# A CLUSTERING METHOD FOR PORTFOLIO OPTIMIZATION

## Iwan Fadilah✉, Rini Setyo Witiastuti

Management Department, Faculty of Economics, Universitas Negeri Semarang, Semarang, Indonesia

## Abstract

This study aims to test the formation of the optimum portfolio using the cluster method. The data used are of financial statements and stock prices of the companies listed on the LQ-45. index The results of this research show that the cluster method can be used to form the optimal portfolio. This is because using the cluster method; the research samples were divided into three clusters that are united to the same characteristics of each company. Further research can be added the research indicators and research sample used in order that the results obtained are more varied.

## *Abstrak*

*Penelitian ini bertujuan untuk menguji secara empiris pembentukan portofolio optimal menggunakan metode kluster. Data yang digunakan adalah laporan keuangan dan harga saham dari perusahaan-perusahaan yang terdaftar di indeks LQ-45. Hasil dari penelitian ini menunjukkan bahwa metode kluster dapat digunakan untuk membentuk portofolio optimal. Hal ini dikarenakan dengan metode kluster, sampel penelitian dibagi menjadi tiga kluster yang disesuaikan dengan karakteristik masing-masing perusahaan. Diharapkan pada penelitian selanjutnya dapat ditambahkan indikator penelitian dan sampel yang digunakan agar hasil yang didapatkan lebih bervariasi.*

**INTRODUCTION**

The last few years, in Indonesia, demand increased in the capital market investments. This is because the capital markets bring together two interests, i.e., the party that excess funds and those who need funds (Mirah & Wijaya, 2013). In addition, activities in the capital markets have developed, due to the increased desire for business travelers looking for financing sources other than banks (Susilowati et al., 2016). Wijayanto (2010) suggests that alternate funding from within the company's profit came from the company who was arrested, while funding from outside the company can either be a creditor of the debt as well as the nature of the funding of the participation in the form of shares (equity).

Yulianto et al. (2016) also states that companies that give priority to profits and seeking increased investment will pay dividends, but the dividend payment it can also be obtained from detained nor profit from debt (investors). Companies that produce a high and stable profit will get a positive response from investors, and investors are willing to pay its share price higher. Return the stock will rise with increasing stock prices (Sudiyanto & Suharmanto, 2011). So that the capital markets have a function as the most allowed to imprint the party capital for the excess funds. The investment could be in the form of financial assets such as stocks, bonds, or mutual funds (Pardosi & Wijayanto, 2015).
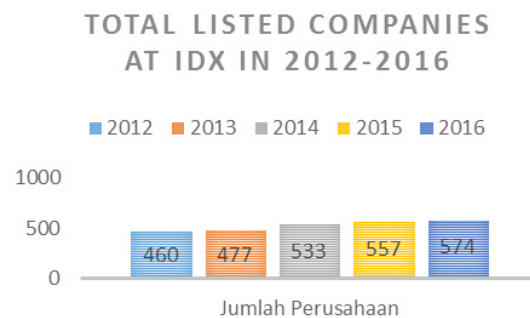
An investor will be faced with profit (return) and the risk if invested in the capital market. Return and risk of investing in the capital market are in price and very volatile. Investors are also aware of the high risks in addition to focusing on the return of investing in capital markets (Wati et al., 2015). The funds that are owned by investors will restrict investors in portfolio formation. So investors can only buy shares at a certain amount. These conditions encourage investors to form the optimal portfolio with the best combination of risk and return (Adiningrum et al., 2016). Because one of the factors taken into consideration in investment decisions is the return (Hermawan, 2012).

Tandelilin (2010) split the risk stake into two types, namely, systematic risk and unsystematic risk. Systematic risk will affect all types of stocks while unsystematic risk only affects one type of stock or a particular sector. Investors are unable to eliminate the systematic risk that impacts the entire existing stocks in the stock market, but investors can reduce unsystematic risk through diversification by forming a portfolio so

that the risks are borne by the investor can minimize.

The funds that are owned by investors will restrict investors in portfolio formation. So investors can only buy shares at a certain amount. These conditions encourage investors to form an optimal portfolio with the best combination of risk and return (Adiningrum et al., 2016).

Portfolio optimization can be done using the model of single-index model and Markowitz. Markowitz's model shows that if wanted to establish a diversified portfolio with good, then the stocks in the portfolio should have a negative correlation coefficient, but in fact in the capital markets there are no two pieces securities that are identical (correlated +1) or counterclockwise (correlated -1) all the time and even the correlation of two securities in the stock market there is not perfect at all (Hartono, 2016).



**TOTAL LISTED COMPANIES AT IDX IN 2012-2016**

■ 2012 ■ 2013 ■ 2014 ■ 2015 ■ 2016

460 477 533 557 574

Jumlah Perusahaan

**Figure 1.** The Graph of Increasing Total Companies at IDX

Based on the above graph, looks each year the number of companies listed on the Indonesia Stock Exchange (IDX) increased during the years of research. This means that every year there is the addition of the company will increase the number of companies listed at the IDX. While the use of the correlation coefficient in the calculation of this will be increasingly difficult to determine the optimal portfolio because currently the greater number of shares traded (Tandelilin, 2010). Thus, the use of the model would complicate the formation of the Markowitz portfolio optimal at this time.

Single index model is a development of the model Markowitz. In contrast to the model in the model, Markowitz single index also takes into account the risk-free assets. This model can simplify calculations in model Markowitz by providing the input parameters required in the model Markowitz (Jayati et al., 2017).

Witiastuti (2012) performs a comparison between the single-index model with the model naive diversification. Such research mentioned

that the risk on a single index model differs significantly with the naive method of portfolio risk diversification. In a small sample, the performance of a single index model of a portfolio of good nor naive diversification equally inferior. Wati et al. (2015) conduct analysis regarding the use of a single index model in the form of an optimal portfolio. Results of the study indicated that the model index of composition return expectations of 2.71% and the risk of 0.14%. This indicates that the single-index model can minimize the risk because the value of the risk is smaller than the value of return expectations.

In addition, refer to some sources that the model will be used in difficult Markowitz to examine with the number of shares in the complex, while the single-index model is often found not closely look at diversification that will be included in the portfolio (Tandelilin, 2010; Pambudi, 2016). While this research has a goal to help, investors choose stocks that are very much using the cluster method. This clustering method is used also to classify stocks based on the characteristic equation (Pambudi, 2016).

The cluster method has advantages, i.e., its ability to incorporate many factors into consideration in the grouping. (Pambudi, 2016). According to Gudono (2014), cluster analysis algorithm is a technique, not a tool of statistical inference, resulting in the implementation does not have to consider the distribution of the data should be normal or not. However, the researches about the stock portfolio analysis using cluster analysis have been widely used. Subekti et al., (2017) outlines the process of a grouping of stocks on the preparation of the portfolio by using two methods of clustering techniques, namely K-means and average linkage. Tola et al. (2008) performs analysis on the formation of portfolio optimal correlation coefficients by grouping in a cluster.

The cluster method in this research use approach to K-means. According to Babu et al. (2012) that the K-means cluster methods are the method most frequently used the cluster. In addition, the selection of the K-means approach is also based on the research of Nanda et al. (2010) as saying that the method of K-means cluster approach produces a diversified portfolio that is better than the other two approaches, namely Fuzzy approach CMeans and Self Organizing Folders (SOM). The first step in categorizing the stocks within a cluster is done by determining factors into selection criteria into clusters (Elton & Gruber, 1971). The selection criteria for these stocks can be done by considering the ratio of the market against the stock (Nanda et al., 2010).

The research using the cluster approach has been done before by some of the world's researchers. (Babu et al. 2012); (Pambudi, 2016); (Guam & Jiang, 2007); (Lahmiri, 2016); (Zhang & Maringer, 2010); (Nanda et al., 2010). The results of these studies state that the cluster method can be used to help investors choose stocks that will be included in the portfolio optimally.

This study aims to test empirically that the selection of stocks with the method of cluster approach can produce an optimal portfolio that has good diversified and also good performance. Optimal portfolio selection model was first introduced by Harry Markowitz. Markowitz published journal articles in the Journal of Finance in 1952. The famous model with the advice "do not put all your eggs into a basket," because if the basket falls, all the eggs in the basket will be broken. In the context of investments, refers to the advice "do not invest all the funds that we have just one asset, because if those assets fail, then all the funds invested will disappear." Overview of the teaching looks simple, but in theory, Markowitz portfolio showed quantitatively why and how diversification could lower risk portfolio (Tandelilin, 2010).

The important contribution of Markowitz was his findings that the return of assets that is correlated between each other, and are not independent. Therefore, the risk of a portfolio should not be calculated from the sum of all of the risks of existing assets in the portfolio, but also consider the effects of linkages between the asset in return estimating risk portfolio. Contributions due to the existence of the relationship between the return of assets can be represented by the value of the covariance. Covariance is a measure of the absolute which shows the extent of the return of both the securities in the portfolio tend to move together (Tandelilin, 2010).

In the world of finance, a portfolio is used to mention a set of investments that are owned by institutions or individuals. Having a portfolio is often a part of the investment and risk management strategy called diversification. By having some assets, certain risks may be reduced (Hapsari & Chandra, 2012).

Investors who think rationally will select the optimal portfolio. Establishment of the optimal portfolio can be done with the model or the single index model Markowitz. To determine the optimal portfolio with these models then required first is an efficient portfolio (Hartono, 2010). A portfolio is considered efficient when the portfolio generates expected a high return and low-risk Lubis (Merliana et al., 2015). To lower the

risk of an investment, the investor needs to diversify (portfolio). Diversification (portfolio) can mean that investors also need to form a portfolio through the combination of election a number of assets in such a way so that the risk can be minimized without having to lower the return expectation ( Tandelilin, 2010).

Portfolio formation method with classical or traditional methods using only the random diversification or the technique randomly without looking at the characteristics of the investment. Investors assume that the more assets went into the portfolio will be the smaller the risks borne. In fact, at some point, if the amount of the asset is increased then decreased risk will make it marginally less (Jayati et al., 2017). This then continued towards the efficient diversification developed by Markowitz in 1952 (Witiastuti, 2012). Unlike the random diversification, diversification Markowitz will consider the characteristics of assets as components of return expectations as well as the industrial classification of an asset (Jayati et al., 2017).

The Markowitz diversification model was introduced in 1952, the return of assets of mutual correlating and not independent. Calculation of the risk of a portfolio is calculated by considering the relationship between return and risk assets portfolio are measured through the correlation coefficient or covariance (Tandelilin, 2010). On the other hand, models use only the rated Markowitz return expectations, and risk considerations did not consider the risk-free asset (Adiningrum et al., 2016).

The Markowitz diversification model was introduced in 1952, the return of assets of mutual correlating and not independent. Calculation of the risk of a portfolio is calculated by considering the relationship between return and risk assets portfolio are measured through the correlation coefficient or covariance (Tandelilin, 2010). On the other hand, models use only the rated Markowitz return expectations, and risk considerations did not consider the risk-free asset (Adiningrum et al., 2016).

**Clustering Method**

The cluster method is a method that is commonly used in multivariate data analysis (Tola et al., 2008). The goal of clustering is to obtain a partition which means in a set of n variables within groups of data in accordance with its characteristics. Gulo (2002) stated that cluster method could be used to classify stocks. This classification is based on the equation some predetermined characteristics into a single cluster,

so that each cluster can be distinguished from each other. As for previous studies that have used cluster method in the field of a stock portfolio.

Tola et al. (2008) introduces the technique of optimization portfolio formation using the c,luster method. This research was conducted to investigate the formation of the optimal portfolio by comparing model Markowitz and cluster method. The data used are the data from the New York Stock Exchange years 1988-1998. The approach used was cluster model approach single linkage and average linkage. Also used Random Matrix Theory (RMT) as filter data using a model from the formation of Markowitz. Portfolio formed by single linkage method has a higher level of risk compared with the model already done Markowitz filtering with RMT (Random Matrix Theory). While the average linkage shows better results than the single linkage method.

Nanda et al. (2010) made a comparative approach in the cluster method to classify stocks. These approaches include K-Means, Fuzzy C-Means, and Self Organizing Folders (SOM). This research was done on the stock market in India by using some consideration. As for consideration were: the rate of return, the ratio of stock price per earning the (p/e), price to book value (P/BV), price per cash EPS (P/CEPS), the value of the company per EBITDA (earnings before interest, depreciation, amortization & taxes), and capital market per sales. The result was the K-Means method showed better results from fuzzy C and SOM.

Babu et al. (2012) doing research using the cluster approach to K-means. K-means took in this research because the rated approach that is often used in the cluster method. On the other hand, the study adds variable, i.e., HAC (hierarchical agglomerative clustering) in K-means forming a combination, i.e., HRK (Hierarchical Agglomerative and Recursive K-Means Clustering). They used five considerations of financial ratios that thinks it is important to measure the performance of the company. Five of these ratios were: operating margin, ROE, ROTA, Equity to Capital, and receivable turnover. The result was the method of K-means method better than HAC (Hierarchical Agglomerative Clustering). HAC will be better than k-means when the level of certainty is 0.9 if the level of certainty is not the number of accuracies when the threshold of high purity resultant cluster would be more fitting. K-means is one of the simplest forms of grouping. The procedure is simple and straightforward to classify the data provided via a number of clusters. Determination of centroid is done by taking

the first data as the centroid of the first, second and second centroid data, and so on until the required number of the centroid. The next step is to calculate the distance from the point to be on the cluster to the centroid of each existing and grouped according to the distance to the closest centroid. When all points are already entered into the grouping, then the first step is completed.

The next step, it is necessary to calculate back k-new centroid as the barycenters of the resulting group. After having k-centroid, which tested back against k-centroid. The calculation of k-centroid and conducted repeated until k-centroid is not moving anymore. Algorithms are as follows:

The firm size is basically a grouping the companies into several groups, including large companies, medium, and small. The company's scale is a measure used to reflect the large to the small; these are based on the total assets of the company . The size of the company is the company scale as seen from the total assets of the company at the end of the year. Total sales can also be used to measure the magnitude of the company. Because total assets can become a benchmark size of the company, then this can be simplified by transforming into natural logarithm (Ghozali, 2006).

The firm size is average total net sales for the year in question until a couple of years. In this case, the sale is greater than the variable costs and fixed costs; it will be retrieved amount of income before taxes. Conversely, if sales are less than the variable costs and fixed costs, then companies will suffer losses (Brigham & Houston, 2011).

Return on Assets (ROA) is a ratio of profitability that could measure the ability of the company in generating profit from assets that are used for the company's operations (Maftukhah, 2013). According to Tandelilin (2010) is the extent to which depictions of the ROA ability assets owned by the company can generate profit. According to Kamsir (2014), ROA is a ratio that shows results over the total assets used in the company. According to Fabozzy (2000), ROA sees how far the investment that has been implanted was able to provide the expected profit returns and investments are in fact the same as the company's assets are implanted or placed.

ROA was able to measure the ability of a company making a profit in the past and then projected into the future. Assets or the assets in question are all company property, which is obtained from its own capital or foreign capital that has transformed the company into a corporate

asset-assets for the company's operations. So the company's performance can be seen from the financial ratios (Asmi, 2014). In addition, according to the Solikhah and Wahyudin (2017) ROA, ROE, and EPS can exhibit the condition of the good company.

Financial ratio analysis can help the business person and the other financial statement users in assessing the company's financial reports (Khajar, 2010). ROE relies heavily on small or big companies; many companies are still categorized a small company has a relatively small capital, this makes the value of the ROE became small. ROE compared net profit after tax with equity has been invested by the shareholders of the company (Van Horne and Wachowicz, 2005). This ratio indicates the power to generate a return on investment based on the book value of the shareholders and is often used in comparing two or more companies top investment opportunities are good and effective cost management.

Eko (2010) performs a comparison between the single-index model with constant correlation model. The study concluded that the establishment of the portfolio using the constant correlation have better performance if compared with the formation of the optimal portfolio that was created by using a single index model.

Pardosi and Wijayanto (2015) performs a comparison of return and risk to the portfolio including stocks and not including portfolio. The approach of Markowitz used to compare between portfolio and not a portfolio. Results of the study state that there is no distinction between risk and return portfolio stocks and not a portfolio.

The above studies do not pay attention to the diversification of stocks that are included in the portfolio. On the other hand, in this research will put a new method that will hopefully generate stock groups that are well diversified. This study will consider the ratio-the ratio of ROA, ROE, and the size of the company. This is because in previous research, conducted by Nanda et al. (2010) chose members who have best performance ratio will be selected to form the optimal portfolio. As for the frame of mind in this research are as Figure 2.
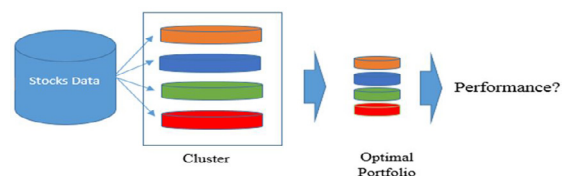


**Figure 2.** Proposed Framework

**METHOD**

This research is quantitative descriptive research types, i.e., calculating the data in order to produce a quantitative interpretation which includes performance measurement (Trihastuti, 2012). The population in this research is the registered shares in the joint stock of har index (IHSG) at the 2012-2016 research period. The sample in this research is the stock listed on the LQ-45 index. Samples were taken from the LQ-45 index because in it there are stocks that traded often. Investors know that inside the LQ-45 index represents stocks of the most liquid. In addition, index LQ-45 is also a composite of several different industrial sectors so that expected to represent the population, i.e., represent stocks from different industry sectors listed on the stock exchange. By using purposive sampling, acquired 22 companies that are examined in this study.

There are two steps in the process data using the cluster method in this research. The first step is done in the cluster method is to calculate the value of the ROE, ROA, the value of the firm's size, return, and variance for each stock. After that, do the normalization of data by changing the previous data into a form the z-score (Pambudi, 2016). Normalization of data useful for the position data and can be compared. In addition, normalization of data is also useful to be able to produce accurate and quality cluster (Mohamad & Usman, 2013). Test of normality in this research was conducted with the help of SPSS 21.0 Kolmogorov Smirnov-test with. Provisions distributed data with normal is when the significance of test results-Kolmogorov Smirnov is greater than 0.05.

The next step is posting validation index to determine how the number of clusters (k) based on the optimal dimensions of the ROA, ROE, and the size of the company. Index validation process carried out with the use of software In their usage, r. package NbClust is validation collation tools in the software R. once known the value of k, then the optimal value of k is used for the grouping to approach K-means (Charrad et al., 2014). The algorithm on the k-means approach used thus categorizing the stocks into k clusters with minimizing the number of squares between the distance of the sample to the center of the cluster for each indicator ROE, ROA, and the size of the company. According to Fan (2010), the algorithm can be formulated as follows:

$$min_c \sum_{i=i}^{n} \overset{min}{\underset{p=1,...,k}{}} \left( \frac{1}{2} \left\| A_i^T - C_p \right\|_2^2 \right)$$

Which are:

$A_i^T$ : Dimension member I in one cluster

$C_p$ : Indicates the center of each cluster from the cluster 1 and cluster k

After doing the first phase in the grouping, the next step is to create a classification into sub-clusters are smaller. The goal is to get a group that has the best performance ratio (Pambudi, 2016). From the results of sub-clusters which then serve as the optimal portfolio Shaper. In which phase two done based on the value of return, excess return against risk, and risk (Sharpe Ratio). By performing the first stage and second stage of clustering methods, expected stocks that go into the formation of optimal portfolios will have a level of diversification with both and have the best performance ratio. After knowing the results of the grouping optimum stock through the cluster method, next up is the measurement of the performance of the stocks.

**Sharpe Index**

Tandelilin (2010) states that Sharpe index basing his calculations against the concept of the capital market line (capital market line) as peg might guess that is by dividing the risk premium with the portfolio standard deviation during measurement, where the standard the deviation is the total risk. Thus, the Sharpe to measure risk premium resulting from each unit the risks involved (Aldias & Yulianto, 2017). With these calculations, the higher the value, the better the performance measurements that are generated. Sharpe index measurement is formulated as follows:

$$Sj = (Ri - Rf) / σj$$

Which are:

Sj : Sharpe Index

Rj : Average portfolio return j measurement period

Rf : Average return risk-free assets over a long period measurements

σj : Standard deviation j portfolio during the period of measurement

**Treynor Index**

According to Halim (2005), in this method, the performance of the portfolios is measured by means of comparing the risk premium between the risk portfolio are stated with the beta. The calculation of this index is not much different with the Sharpe index, only in this index standard deviation in Sharpe's index was replaced with the beta. Here is the formula:

$$Tj = (Ri - Rf ) / βj$$

which are:
Tj : Treynor Index
Ri : Average portfolio return j measurement period Rf = average return risk-free assets over a long period measurements
βj : Portfolio beta j during the measurement period

**Jensen Index**

This method is based on the concept of the securities market line (the securities market line) which is the line that connects the market portfolio with a risk-free investment opportunity (Halim, 2005). This index shows the difference between the actual return rate with the retrieved portfolio with the level of return expectations if the portfolio is on the capital market line. The formula is:

$$J_p = R_p - [RF + (R_M - RF) β_p]$$

Which are:
$J_p$ : Jensen index
$R_p$ : Average return portfolio p during the period of observation
RF: The average rate of return risk-free during the period of observation
$β_p$ : Beta portfolio p

**RESULTS AND DISCUSSIONS**

This research uses the object of research of stocks listed in the index of LQ 45 in 2012-2016. The object of the research is then given criteria-criteria for selecting samples that will be used in this research. The criteria of a sample to be taken for research include: LQ 45 index listed on, never delisting during the period of research, always publish the financial statements during the period of research, and entered into a period of research. Shares of companies that are used as the sample of the study after the criteria imposed totaling 22 company samples. A list of company names that are used can be seen in Table 1.

**Table 1.** Lists of Research Sample

| Code | Company |
| --- | --- |
| AALI | PT Astra Agro Lestari Tbk |
| ADRO | PT Adaro Energy Tbk |
| ASII | PT Astra International Indonesia Tbk |
| ASRI | PT Alam Sutera Realty Tbk |
| BBCA | PT Bank Central Asia (Persero)Tbk |
| BBNI | PT Bank Negara Indonesia (Persero)Tbk |
| BBRI | PT Bank Rakyat Indonesia (Persero)Tbk |
| BMRI | PT Bank Mandiri (Persero)Tbk |
| CPIN | PT Caroen Pokphand Indonesia Tbk |
| GGRM | PT Gudang Garam Tbk |
| INDF | PT Indofood Sukses Makmur |
| INTP | PT Indocement Tunggal Perkasa Tbk |
| JSMR | PT Jasa Marga (Persero) Tbk |
| KLBF | PT Kalbe Farma Tbk |
| LPKR | PT Lippo Karawaci Tbk |
| LSIP | PT London Sumatera Tbk |
| PGAS | PT Perusahaan Gas Negara (Persero) Tbk |
| PTBA | PT Tambang Batubara Bukit Asam (Persero) Tbk |
| SMGR | PT Semen Indonesia (Persero) Tbk |
| TLKM | PT Telekomunikasi Indonesia (Persero) Tbk |
| UNTR | PT United Tractor Tbk |
| UNVR | PT Unilever Indonesia Tbk |

**Normality Test**

Prior to clustering the data and test the performance of the portfolio, all data must be ensured to be distributed normally. Therefore, the test should be done against data normality ROA, ROE, company size, and variances against return first. The data that is to be tested is converted in the form data after the Z-Score. The purpose of the convert into the value of the Z-score is so that the scale can be likened or grouping is normalized so that the values between variables not be embedded too far. The value of the Z-Score is used only at the time of determining the grouping.

Based on the test results, obtained Kolmogorov Smirnov-value significance of 0.972. This means that the significance of the value data, which provided the sample research has distributed normally. This is due to the value is more than 0.05 significance value.

**Clustering Methods**

The cluster method is done with the help of software R and SPSS 21.0 approach to K-Means clustering. After a test of normality, then the data already entered in the normal distributed software R to determined how many clusters should form.

Based on the results of the output of the NbClust in the software R, the number of clusters should be formed in the method of K-Means clustering was as much as 3 (three) cluster. Thus, the stock data will be grouped in three clusters or in other words the stock data will form three diversified stocks that make up the optimum portfolio diversification.

After it is determined how the values of k are included in the method of k-means clustering, the picture below is the result of the output of k-means clustering with the help of SPSS software: 21.0:

**Table 2.** K-Means Clustering

|  |  | **Members** |
|---|---|---|
| Cluster | 1 | 1 |
|  | 2 | 14 |
|  | 3 | 7 |
| Valid |  | 22 |
| Missing |  | 0 |

Table 2 shows that by the method of k-means clustering there are 3 clusters are formed. Cluster-1 consists of 1 issuer, cluster-2 consists of 14 issuers, cluster-3 consists of 7 issuers.

Based on grouping the data stocks using k-means clustering approach before, then formed three clusters. The result of this grouping is also the establishment of diversification. This is due to the k-means clustering have grouped shares according to the characteristics of each cluster. ROA, ROE, company size, variance, and the return is an indicator of consideration in this grouping.

The data in the tables below are preliminary data before converted in the form of z-score. This is because the value of the Z-score is already in use at the time will form clusters using k-means clustering approach. Therefore, to analyze how the diversification that forms required initial data ROA, ROE, company size, return, and the variance to be seen how the characteristics of diversification that is formed. As for the data that is available in the Table 3.

**Table 3.** Cluster-1 Result

| Code | X1 | X2 | X3 | X4 | X5 |
|---|---|---|---|---|---|
|  | ROA | ROE | Size | Return | Variants |
| UNVR | 45.48 | 125.92 | 13,245,308 | .2219 | .0141 |

Based on the previous analysis have been explained that the table above is a result of the cluster formed by category the value of ROA, ROE, and return the above average. Members of the cluster-1 were only one Member, so in this cluster are not formed diversification. Thus, the result of cluster-1 cannot be used as a combining optimal portfolio.

**Table 4.** Cluster-2 Result

| Code | X1 | X2 | X3 | X4 | X5 |
|---|---|---|---|---|---|
|  | ROA | ROE | Size | Return | Variants |
| AALI | 11.82 | 17.11 | 18,336,420 | -.0445 | .0346 |
| ADRO | 3.94 | 7.81 | 79,385,374 | -.1021 | .0779 |
| ASRI | 6.23 | 16.04 | 16,196,619 | .0750 | .0999 |
| CPIN | 12.70 | 21.17 | 19,768,408 | .1351 | .0672 |
| GGRM | 9.03 | 15.12 | 55,364,573 | .0750 | .0365 |
| INTP | 17.32 | 19.74 | 27,207,263 | .0349 | .0212 |
| JSMR | 4.26 | 11.88 | 35,040,630 | .0841 | .0430 |
| KLBF | 16.03 | 20.47 | 12,308,594 | .1988 | .0486 |
| LPKR | 4.77 | 10.30 | 36,172,224 | .1089 | .0562 |
| LSIP | 9.66 | 11.66 | 8,497,940 | -.0528 | .0413 |
| PGAS | 13.33 | 23.98 | 71,129,044 | -.0204 | .0656 |
| PTBA | 14.10 | 22.99 | 14,672,566 | -.1438 | .0162 |
| SMGR | 14.85 | 21.05 | 34,813,330 | .0359 | .0548 |
| UNTR | 8.06 | 12.58 | 58,732,307 | -.0492 | .0169 |

Table 4 is a result of the cluster formed by the average value of the ROA, ROE, company size, and return a negative or under on average, but the value of risk that is above average. Category optimal portfolio selection is minimal risk and maximum return. Results on cluster-2 are inversely proportional to the optimal portfolio of the common stock category. Thus, the results of the cluster-2 is not a share which categorized the optimal portfolio formation.

**Table 5.** Cluster-3 Result

| Code | X1 | X2 | X3 | X4 | X5 |
|------|------|------|-----------|--------|----------|
|      | ROA  | ROE  | Size      | Return | Variance |
| ASII | 9.12 | 18.95 | 227,917,400 | .04 | .0067 |
| BBCA | 2.91 | 20.89 | 552,566,837 | .14 | .0061 |
| BBNI | 2.14 | 15.47 | 449,631,839 | .07 | .0088 |
| BBRI | 3.06 | 24.17 | 772,309,095 | .13 | .0087 |
| BMRI | 2.24 | 17.82 | 834,505,512 | .09 | .0039 |
| INDF | 5.37 | 10.62 | 81,523,447 | .08 | .0226 |

Table 5 is a result of the cluster formed by the average value of the return and the size of the companies that are above average. Stocks that are formed from the results of the cluster, this is the best category in between the other two. This is because on the previous explanation that the cluster member has the value of return above average and risks under the average. Therefore, the members who are present in the cluster-3 best stocks is the optimal portfolio formation.

Weighting the proportion of shares the results of the cluster is performed using a single index of the model formulation. This is because the cluster method is just an algorithm, not the gauge inferential statistics. Thus, to measure weighting can be done with the absorption of a single index model formulation. The calculation of the weighting, then the formulation can be shown in the following table:

**Table 6.** The weighting of Optimal Portfolio

| Kode | Bobot |
|------|-------|
| ASII | 3.96% |
| BBCA | 21.39% |
| BBNI | 5.84% |
| BBRI | 21.50% |
| BMRI | 30.32% |
| INDF | 3.97% |
| TLKM | 13.03% |
| Total | 100% |

Based on the table above, it can be noted that the proportion of shares of the highest is low and most BMRI was INDF. Method of cluster successfully formed the shares forming the optimal portfolio with a number of issuers as 6 corporations.

After knowing the grouping shares based on the method of the cluster, the next step is to analyze the performance of a portfolio that included in the optimal portfolio. As for the projections to be used is the Sharpe index Jensen, index, and index Treynor.

**Table 6.** Portfolio Performance

| Kluster | Indeks Jensen | Indeks Sharpe | Indeks Treynor |
|---------|---------------|---------------|----------------|
| 2 | .1901 | .1509 | .0146 |
| 3 | .1856 | 1.1452 | .1251 |

Based on the results of the calculation of an index of Jensen, Sharpe, Treynor and from the table above, the stocks are selected to form the optimal portfolio are stocks that are included in cluster 3. In this case which is shares in cluster 3 potentially can form an optimal portfolio. In addition, the results of the previous cluster output the result in accordance with the results of calculations using index Jensen, Sharpe, and Treynor.

The use of the method for the formation of an optimum portfolio cluster is done using multiple stages. The first stage by calculating the entire value of the indicators used to establish an optimal portfolio. After calculating the indicator, then convert the values into the form z-score. The next step is to enter data into the software z-score R to determine how the number of clusters (amount k on approach to k-means) is the best. Thus, this led to the k-means clustering approach can classify data in accordance with its characteristics form the cluster 3.

Cluster-cluster 1, 2, and 3 were formed based on the cluster with the characteristics of each. Cluster-1 consists of the UNVR, where the member is based on the value of the ROA, ROE, and return a far above total average. This is evidenced by the positive value of the ROA, ROE, and the return on the results of the k-means cluster, which indicates that the value is the value above average. While the value of the company's size and variance (risk) get value negative. Because cluster-1 only consists of one company, so there can be seen the diversification, and in this cluster are not allowed to be analyzed further.

Cluster-2 has 14 members. Based on the results of k-means clustering, the cluster was found

to have the characteristic value of the ROA, ROE, company size, and return below the average. While the risk of this cluster has a value above average so that within the cluster is certainly not in line and reject the theory of optimal portfolio that was introduced by Markowitz (1952). This is because in the cluster-2 return value is smaller than the value of the risks involved.

Cluster-3 has 6 members. Each Member has the characteristic value of the ROA, ROE, and the risks that under average. While the value of the return and the size of the company is above average. The results of this cluster support the optimal portfolio theory introduced by Markowitz (1952). This is because in the cluster-3, its members have a value higher than the return value of the risk, so it is in the cluster-3 can be analyzed more about optimal portfolio proportion with the performance of its portfolio.

After calculation, the result of portfolio performance clusters with k-means clustering approaches has in common the result with index Jensen, Sharpe, and Treynor. It is shown that the cluster-3 has a better index value of cluster-2. Thus, the conclusion can be drawn that the cluster-cluster 3 is best to form the optimal portfolio with also saw the diversification of its shares.

The results of this study support previous studies. Research Babu et al. (2012) stated that the method of k-means clustering could divide multiple groups with the same characteristics of the stock so that it will maximize the yield and minimizing the risks in investing. The study also supports the research of Nanda et al. (2010) stating that the method of k-means clustering has better performance than the Fuzzy cluster method and SOM. This statement is answered in this research, due to the approach of the k-means clustering in this study were able to select stocks that have a risk rating is lower than the return and vice versa, are able to classify stocks that can provide a greater advantage of the level of risk.

This research can provide an overview to the investor that the cluster methods can be grouped based on stocks with a value of minimum risk and maximum return. The study also supports the statement of the Tandelilin (2010) stating that the model Markowitz would be hard is used to analyze the stocks amounted to much. This is because the cluster model results are more easily understood and can form a diversified stock with more simple. In addition, with the formation of diversification which can be seen from each Member of the cluster, the results of this research also supports research Pambudi (2016) stating that a single index model made by some researchers previously do not look at diversification so that the cluster model can be used as an alternative to forming the optimal portfolio.

Portfolio performance measurement index measured from Jensen, Sharpe, Treynor and produced that cluster-3 has better performance than the cluster-2. So the portfolio that will produce the maximum return and minimal risk are located on the cluster member shares-2. As for the proportion of each of its shares: ASII (3.96%), BBCA (21.39%), BBNI (5.84%), BBRI (21.50%), the BMRI (30.32%), INDF (3.97%), and (TLKM 13.03%).

## CONCLUSION AND RECOMMENDATION

Based on the results of such research could be taken to the conclusion that the cluster can be used as a method of forming the optimal portfolio based on characteristic value variables that serve as indicators of measurement against the return of a company. K-means clustering method used in this research cluster 3 best tailored to each company's characters in common. This research supports research conducted by Nanda et al. (2010); Babu et al. (2012) concerning the use of the method for the formation of an optimum portfolio cluster. Nanda et al. (2010) States that the k-means clustering method is the best method if compared with the method of fuzzy c-means and SOM. While Babu et al. (2012) stated that the use of k-means clustering method in his research because of this approach is the approach most often used in analyzing a grouping based on similarity to object/ characteristic. For investors, the cluster method with k-means clustering approach can be used to create an optimal portfolio diversification. Further research is expected to make the period of study is longer and use objects as well as indicators of research that different.

## REFERENCES

Adiningrum, T. R., Hidayat, R. R., & Sulasmiyati, S. (2016). Penggunaan Metode Single Index Model dalam Menentukan Portofolio Optimal Tahun 2012-2015 (Studi pada Saham-Saham yang Terdaftar dalam Indeks IDX30 di Bursa Efek Indonesia Periode Februari 2012 - Agustus 2015). *Jurnal Administrasi Bisnis (JAB)*, *38*(2), 89-96.

Aldias, W. A., & Yulianto, A. (2017). Analisis Kinerja Reksadana Saham dengan Metode Sharpe Treynor dan Jensen. *Management Analysis Journal*, *6*(1), 13-22.

Asmi, T. L. (2014). Current Ratio, DER, TAT, ROA, PBV, sebagai Penentu Return Saham. *Manage-*

*ment Analysis Journal, 3*(2), 1-12.

Babu, M. S., Geethanjali, N., & Satyanarayana, P. B. (2012). Clustering Approach to Stock Market Prediction. *International Journal Advanced Networking and Applications, 1291*, 1281-1291.

Brigham, E. F., & Houston, J. F. (2011). Dasar-Dasar Manajemen Keuangan. Edisi 10. Jakarta: Salemba Empat.

Chandra, L., & Hapsari, Y. D. (2014). Analisis Pembentukan Portofolio Optimal dengan Menggunakan Model Markowitz untuk Saham LQ45 Periode 2008-2012. *Jurnal Manajemen, 11*(1), 41-59.

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: an R Package for Determining the. *Journal of Statistical Software, 61*(6), 1-5.

Eko, U. (2010). Analisis dan Penilaian Kinerja Portofolio Optimal Saham-Saham LQ-45. *Jurnal Ilmu Administrasi dan Organisasi, 15*, 178-187.

Elton, E. J., & Gruber, M. J. (1971). Valuation and the Cost of Capital for Regulated Industries. *The Journal of Finance, 26*(3), 661-670.

Fan, Y. (2010). Optimization Models and Algorithms or Sample-Preserved Classification and Clustering. *Dissertation.* New Brunswick, New Jersey.

Guam, H. S., & Jiang, Q. S. (2007). Cluster Financial Time Series for Portfolio. In *Wavelet Analysis and Pattern Recognition, 2007. ICWAPR'07. International Conference, 2*, 851-856.

Fabozzi, F. J. (2006). *Fixed Income Mathematics, Analytical and Statistical Techniques.* 4th Edition. New York: McGraw-Hill.

Ghozali, I. (2006). *Aplikasi Analisis Multivariate dengan Program SPSS.* Semarang: Badan Penerbit Universitas Diponegoro.

Gudono. (2014). *Analisis Data Multivariat.* Yogyakarta: BPFE.

Gulo, W. (2002). *Metode Penelitian.* Jakarta: PT. Grasindo.

Hermawan, D. A. (2012). Pengaruh Debt to Equity Ratio, Earning per Share dan Net Profit Margin terhadap Return Saham. *Management Analysis Journal, 1*(5), 1-7.

Jayati, N., Handayani, S. R., & Zahro Z.A, Z. (2017). Analisis Metode Single Index Model dalam Pembentukan Portofolio Optimal Untuk Menurunkan Risiko Investasi (Studi pada Perusahaan yang Terdaftar dalam Indeks IDX30 Periode Agustus 2013-Juli 2016). *Jurnal Administrasi Bisnis (JAB), 49*(1), 96-105.

Kasmir. (2014). *Analisis Laporan Keuangan.* Edisi Pertama, Cetakan Ketujuh. Jakarta: PT. Rajagrafindo Persada.

Khajar, I. (2010). Pengaruh Right Issue terhadap Kinerja Keuangan Perusahaan. *Jurnal Dinamika Manajemen, 1*(1), 27-33.

Lahmiri, S. (2016). Clustering of Casablanca Stock Market Based on Hurst Exponent Estimates. *Physica A: Statistical Mechanics and its Applications, 456,* 310-318.

Maftukhah, I. (2013). Kepemilikan Manajerial, Kepe-

milikan Institusional dan Kinerja Keuangan sebagai Penentu Struktur Modal Perusahaan. *Jurnal Dinamika Manajemen, 4*(1), 69-81.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance, 7*(1), 77-91.

Mirah, & Wijaya, T. (2013). Analisis Model Indeks Tunggal Portofolio Saham di Bursa Efek Indonesia (BEI) Periode 2009-2011. *Journal MDP,* 1-10.

Mohamad, I. Bin, & Usman, D. (2013). Standardization and Its Effects on K -Means Clustering Algorithm, *6*(17), 3299-3303.

Nanda, S. R., Mahanty, B., & Tiwari, M. K. (2010). Clustering Indian Stock Market Data for Portfolio Management. *Expert Systems with Applications, 37*(12), 8793-8798.

Pambudi, B. S. (2016). Pendekatan Metoda Klaster terhadap Model Indeks Tunggal dalam Pembentukan Portofolio Optimal. *Doctoral Dissertation.* Universitas Gadjah Mada.

Pardosi, B., & Wijayanto, A. (2015). Analisis Perbedaan Return dan Risiko Saham Portofolio Optimal dengan Bukan Portofolio Optimal. *Management Analysis Journal, 4*(1), 1-9.

Merliana, N. P. E., & Santoso, A. J. (2015). Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means Clustering. *Proceeding Sendi_U. Prosiding Seminar Nasional Multidisiplin Ilmu dan Call for Paper Unisbank*, 978-979.

Subekti, R., Kusumawati, R., & Sari, E. R. K-Means Clustering dan Average Linkage dalam Pembentukan Portfolio Saham. *Seminar Matematika dan Pendidikan Matematika,* 219-223.

Sudiyanto, B., & Suharmanto, T. (2011). Kinerja Keuangan Konvensional, EVA dan Return Saham. *Jurnal Dinamika Manajemen, 2*(2), 153-161.

Susilowati, M., Rahmawati, R., & Prahutama, A. (2016). Analisis Kinerja Portofolio Optimal dengan Metode Mean-Gini. *Jurnal Gaussian, 5,* 498-504.

Tandelilin, E. (2010). *Portofolio dan Investasi Teori dan Aplikasi.* Yogyakarta: Kanisius.

Tola, V., Lillo, F., Gallegati, M., & Mantegna, R. N. (2008). Cluster Analysis for Portofolio Optimization. *Journal of Economic Dynamics and Control, 32*(1), 235-258.

Trihastuti, K. (2012). Analisis Kinerja Perusahaan dengan Metode BSC. *Management Analysis Journal, 1*(1), 1-9.

Van Horne, J. C., & Wachowicz, J. M. (2005). *Prinsip-prinsip Manajemen Keuangan.* Jakarta: Salemba Empat.

Wahyudin, A., & Solikhah, B. (2017). Corporate Governance Implementation Rating in Indonesia and Its Effects on Financial Performance. *Corporate Governance: The International Journal of Business in Society, 17*(2), 250-265.

Wati, N. S., & Sulasmiyati, S. (2016). Analisis Single Index Model untuk Menentukan Komposisi Portofolio Optimal (Studi pada Saham yang Termasuk 50 Leading Companies in Market

Capitalization Periode 2012-2015). *Jurnal Administrasi Bisnis, 35*(1), 43-52.

Wijayanto, A. (2010). Analisis Pengaruh ROA, EPS, Financial Leverage, Proceed terhadap Initial Return. *Jurnal Dinamika Manajemen*, *1*(1), 68-78.

Witiastuti, R. S. (2012). Analisis Kinerja Potofolio Pengujian Single Index Model dan Naive Diversification. *Jurnal Dinamika Manajemen*, *3*(2), 122-132.

Yulianto, A., Suseno, D. A., & Widiyanto, W. (2016). Testing Pecking Order Theory and Trade Off Theory Models in Public Companies in Indonesia. *International Journal of Economic Perspectives*, *10*(4), 21-28.

Zhang, J., & Maringer, D. (2010). A Clustering Application in Portfolio Management. *Electronic Engineering and Computing Technology*, *27*, 390-399.