Current Science Research Bulletin

Volume 02 (06), June 2025, pp. 183-191

CrossRef DOI: 10.55677/csrb/02-V02I06Y2025

Available on: https://csrbjournal.org/index.php/csrb

e-ISSN: 3050-7103

p-ISSN: 3050-709X

Stock Selection Using Ward Clustering and Mean Absolute Deviation Portfolio Construction on the IDX Quality30 Index

Fairuz Dwi Najla[1], Triastuti Wuryandari*[2], Rahmila Dapa[3]

[1,2,3]Department of Statistics, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia

ABSTRACT: Investment in stocks offers high potential returns but also involves risks that need to be managed through diversification. This study aims to find the best stocks for constructing a Mean Absolute Deviation (MAD) portfolio by using Ward Clustering to select stocks from the IDX Quality30 index. The data includes stocks listed continuously from November 2023 to October 2024. Stocks were grouped based on financial ratios: Earnings Per Share, Price to Earnings Ratio, Debt to Equity Ratio, and Return on Equity. Ward Clustering was applied to create clusters with low internal variation by merging groups that minimize the increase in Sum of Squared Error. The Silhouette Coefficient was used to determine that five clusters were optimal. From each cluster, stocks with the highest positive expected return were chosen. The MAD model then calculated the best portfolio weights by minimizing risk measured by absolute deviation, with limits on minimum return and stock weights. The final portfolio contains four stocks: ADRO (30%), BBCA (30%), MIKA (10%), and UNTR (30%), while ACES was excluded due to its insignificant contribution to portfolio performance . The portfolio achieved an expected return of 0.00090 and risk of 0.01018. A Sharpe Ratio of 0.07224 indicates the portfolio outperforms risk-free investments, making it a suitable option for investors looking for a balanced risk-return portfolio within the IDX Quality30 index.

KEYWORDS: IDX Quality30 Index, Ward Clustering, Mean Absolute Deviation Portfolio, Sharpe Index

1. INTRODUCTION

Investment is a strategy to achieve financial stability by allocating funds into productive assets for future returns. Stocks are popular investment instruments traded on the Indonesia Stock Exchange (IDX), which includes 45 indices such as IDX Quality30. This index contains 30 stocks with strong fundamentals and stable financial performance, making it a reliable reference for stock selection. Investors seek returns through dividends or capital gains, but higher returns involve higher risks. Diversification through portfolio construction can reduce risk and optimize returns, with stock proportions determined by cluster analysis (Gubu et al., 2020).

Financial ratios take a crucial role in investment decisions. Ratih et al. (2013) identified four key ratios affecting stock prices: Earnings Per Share (EPS), Price to Earnings Ratio (PER), Debt to Equity Ratio (DER), and Return on Equity (ROE). Ward Clustering groups stocks based on similarities in these ratios, creating more homogeneous clusters.

The Mean Variance Efficient Portfolio (MVEP) model by Markowitz (1952) is a foundational approach but has limitations due to assumptions and computational complexity (Bower & Wentz, 2005). The Mean Absolute Deviation (MAD) model by Konno and Yamazaki (1991) uses absolute deviation as a risk measure and can be solved with linear programming. Anugrahayu and Azmi (2023) found that MAD produces higher Sharpe Index and returns with lower risk compared to MVEP.

This study constructs a MAD portfolio by selecting stocks through Ward Clustering, validated by the Silhouette Coefficient, within the IDX Quality30 index. Stocks with the highest positive expected returns from each cluster are chosen, and their optimal weights are calculated using the Simplex method. Portfolio performance is evaluated using the Sharpe Index to support more efficient investment decisions.

2. LITERATURE REVIEW

Investment involves postponing current consumption by allocating funds into productive assets to achieve capital appreciation and returns, protect against inflation, and meet long-term financial goals (Hartono, 2022). Stocks represent partial ownership in companies and offer returns through dividends and capital gains, allowing flexible investment strategies (Adnyana, 2020). The IDX Quality30 index tracks 30 stocks selected from IDX80 based on high profitability, strong solvency, consistent

Corresponding Author: Triastuti Wuryandari

earnings growth, and positive liquidity, using criteria such as Return on Equity (ROE), Debt to Equity Ratio (DER), and Earnings Per Share (EPS) growth.

Financial ratios are tools in fundamental analysis to evaluate a company's performance and financial health. This study uses three types: profitability ratios (Earnings Per Share [EPS] and Return on Equity [ROE]), a market ratio (Price to Earnings Ratio [PER]), and a solvency ratio (Debt to Equity Ratio [DER]). EPS measures profit per share, ROE shows net income relative to equity, PER indicates stock valuation, and DER reflects financial leverage.

$$EPS_i = \frac{Net\ income\ attributable\ to\ the\ shareholders\ of\ stock\ i}{Number\ of\ outstanding\ shares\ of\ stock\ i} \quad (1)$$

$$PER_i = \frac{Market\ price\ per\ share\ of\ stock\ i}{EPS_i} \quad (2)$$

$$DER_i = \frac{Total\ liabilities\ of\ the\ company\ issuing\ stock\ i}{Total\ equity\ invested\ by\ shareholders\ in\ stock\ i} \quad (3)$$

$$ROE_i = \frac{Net\ income\ attributable\ to\ the\ shareholders\ of\ stock\ i}{Total\ equity\ invested\ by\ shareholders\ in\ stock\ i} \times 100\% \quad (4)$$

Cluster analysis is a multivariate method used to group objects based on similarities across multiple variables, aiming for lower variation within clusters than between them (Hair et al., 2019). This study uses a hierarchical approach with Ward Clustering. As clustering relies on distance measures sensitive to variable scales, standardization is necessary to prevent bias from variables with larger ranges. To ensure consistent scales and reduce outlier effects, this study applies the Maximum Absolute Scaler, which preserves data relationships during the standardization process.

$$x_{ik_{scaled}} = \frac{x_{ik}}{max(|x_k|)} \quad (5)$$

where $x_{ik_{scaled}}$: the standardized value of the i-th object on the k-th variable; $max(|x_k|)$: the maximum absolute value of the data for the k-th variable.

Before performing clustering to obtain optimal groupings, two assumptions must be satisfied (Hair et al., 2019).

1. The assumption of a representative sample is evaluated using the Kaiser–Meyer–Olkin (KMO) test, which measures the adequacy of the sample in relation to each indicator. The KMO statistic ranges from 0 to 1, where a value between 0.5 and 1 indicates that the sample is appropriate and representative of the population's characteristics.

$$KMO = \frac{\sum_{k=1}^{p} \sum_{l=1,l\neq k}^{p} r_{kl}^2}{\sum_{k=1}^{p} \sum_{l=1,l\neq k}^{p} r_{kl}^2 + \sum_{k=1}^{p} \sum_{l=1,l\neq k}^{p} \rho_{kl}^2} \quad (6)$$

where p : the number of variables; r_{kl}^2 : the squared correlation coefficient between variable k and variable l; ρ_{kl}^2 : the squared partial correlation coefficient between variable k and variable l.

2. The assumption of non-multicollinearity is assessed using the Variance Inflation Factor (VIF), which detects strong linear relationships among independent variables that may bias clustering results. This assumption is considered fulfilled if all VIF values are below 10.

$$VIF_k = \frac{1}{1 - {R_k}^2} \quad (7)$$

where ${R_k}^2$: the coefficient of determination between the independent variable x_k d and the other independent variables, ${R_k}^2 = \frac{SSR_k}{SST_k}$; SSR_k : Sum of Squares Regression where $SSR_k = \sum_{i=1}^{n}(\hat{x}_{ik} - \bar{x}_k)^2$; SST_k : Sum of Squares Total where $SST_k = \sum_{i=1}^{n}(x_{ik} - \bar{x}_k)^2$, n : the number of objects.

In the Ward method, inter-object dissimilarity is measured using the Squared Euclidean distance, since this approach evaluates cluster similarity by the total squared deviation across all variables rather than a single metric (Hair et al., 2019). By omitting the square root in the Euclidean distance, larger distances are emphasized and computation is expedited. The formula appears in Equation (8) (Murtagh & Legendre, 2014).

$$d_{ij}^2 = \sum_{k=1}^{p}(x_{ik} - x_{jk})^2 \quad (8)$$

where d_{ij}^2 : the Squared Euclidean distance between object x_{ik} and object x_{jk}; x_{ik} : the value of the i-th object for the k-th variable; x_{jk} : the value of the j-th object for the k-th variable.

Ward's method operates by iteratively merging the pair of clusters whose fusion results in the smallest increase in total within-cluster variance, thereby producing a new cluster with minimal internal dispersion (Hair et al., 2019). According to Ward (1963), the loss of cluster homogeneity is quantified by the increase in the Sum of Squared Errors (SSE) following each merger. Initially, each object is treated as its own cluster, so that the number of clusters c equals the total number of objects n. At each step,

Corresponding Author: Triastuti Wuryandari

the increase in SSE between two clusters (ΔSSE_{ij}) is computed. If a cluster consists of a single object, its SSE is defined as zero. For a cluster containing more than one object, SSE is calculated using Equations (9).

$$SSE = \sum_{i=1}^{n} (\boldsymbol{x_i} - \boldsymbol{\bar{x}})^T (\boldsymbol{x_i} - \boldsymbol{\bar{x}}) \tag{9}$$

where $\mathbf{x_i}$: the column vector of the i-th object's values; $\bar{\mathbf{x}}$: the column vector of the cluster mean.

Murtagh and Legendre (2014) demonstrated that when two clusters have equal size, ΔSSE_{ij} equals one-half the squared Euclidean distance between their centroids. The formula for the ΔSSE_{ij} measure is presented in Equation (10) (Mitran, 2019). The pair of clusters yielding the smallest ΔSSE_{ij} are merged at each iteration, reducing the total number of clusters by one. This process is repeated until all objects form a single cluster or until the desired number of clusters is reached.

$$\begin{aligned} \Delta SSE_{ij} &= \frac{1}{2}\left(\boldsymbol{x_i} - \boldsymbol{x_j}\right)^T \left(\boldsymbol{x_i} - \boldsymbol{x_j}\right) \\ \Delta SSE_{ij} &= \frac{1}{2}\sum_{k=1}^{p}\left(x_{ik} - x_{jk}\right)^2 \\ \Delta SSE_{ij} &= \frac{1}{2} d_{ij}^2 \end{aligned} \tag{10}$$

The evaluation of clustering results is conducted using the Silhouette Coefficient, which measures how well each object lies within its assigned cluster and how distinctly it is separated from other clusters. This metric provides an assessment of the cohesion and separation of the formed clusters.

$$s(i) = \frac{b(i) - a(i)}{maks(a(i), b(i))} \tag{11}$$

$$SC = \frac{1}{n}\sum_{i}^{n} s(i) \tag{12}$$

where $a(i)$: the average distance between object i and all other objects within the same cluster; $b(i)$: the minimum average distance between object i and all objects in any other cluster; $s(i)$: the Silhouette Coefficient value for the i-th object; SC : the overall Silhouette Coefficient calculated across all n data objects.

The range of the Silhouette Coefficient value is between –1 and 1 for each object i, with interpretations as proposed by Kaufman and Rousseeuw (1989).

Table 1. Silhouette Coefficient Value Interpretation

Silhouette Coefficient Range	Interpretation
0,71 – 1,00	The resulting clustering structure is very strong.
0,51 – 0,70	The resulting clustering structure is moderately strong.
0,26 – 0,50	The resulting clustering structure is weak.
≤ 0,25	There is no substantial clustering structure.

Return and risk constitute the two principal, interrelated dimensions of investment performance. Higher returns are typically associated with greater risk, reflecting the "high risk, high return" principle. Realized return is the actual return computed from historical data, whereas expected return represents the investor's forecast of future performance. Realized return is calculated using continuously compounded (log) returns as defined in Equation (13) (Tsay, 2005), while expected return follows Equation (14) (Maruddani, 2019).

$$R_{i,t} = ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right) \tag{13}$$

$$E(R_i) = \frac{1}{T}\sum_{t=1}^{T} R_{i,t} \tag{14}$$

where $R_{i,t}$: the return of stock i in period t; $P_{i,t}$ and $P_{i,t-1}$: the prices of stock i at periods t and t−1; $E(R_i)$: the expected return of stock i; T : the total number of daily return observations.

Corresponding Author: Triastuti Wuryandari

Portfolio realized return is obtained as the weighted average of each constituent's realized returns over the sampling period, and portfolio expected return is the weighted average of each constituent's expected returns. The portfolio realized return formula appears in Equation (15) (Tsay, 2005), and the portfolio expected return formula is provided in Equation (16) (Maruddani, 2019).

$$R_{p,t} = \sum_{i=1}^{n_p} w_i R_{i,t} \tag{15}$$

$$E(R_p) = \mu_p = \sum_{i=1}^{n_p} w_i E(R_i) \tag{16}$$

where $R_{p,t}$: return portfolio in period t; $E(R_p)$: portfolio's expected return; n_p : the number of stocks in the portfolio, w_i : the weight of stock i such that $\sum_{i=1}^{n_p} w_i = 1$.

Portfolio risk measures the dispersion of realized returns around the expected portfolio return. This risk is quantified by the portfolio standard deviation, which indicates the extent to which portfolio returns deviate from their mean.

$$\sigma_p = \sqrt{\frac{\sum_{t=1}^{T}(R_{p,t} - E(R_p))^2}{T-1}} \tag{17}$$

According to Konno and Yamazaki (1991), the Mean Absolute Deviation (MAD) portfolio optimization model measures portfolio risk as the average absolute deviation between realized returns and expected returns. MAD minimizes risk by directly minimizing these absolute deviations. As an extension of the Mean-Variance Efficient Portfolio (MVEP) model, MAD dispenses with the normality assumption and obviates the need for complex variance–covariance matrix calculations (Konno & Yamazaki, 1991). The MAD value as a measure of stock risk in the MAD portfolio can be calculated by averaging the absolute deviations between the return values and the expected returns of the stocks, as shown in Equation (18) (Bower & Wentz, 2005).

$$MAD_i = \frac{1}{T}\sum_{t=1}^{T} |R_{i,t} - E(R_i)| \tag{18}$$

The MAD optimization model seeks to minimize total portfolio risk, defined as the weighted sum of each asset's MAD, subject to a minimum return constraint and weight bounds. Formally, the linear programming formulation is:
Objective function (minimize portfolio risk):

$$\sigma(w) = \sum_{i=1}^{n_p} MAD_i w_i \tag{19}$$

Subject to:

$$\sum_{i=1}^{n_p} E(R_i)\, w_i \geq R_{min} \tag{20}$$

$$\sum_{i=1}^{n_p} w_i = 1 \tag{21}$$

$$0 \leq w_i \leq u_i \text{ untuk } i = 1,2,\dots,n_p \tag{22}$$

where: $\sigma(w)$: objective function; R_{min} : The minimum portfolio return, $R_{min} = \frac{1}{n_p}\sum_{i=1}^{n_p} E(R_i)$; w_i : the weight of stock i; n_p : the number of stocks in the portfolio, u_i : the maximum weight allowed for each stock in the portfolio.

The portfolio optimization problem based on linear programming is solved using the Simplex method to determine the optimal investment weights for each stock by utilizing Excel Solver software. The Simplex method is a mathematical approach that starts from a feasible basic solution and iteratively moves to adjacent feasible basic solutions until reaching the optimal solution. Before applying the Simplex method, the objective function and constraint set in the MAD model must be converted into standard form (Wirdasari, 2009), which involves the following steps:

1. The right-hand side (RHS) value of the objective function must be zero.
2. The right-hand side (RHS) values of the constraints must be positive.
3. Constraints that are not in standard equation form must be transformed as follows:
   a. For "≤" constraints, add a non-negative slack variable (s) to the left-hand side (LHS) to convert it into an equation.
   b. For "≥" constraints, subtract a surplus variable (s) and add an artificial variable (r) to the LHS to form an equation.
   c. For "=" constraints, add an artificial variable (r) to the LHS.

Portfolio performance evaluation is essential to assess whether the portfolio's return can exceed the return of the benchmark portfolio and whether the return level is proportional to the risk borne by the investor. The Sharpe Ratio is measured by comparing the difference between the portfolio return and the risk-free rate with the portfolio's standard deviation or risk. The formula for the Sharpe Ratio is presented in Equation (23) (Anugrahayu and Azmi, 2023).

$$SR_p = \frac{E(R_p) - R_f}{\sigma_p} \quad (23)$$

where SR_p : sharpe Ratio of the portfolio; $E(R_p)$: expected return of the portfolio; R_f : risk-free-rate; σ_p : portfolio risk.

3. DATA AND METHODOLOGY

This study utilizes secondary data, including the biannual IDX Quality30 major index evaluations published from August 2023 to February 2025. Financial ratio data—Earnings Per Share (EPS), Price to Earnings Ratio (PER), Debt to Equity Ratio (DER), and Return on Equity (ROE)—for 2023 and 2024 were collected from 21 stocks consistently included in the IDX Quality30 index. Additionally, daily closing prices and monthly Bank Indonesia interest rates (BI Rate) from November 1st 2023, to October 31th 2024, were used. Data analysis was conducted using R Studio and Excel Solver through the following steps:

1. Preparing data.
2. Standardizing EPS, PER, DER, and ROE using the Maximum Absolute Scaler.
3. Testing assumptions of representative samples and non-multicollinearity.
4. Calculating distances between data points using Squared Euclidean distance.
5. Clustering stocks with Ward Clustering method.
6. Determining the optimal number of clusters via the Silhouette Coefficient.
7. Calculating return and expected return for each stock.
8. Selecting representative stocks from each cluster for the portfolio.
9. Calculating the Mean Absolute Deviation (MAD) for each stock.
10. Computing the portfolio's minimum return.
11. Identifying the objective function and constraints.
12. Determining optimal investment weights for each stock using the Simplex method.
13. Calculating the portfolio's expected return and risk.
14. Evaluating portfolio performance using the Sharpe Index.

4. RESULTS AND DISCUSSION

The data used for clustering consist of EPS, PER, DER, and ROE for 21 stocks, totaling 84 data points. Descriptive statistics for these four financial ratio variables are presented in Table 2.

Table 2. Descriptive Statistics of Financial Ratio Variables

Variables	N	Minimum	Mean	Maximum	Std. Deviation
EPS (IDR)	21	5,81	569,12	5541,62	1176,33
PER (x)	21	2,86	15,14	37,39	10,60
DER (x)	21	0,09	1,15	5,84	1,72
ROE (%)	21	5,04	16,41	37,26	8,68

Since these financial ratio variables have different scales, a data standardization process is required. The Maximum Absolute Scaler method is applied to normalize the variable scales, as described by Equation (5). The standardized data must meet two assumptions required for cluster analysis: representative sample and non-multicollinearity.

1. Representative Sample
   The output generated using R Studio shows a Kaiser-Meyer-Olkin (KMO) value of 0.61, which exceeds the minimum threshold of 0.50. This indicates that the assumption of a representative sample is satisfied, and the sample adequately reflects the characteristics of the population, making the data suitable for the clustering process.
2. Non-multicollinearity
   The Variance Inflation Factor (VIF) values were examined to ensure the absence of multicollinearity. The VIF values obtained using R Studio are 1.14990 for EPS, 1.26878 for PER, 1.03715 for DER, and 1.20492 for ROE. All variables have

Corresponding Author: Triastuti Wuryandari

VIF values below 10, indicating that there is no significant multicollinearity or strong linear relationships among the variables.

Stock grouping using the Ward Clustering method was carried out through the following stages. As the initial step of the clustering process, each object was considered as an individual cluster, resulting in the initial number of clusters (c) being equal to the number of objects (n), which in this case is 21 clusters. Subsequently, the increase in the Sum of Square Error (SSE) between two clusters or pairs of objects (ΔSSE_{ij}) was calculated using the formula provided in Equation (10).

$$\Delta SSE_{11} = \frac{1}{2}((0{,}00869 - 0{,}00869)^2 + (0{,}41915 - 0{,}41915)^2 + (0{,}04366 - 0{,}04366)^2 + (0{,}36715 - 0{,}36715)^2) = 0$$

$$\Delta SSE_{12} = \frac{1}{2}((0{,}00869 - 0{,}02803)^2 + (0{,}41915 - 0{,}23953)^2 + (0{,}04366 - 0{,}09675)^2 + (0{,}36715 - 1)^2) = 0{,}21797$$

$$\vdots$$

$$\Delta SSE_{20\,21} = \frac{1}{2}((0{,}04227 - 1)^2 + (0{,}37903 - 0{,}11917)^2 + (0{,}14555 - 0{,}15154)^2 + (0{,}41016 - 0{,}65728)^2) = 0{,}52294$$

$$\Delta SSE_{21\,21} = \frac{1}{2}((1 - 1)^2 + (0{,}11917 - 0{,}11917)^2 + (0{,}15154 - 0{,}15154)^2 + (0{,}65728 - 0{,}65728)^2) = 0$$

The next step involves merging the two clusters that yield the smallest ΔSSE_{ij} value. In the initial process, the smallest ΔSSE_{ij} was found between clusters 10 and 17, with a value of 0.00385. Therefore, the first merging was performed between cluster 10 and cluster 17. Steps (2) and (3) were then repeated iteratively until the number of clusters ranged from 10 to 2.

The quality of the clustering results was evaluated using the Silhouette Coefficient validity index to determine the most optimal number of clusters. The Silhouette Coefficient values, calculated using Equations (11) and (12), are presented in Table 3.

Table 3. Silhouette Coefficient Value

Number of Clusters	Silhouette Coefficient Value
2	0,59109
3	0,55122
4	0,59299
5	0,66655
6	0,56797
7	0,51887
8	0,47919
9	0,42811
10	0,40266

Based on Table 3, the optimal number of clusters is five, which corresponds to the highest Silhouette Coefficient value of 0.66655. This value falls within the range of 0.50 to 0.70, indicating that the resulting clustering structure is moderately strong. The stock members of the five clusters are presented in Table 4.

Table 4. Stock Members of the 5 Clusters

Cluster	Stock Members	Number of Members
Cluster 1	ACES, ASII, BFIN, BTPS, INCO, INTP, KLBF, MNCN, TKIM, TLKM	10
Cluster 2	ADMR, ADRO, PTBA, SIDO	4
Cluster 3	BBCA, BBRI, BMRI	3
Cluster 4	CPIN, MIKA, SCMA	3
Cluster 5	UNTR	1

Stock returns were calculated using daily closing prices, as defined in Equation (13), over 239 periods for each of the 21 stocks. An example of the return calculation for the stock ADRO for the second period on November 2, 2023, is presented below.

$$R_{ADRO,2} = \ln\left(\frac{P_{ADRO,2}}{P_{ADRO,1}}\right) = \ln\left(\frac{2390}{2410}\right) = -0{,}00833$$

Corresponding Author: Triastuti Wuryandari

Representative stocks from each cluster were selected based on the highest positive expected return, which was determined by averaging the stock's return values using Equation (14). Table 5 presents the five portfolio stocks that exhibit the highest positive expected return within each cluster. An example of the expected return calculation for the stock ADRO is provided below.

$$E(R_{ADRO}) = \frac{-0{,}00833 + 0{,}04099 + \cdots + (-0{,}00551)}{238} = 0{,}00171$$

Table 5. Top Expected Return of Representative Stocks from Each Cluster

No	Stock	Expected Return
1	ACES	0,00052
2	ADRO	0,00171
3	BBCA	0,00074
4	MIKA	0,00005
5	UNTR	0,00055

The constructed portfolio was optimized using the Mean Absolute Deviation (MAD) approach, a method aimed at minimizing risk while maintaining the expected return desired by investors. The MAD value of each portfolio stock was calculated using Equation (18) and is presented in Table 6. An example of the MAD calculation for the stock ACES is also provided.

$$MAD_{ACES} = \frac{|0 - 0{,}00052| + |0{,}03096 - 0{,}00052| + \cdots + |0{,}00557 - 0{,}00052|}{238}$$

$$MAD_{ACES} = \frac{0{,}00052 + 0{,}03044 + \cdots + 0{,}00505}{238} = 0{,}01709$$

Table 6. MAD Values of Each Portfolio Stock

No	Stock	MAD
1	ACES	0,01709
2	ADRO	0,01353
3	BBCA	0,01000
4	MIKA	0,01396
5	UNTR	0,01155

Before calculating the investment weights using the Simplex method, the objective function and set of constraints must be converted into standard form. For the first constraint, the minimum return of the portfolio must be calculated in advance. For the third through seventh constraints, the upper limit of investment weight for each stock was determined based on investor preferences, which was set at 30%. The complete objective function and constraints, expressed in standard form, are formulated as follows:

Minimize:

$$\sigma(w) = 0{,}01709w_1 + 0{,}01353w_2 + 0{,}01000w_3 + 0{,}01396w_4 + 0{,}01155w_5 + M(r_1 + r_2)$$

Subject to:

$$0{,}00052w_1 + 0{,}00171w_2 + 0{,}00074w_3 + 0{,}00005w_4 + 0{,}00055w_5 - s_1 + r_1 = 0{,}00071$$
$$w_1 + w_2 + w_3 + w_4 + w_5 + r_2 = 1$$
$$w_1 + s_2 \leq 30\%$$
$$w_2 + s_3 \leq 30\%$$
$$w_3 + s_4 \leq 30\%$$
$$w_4 + s_5 \leq 30\%$$
$$w_5 + s_6 \leq 30\%$$

The portfolio optimization problem was solved using Excel Solver. A linear programming model was formulated and solved using the Simplex method provided in Excel Solver. The optimization process involved seven iterations to obtain the optimal investment weights for each stock in the portfolio, as shown in Table 7.

Corresponding Author: Triastuti Wuryandari

Table 7. Investment Weights of Each Portfolio Stock

No	Stock	Weight
1	ACES	0
2	ADRO	0,3
3	BBCA	0,3
4	MIKA	0,1
5	UNTR	0,3

The expected return of the portfolio reflects the investor's anticipated return on investment. This value was obtained by aggregating the products of each stock's expected return and its investment weight in the portfolio, as calculated using Equation (16).

$$E(R_p) = 0{,}00052(0) + 0{,}00171(0{,}3) + 0{,}00074(0{,}3) + 0{,}00005(0{,}1) + 0{,}00055(0{,}3)$$

$$E(R_p) = 0{,}00090$$

Portfolio risk was assessed by first determining the realized return of the portfolio for each period t, using Equation (15). The standard deviation of the portfolio return, calculated using Equation (17), served as a measure of portfolio risk.

$$\sigma_p = \sqrt{\frac{(0 - 0{,}00090)^2 + (0{,}00015 - 0{,}00090)^2 + \cdots + (0 - 0{,}00090)^2}{237}}$$

$$\sigma_p = 0{,}01018$$

Portfolio performance was evaluated using the Sharpe Ratio, which measures the excess return per unit of risk by comparing the difference between the expected return of the portfolio and the risk-free rate against the portfolio's risk. The risk-free rate used in this analysis referred to the Bank Indonesia interest rate. The Sharpe Ratio was computed using Equation (23), and the result, rounded to five decimal places, was 0.07224. This value indicates that the portfolio exhibits strong performance by offering returns exceeding the risk-free rate for each unit of risk undertaken by the investor.

$$R_f = \frac{1}{365}\left(\frac{6\% + 6\% + 6\% + \cdots + 6{,}25\% + 6{,}25\% + 6\%}{12}\right) = 0{,}00017$$

$$SR_p = \frac{0{,}00090 - 0{,}00017}{0{,}01018} = 0{,}07224$$

5. CONCLUSIONS

The stock grouping of the IDX Quality30 index, conducted using the Ward Clustering method and validated through the Silhouette Coefficient, determined that the optimal number of clusters is five. The composition of the clusters is as follows: Cluster 1 consists of 10 stocks, Cluster 2 consists of 4 stocks, Cluster 3 consists of 3 stocks, Cluster 4 consists of 3 stocks, and Cluster 5 consists of 1 stock. Representative stocks from each cluster were selected based on the highest positive expected return, resulting in a Mean Absolute Deviation portfolio comprising four stocks with the following optimal investment weights: ADRO (30%), BBCA (30%), MIKA (10%), and UNTR (30%). The stock ACES was excluded from the portfolio due to its insignificant contribution to overall portfolio performance. The resulting portfolio has an expected return of 0.00090 and a risk level of 0.01018. The portfolio's performance, measured by the Sharpe Ratio, yielded a positive value of 0.07224, indicating that the portfolio outperforms the risk-free investment alternative and can be recommended as a viable stock investment option for investors.

REFERENCES

1. Adnyana, I. M. (2020). *Manajemen investasi dan portofolio*. Jakarta: Lembaga Penerbitan Universitas Nasional (LPU-UNAS).
2. Anugrahayu, M. R., & Azmi, U. (2023). Stock portfolio optimization using mean-variance and mean absolute deviation model based on K-medoids clustering by dynamic time warping. *Jurnal Matematika, Statistika dan Komputasi*, *20*(1), 164–183.
3. Bower, B., & Wentz, P. (2005). Portfolio optimization: MAD vs. Markowitz. *Rose-Hulman Undergraduate Mathematics Journal*, *6*(2), Article 3. https://scholar.rose-hulman.edu/rhumj/vol6/iss2/3
4. Gubu, L., Rosadi, D., & Abdurakhman. (2020). Robust mean-variance portfolio selection with Ward and complete linkage clustering algorithm. *Economic Computation and Economic Cybernetics Studies and Research*, *54*(3), 111–127.
5. Hair, J. F., Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Andover: Cengage Learning EMEA.

Corresponding Author: Triastuti Wuryandari

6. Hartono, J. (2022). *Portofolio dan analisis investasi pendekatan modul* (2nd ed.). Yogyakarta: ANDI.
7. Kaufman, L., & Rousseeuw, P. J. (1989). *Finding groups in data: An introduction to cluster analysis*. John Wiley & Sons, Inc.
8. Konno, H., & Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Science, 37*(5), 519–531. https://doi.org/10.1287/mnsc.37.5.519
9. Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77–91. https://doi.org/10.2307/2975974
10. Maruddani, D. A. I. (2019). *Value at risk untuk pengukuran risiko investasi saham: Aplikasi dengan program R*. Ponorogo: WADE GROUP.
11. Mitran, S. (2019). Hierarchical clustering techniques. University of North Carolina. http://mitran-lab.amath.unc.edu/courses/MATH590/biblio/Clustering.ch7.HierarchicalClustering.pdf
12. Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classification, 31*, 274–295. https://doi.org/10.1007/s00357-014-9161-z
13. Ratih, D., Prihatini, A. E., & Saryadi, S. (2013). Pengaruh EPS, PER, DER, ROE terhadap harga saham pada perusahaan sektor pertambangan yang terdaftar di Bursa Efek Indonesia (BEI) tahun 2010-2012. *Jurnal Ilmu Administrasi Bisnis, 3*(1), 83–94.
14. Tsay, R. S. (2005). *Analysis of financial time series* (2nd ed.). John Wiley & Sons, Inc.
15. Wirdasari, D. (2009). Metode Simpleks dalam Program Linier. *Jurnal SAINTIKOM, 6*(1), 276–285.

Corresponding Author: Triastuti Wuryandari