictions are correct. The Machine Learning approach essentially halved the risk and tripled the returns.

3.3.3.2 Strategic Insights and Portfolio Construction

The model demonstrated true “market intelligence” in terms of allocations. The model was very optimistic about Oil Gas, allocating 30.04% to Total Nigeria, as well as Telecommunications, with 12.55% allocated to MTN. The truly impressive part, however, was how “picky” the model was within the Consumer Goods sector, not only allocating to the sector but picking winners while ignoring the losers. This implies that the Random Forest model is much superior to traditional econometric models for sector identification.

3.3.3.3 Emerging Market Implications

This shows that the Nigerian market may not be as efficient as it appears. The fact that it’s possible to find these patterns shows that the information isn’t priced in real time, leaving room for smart money to take advantage of it. For a portfolio manager in Nigeria, this isn't theory—it's a guidebook to how to increase profits. Though the NGX has its problems, it's clear that sophisticated techniques, such as machine learning, are equally valid in Nigeria as elsewhere. The next step needs to be to pit these models against real-world costs.

CHAPTER FOUR

SUMMARY, CONCLUSION AND RECOMMENDATIONS

4.1 Summary

In summary, this study compares time series models (ARIMA-GARCH) and a machine learning model (Random Forest) to see the better model at forecasting some selected 19 Nigeria Stock returns and volatility used to implement four portfolio optimization strategies. Four models were used in this research, the first is the Autoregressive Integrated Moving Average (ARIMA) model used to forecast returns and the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model used to forecast volatility, the second is the Random Forest Machine learning model used to forecast both returns and volatility, the third is the Random Walk model which assumes Weak-Form Efficient Market Hypothesis (EMH) also used to forecast returns and volatility, the fourth is the Hybrid model selection which is basically made up of the first three models but selected by the best performer for each by RMSE. These four models were then combined with four different portfolio optimization strategies: the Minimum Variance Portfolio (MVP), Risk Parity, the Equal Weight Portfolio (EWP), and the traditional Mean-Variance framework. This created 16 model strategy combination and the goal was to see which combination actually delivers the best risk-adjusted returns in the Nigerian market. The analysis showed that the ARIMA-GARCH was better mostly in forecasting returns and the Random Forest was better in volatility forecasting. The portfolio optimization results showed that the Machine learning model (Random Forest) had the top performing model strategy combinations, combined with the mean-variance optimization had the highest risk-adjusted returns while combined with the minimum variance portfolio had the lowest risk. This research shows that Nigerian stocks from different sectors has low correlation, which is great for diversification. Their movements show non-normality and volatility clustering, which justifies using machine learning models like Random Forest. This study provides an evidence for building better portfolios in the Nigerian stock market by implementing machine learning models with established optimization strategies.

4.2 Conclusion

This study proves that machine learning models, particularly Random Forest, is better for portfolio optimization in the Nigerian stock Market. The performance of Random Forest MVO, achieving a Sharpe Ratio of 1.5340, shows that machine learning can enhance both return and risk (volatility) forecasting compared to traditional models like ARIMA and GARCH. The success of our forecasts, especially against the Random Walk, shows that the Nigerian market isn't purely efficient, past data contains useful patterns that can be used to forecast, which investors can use. This study also shows that traditional models like ARIMA and GARCH can be weak sometimes due their limitation in only being able to model linear patterns and inability to model non-linear complex market structures, so it would be needed to incorporate more complex, non-linear models like the Random Forest especially in the Nigeria market. The effectiveness of Risk Parity strategies, particularly when enhanced with machine learning inputs, the ability to balance risk among the stocks using the correlation works well for diversification. Like wise, the Minimum Variance strategy is great for investors that want to minimize the portfolio risk as much as possible, also made much better with machine learning inputs. This research has provided investors with strategies recommended based on their investment goals. This research helps us get a clearer picture of how modern tools can actually work in a market like ours. The results show that while the Nigerian market has its own problems, it still holds a lot of potential for investors who are willing to use the right tools. Seeing how well the machine learning model performed here is really encouraging, and I believe it is time we start pushing for more of these advanced methods across other African financial markets

4.3 Recommendations

Based on the findings of this research, the following recommendations are offered for different investors in the Nigerian financial ecosystem:
For investors with high risk tolerance, the Random Forest Mean-Variance Optimization strategy is the recommended strategy because it offers a high expected daily returns 0.6599% with a high Sharpe ratio of 1.5340 which indicates a high return for risk performance. Investors looking to use this strategy should also be also be aware of the concentrated nature, in this case with just 12 out of 19 stocks. For investors who would like to diversify the risk in their portfolio, the Random Forest Risk Parity Portfolio would be the recommended choice because of its diversified nature. The performance in terms of return for this strategy was the highest at 0.3116% with well-distributed risk among the assets with a maximum stock weight of 6.38%. This strategy would be suitable for investors who seek stability in their performance. For those with fewer resources, they can look at the Random Forest Equal Weight Portfolio. Although not the best, they still provide a good return of 0.3284% for an equally weighted portfolio of all stocks.

References

Adediran, I. A., & Adeleke, I. A. (2013). Forecasting stock returns on the nigerian stock market: A comparative analysis of arima, arch and garch models. *International Journal of Economics and Financial Issues*, *3*(3), 662–671.

Adhikari, R., Agrawal, S., & Singh, R. (2024). A comparative study of arima and machine learning models for stock price forecasting in emerging markets. *Journal of Financial Analytics*, *12*(1), 45–62.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307–327.

Box, G. E., & Jenkins, G. M. (1976). *Time series analysis: forecasting and control*. Holden-day.

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*, *22*(5), 1915–1953.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, 987–1007.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of finance*, *25*(2), 383–417.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.

Khaidem, L., Saha, S., & Dey, S. R. (2016). Predicting the direction of stock market prices using random forest. *arXiv preprint arXiv:1605.00003*.

Maillard, S., Roncalli, T., & Teiletche, J. (2010). The properties of equally weighted risk contribution portfolios. *The Journal of Portfolio Management*, *36*(4), 60–70.

Markowitz, H. (1952). Portfolio selection. *The journal of finance*, *7*(1), 77–91.

Oguntade, O., & Adeyemi, A. (2019). Machine learning approaches for stock price prediction: A case study of the nigerian stock exchange. *Journal of Computer Science and its Applications*, *26*(2), 89–102.

Olowe, R. A. (2009). Volatility, mean reversion and the nigerian stock market. *African Review of Money Finance and Banking*, 77–104.

Singh, A., & Singh, P. (2021). A comparative analysis of machine learning techniques for stock market prediction. *International Journal of Financial Engineering*, *8*(3), 2150024.