p-ISSN : 2774-9576
e-ISSN : 2747-2981

PCA Implementation in Identifying Risk and Return of LQ45 Stocks

Irawan Wingdes[1], Sahrul Nurfaizi[2], Muhamad Rifki[3]
Information Systems, STMIK Pontianak[1,2,3]
Jln Merdeka no 372, West Borneo /0561-735555 [1,2,3]
Indonesia [1,2,3]
e-mail: irawan.wingdes@gmail.com[1]

To cite this document:
Wingdes, I., Nurfaizi, S. ., & Rifki, M. . (2022). PCA Implementation In Identifying Risk And Return of LQ45 Stocks. Conference Series, 4(1), 32–42. https://doi.org/10.34306/conferenceseries.v4i1.668

Hash:ABCLneWpMmeh175i2ymnhqcx14Q85So35w6pCNCFeg3U5sr0VrXjseXGbk56dMOA

Abstract

Stock investing is a high-risk investment, so investors are constantly looking for ways to reduce risk while adapting returns. Lower risk stock groups are represented in Indonesia by the LQ45 index, which includes stocks with high liquidity, large capitalization, and good fundamentals. However, the index contains 45 members with varying risk and return. As a result, the price action analysis of the index's stock members is required to further reduce risk. This study employed principal component analysis to identify the index's various price action groups. PCA was used in Python 3 and produced two groups that explain 70.91% of the variance in the index's price movement. The stock groups derived from PCA were further analyzed, and ANOVA revealed that the groups differed significantly in return but not significantly in risk. The implemented PCA demonstrated potential return for a variety of investors interested in LQ45 stocks.
Keywords: LQ45 Index, PCA, Price Action

p-ISSN : 2774-9576
e-ISSN : 2747-2981

I. INTRODUCTION

Stock investing is a high-risk activity, and as such, caution and informed decision-making are required. Individual stocks that can be purchased independently by Indonesian investors carry the risk of a stock price decline, corporation liquidity, or the inability of the stock to be transferred [1]. There is also a danger of diminishing investment value over time, liquidity risk, or, in the worst-case situation, a default in mutual funds managed by professional investment managers [2].

In 1997, the Indonesian stock exchange launched the LQ45 index[3], which consists of lower-risk companies associated with higher-performing corporations and better liquidity stocks. Investors will gain a better understanding of which equities are less risky as a result of this index. However, the LQ45 index is comprised of 45 companies with varying risk and return characteristics. As a result, additional research on the risk and return characteristics of the 45 members of the LQ45 index is necessary for investors to make better informed decisions.

One of the most often used analytical tools in stock trading is price movement analysis, which is colloquially referred to as price action trading [4]. Each company in the LQ45 index exhibits an unique price behavior. Because the latter is difficult to see, a tool for simplifying price action movement will aid investors in gaining a better understanding of the latter. Simplifying can be accomplished in a variety of ways, from forecasting with several specific algorithm[5] to recognizing price action similarities by identifying patterns[6] and finally forecasting by pre-classifying stocks with multiple variables[7].

The analytical tools were frequently investigated in relation to stock indices, especially in developed countries. In emerging countries such as Indonesia, it was less common to investigate an index's price movement from its price action. Earlier research concentrated on machine learning historical prices in order to obtain a better understanding of expected future prices[8], forecasting prices using a variety of methodologies[5], and eventually trying to reduce indicators utilized in stock investment [9]. This research will try to narrow the gap by discovering stocks in the LQ45 index with similar short-term price movements. This research was unique in that it examines price movement without attempting to forecast future price direction. As such, the objective of this research was to find groups of stocks that exhibit similar price movements in order to extract additional information from risk and return of such group. By identifying companies according to their price movement, we intended to profile stocks risk and return in the LQ45 index based on its price movement. Identifying stocks risk and return with distinct price movement characteristics enables investors to make more informed decisions, since there will be contrasts from which they can choose.

Principal component analysis was used to accomplish the research purpose. This technique was employed because it maximizes the variation extracted from price movement while simultaneously decreasing the dimensions of the multiple stocks price movements. In order to analyze the data, the time series approach with purposive sampling was used. The data was derived from the daily closing prices of LQ45 index members in 2021.

The year 2021 was chosen because it corresponded to the most recent price change, which was more relevant for short-term trading. Additionally, 2021 presented a more stable market following the pandemic havoc of 2020. After performing PCA, the analysis's results can be studied in better detail. Reduced price groups will then be determined further for their daily rate of return and associated daily risk. The groups created from price movement were then compared using anova to determine their statistical significance difference.

II. LITERATURE REVIEW

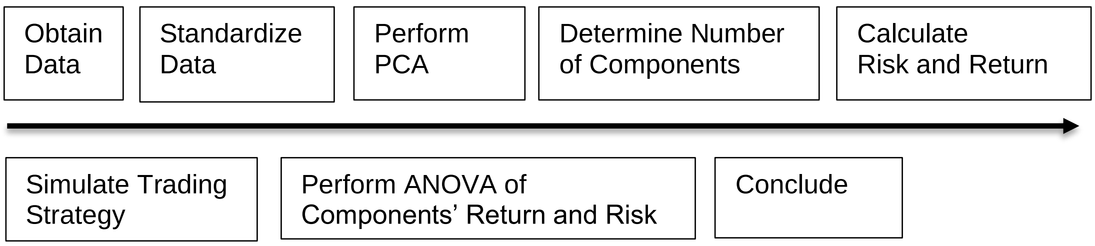
Principal components analysis is a technique for reducing the dimension of large datasets in an interpretable direction. This encompasses retaining as much variation as possible. The dimension is reduced by identifying new variables that are linear functions of the data in such a way that the new variables are uncorrelated with each other [10][11]. PCA has been used in a variety of studies, including [5][7][9][12]. In [5], PCA was used as an input to an LSTM model to predict the direction of a single stock, which was then compared to other dimension reduction techniques such as lasso and elastic net. According to the research, PCA was more accurate than the other techniques used. In [7], traditional simple PCA was compared to Fuzzy Robust PCA and Kernel-based PCA as an input to an ANN; again, the results indicated that traditional PCA was more accurate at predicting the daily direction of the stock market return. According to [9], PCA can be used to reduce macroeconomic indicators, which can then be used to attempt to forecast stock prices using regression. PCA was used in [12] to group sentiment metrics that have an effect on stock return.

Prior research in Indonesia [5][9][12] indicated that, while still scarce, PCA has been implemented successfully. Although none of the studies identified provide additional information about the use of pca to identify price movement groups. Thus, this research will help bridge the gap further by utilizing PCA to determine the similarity of price movements in addition to attempting to anticipate dynamic stock market movement.

III. RESEARCH METHODS

The steps taken during this research can be summarized as follows (Figure 1):

Figure 1. Research Steps



To obtain the necessary data, the most recent index LQ45 members were obtained from [13], and tickers / stock data were obtained from the yahoo finance website for the 2021 daily trading prices [14]. The data were obtained in Python 3 [15] with the numpy and pandas libraries and then standardized to the z-score using the following pseudo code block:

```
Dataset = Get_tickers[
'ACES.JK', 'ADRO.JK', 'AKRA.JK', 'ANTM.JK', 'ASII.JK',
'BBCA.JK', 'BBNI.JK', 'BBRI.JK', 'BBTN.JK', 'BMRI.JK',
'BRPT.JK', 'BSDE.JK', 'CPIN.JK', 'ERAA.JK', 'EXCL.JK',
'GGRM.JK', 'HMSP.JK', 'ICBP.JK', 'INCO.JK', 'INDF.JK',
'INKP.JK', 'INTP.JK', 'ITMG.JK', 'JPFA.JK', 'JSMR.JK',
'KLBF.JK', 'MDKA.JK', 'MEDC.JK', 'MIKA.JK', 'MNCN.JK',
'PGAS.JK', 'PTBA.JK', 'PTPP.JK', 'PWON.JK', 'SMGR.JK',
'SMRA.JK', 'TBIG.JK', 'TINS.JK', 'TKIM.JK', 'TLKM.JK',
'TOWR.JK', 'TPIA.JK', 'UNTR.JK', 'UNVR.JK', 'WIKA.JK']

Z_score(dataset)
```

Standardization was used to center the data in light of the varying stock prices. Data centering was a necessary procedure for performing PCA. The following formula was used to calculate the z score:

$$z = \frac{x - \bar{x}}{s} \tag{1}$$

where x represents the observed data, x bar represents the mean of the data for each stock, and s represents the standard deviation of the observed data for each stock.

 p-ISSN : 2774-9576
e-ISSN : 2747-2981

PCA was implemented firstly by identifying data. There are p variables (45 stocks of LQ45 members) associated with the data in this study, each with n observations (trading days). The data matrix will be linearly projected in a way that maximizes variance. The projection of data revealed the importance of the covariance matrix and ultimately, solving for eigenvalues with corresponding eigenvectors managed to solve the maximization problem. Formula for PCA is given below:

$$\max \ var(\alpha_1^T \mathrm{x}) = \alpha_1^T \Sigma \alpha_1 \qquad (2)$$
$$s.t. \quad \alpha^T \alpha = 1$$

To maximize $\alpha_1^T \Sigma \alpha_1$ subject to $\alpha^T \alpha = 1$, the standard approach is to use Langrange multipliers. Maximize

$$\alpha_1^T \Sigma \alpha_1 - \lambda(\alpha_1^T \alpha_1 - 1) \qquad (3)$$

where λ is a Lagrange multiplier. Differentiation with respect to α1 gives

$$\left(\Sigma - \lambda \boldsymbol{I}_p\right)\alpha_1 = 0 \qquad (4)$$

Where the quantity to be maximized is

$$\alpha_1^T \Sigma \alpha_1 = \alpha_1^T \lambda \, \alpha_1 = \lambda \alpha_1^T \alpha_1 = \lambda \qquad (5)$$

so λ must be as large as possible with the corresponding eigenvector. The second component follows the similar solving steps where the second component equals to second largest eigenvalue. PCA was implemented in Python 3 firstly by performing eigen decomposition using the following pseudo code with numpy.

```
Eigen_value, Eigen_vectors = numpy.linalg(covariance_matrix of data)
```

In order to reduce the dimension for interpretation and group similar price movements, the 45 eigenvalues were sorted and then their percentage of total eigenvalues was calculated. The following was the formula:

$$\pi_j = \frac{\lambda_j}{\sum_{j=1}^{p} \lambda_j} * 100\% \qquad (6)$$

To determine the number of components, a scree plot of the extracted eigenvalues and variance was plotted, along with the proportion of eigenvalues relative to the summed eigen values. A cutoff value of 70% is frequently used [11], as is a steepest eigenvalue in screeplot. It was implemented in Python and utilizes the following pseudo code.

```
Sorted_eigenvalues = Eigenvalues.sort
Biggest_variance1 = each_sorted_eigenvalues / summed_eigenvalues
print_out(components, total_variance)
```

To aid in interpretation, the varimax method was used to create a rotation from the components selected. Rotation has no effect on variance but greatly enhances interpretation because loadings after rotation were similar to correlation coefficients . By determining which variable was associated with which component, the interpretation of reduced components can be more precisely examined [16].

Following rotation and variable identification, the formula below was used to calculate the rate of return on the corresponding stocks for each component.

$$Daily\ Rate\ of\ Return = \frac{Open - Close}{Open} \qquad (7)$$

The following formula computed a risk indicator, which was the standard deviation of the daily return.

$$Standard\ Deviation = \sqrt{\frac{\Sigma(x_i - \overline{x})^2}{n}} \tag{8}$$

Risk is intuitively equal to the standard deviation of the rate of return, as the standard deviation reflects the percentage change in the expected rate of return. The risk is the change due to the element of uncertainty [17].

Following the computation of RoR and risk, the ror and risk of each new component group were further compared using anova. Anova was used to test the following hypothesis with a confidence level of 95%.

H_0: there exist no difference of RoR and Risk between newly created components
H_{a1} : there is significant difference of RoR from newly created components
H_{a2} : there is significant difference of risk from newly created components

Following anova, each stock belonging to newly created components was further compared using a buy and hold strategy to simulate real trading. Purchase on the first day and sell on the final day. The results were then further analyzed to meet the particular needs of each investor.

IV. FINDINGS AND RESULTS

The preceding method steps was implemented, with the first step being to collect data and preprocessing by standardizing it, yielded the following result:

Table 1. Data Obtained From Yahoo (Partial)

ACES.JK	Date	Open	High	Low	Close	Adj Close	Volume
	2021-01-04	1715	1745	1670	1700	1660.820801	12191600
	2021-01-05	1700	1760	1675	1760	1719.437988	19742400
	2021-01-06	1760	1780	1690	1730	1690.129395	14464700
ADRO.JK	Date	Open	High	Low	Close	Adj Close	Volume
	2021-01-04	1430	1460	1360	1455	1377.850098	110366200
	2021-01-05	1455	1470	1420	1425	1349.440796	107023500
	2021-01-06	1415	1420	1340	1375	1302.092041	203948800
BBRI.JK	Date	Open	High	Low	Close	Adj Close	Volume
	2021-01-04	4150	4320	4150	4310	4217.732422	96568200
	2021-01-05	4300	4300	4240	4270	4178.588867	97239100
	2021-01-06	4280	4300	4160	4200	4110.087402	116634000

Standardizing the data (daily closing price) resulted in:

Table 2. Data Standardization (Partial)

ACES	ADRO	AKRA	ANTM	ASII	BBCA	BBNI	BBRI	BBTN
1.9309	0.1401	-0.7403	-1.2231	1.3747	0.2256	0.6664	0.4077	0.8654
2.4023	0.0302	-0.8266	-1.025	1.2729	0.7551	0.839	0.2891	0.8176
2.1666	-0.153	-0.9344	-1.0745	1.222	0.454	0.7009	0.0817	0.7938
2.5594	-0.0431	-0.4383	0.8076	1.0694	0.4955	0.7354	0.3188	0.8176

Following the preprocessing of obtaining and standardization of daily closing prices for each stock in LQ45. PCA was carried out by first generating a covariance matrix. The output was as follows:

Figure 1. Covariance Matrix From Data (Partial)

p-ISSN : 2774-9576
e-ISSN : 2747-2981

```
          BBRI      BBTN      BMRI  ...      SMRA      TBIG      TINS  \
ACES  0.644616  0.498964  0.055365  ... -0.094093 -0.819265  0.681490
ADRO -0.076245  0.147507  0.685127  ...  0.175647  0.329344 -0.278533
AKRA -0.171526  0.113741  0.562129  ...  0.285958  0.372427 -0.283497
ANTM  0.415673  0.277244  0.043546  ... -0.249949 -0.387697  0.833111
ASII  0.614132  0.699644  0.852215  ...  0.140401 -0.482469  0.474145
BBCA  0.365453  0.589861  0.899109  ...  0.244900 -0.095505  0.156616
BBNI  0.556993  0.759259  0.923824  ...  0.411655 -0.296620  0.234362
BBRI  1.004082  0.877183  0.520459  ...  0.211185 -0.750869  0.723811
BBTN  0.877183  1.004082  0.658812  ...  0.366548 -0.695652  0.605871
```

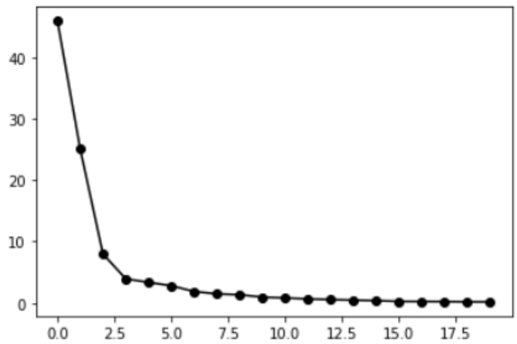
The covariance matrix was symmetric with the diagonal contained variances. From the covariance matrix, eigenvalue and eigenvector was calculated, sorted and compared to total summed eigenvalues with following output of the first 8 eigenvalues:

Table 3. Sorted Eigenvalues and Total Variance

No.	EigenValues	% Variance
1	45.877	45.88%
2	25.026	25.03%
3	7.958	7.96%
4	3.931	3.93%
5	3.349	3.35%
6	2.809	2.81%
7	1.866	1.87%
8	1.481	1.48%

According to the eigenvalues in table 3, the first two components accounted for 70.91% of the variance in the price movement of LQ45 index stocks. The variance of the eigenvalues decreased as more components are extracted. To verify the eigenvalues further, a screeplot or eigenspectra was created, as shown in figure 2.

Figure 2. Eigen Spectra / Scree Plot



The purpose of the scree plot was to visually identify which component contributes the most variance by observing the plot's steepest effect. Observation revealed that the steepest was produced by the first two components. The total variance and screeplot indicated that two components account for 70% of the price movement in the LQ45 index. Following the purpose of this pca, which was differentiated from trying to fit data in order to predict, 70% variance was deemed sufficient to represent the total variance of price movement.

After selecting the number of components, the following table displays the stock loadings on the components, along with the result of the varimax rotation utilised. Both results were expressed in absolute loading values (eigenvector times square root of eigenvalue) to avoid additional confusion and to improve readability. Loadings can be thought of as the correlation coefficient of each stock with respect to its group.

Table 4. Component Loading and Rotation Result (Partial)

Before Rotation			Varimax Rotation		
	comp.1	comp.2		comp.1	comp.2
ACES	0.912	0.109	ACES	0.776	0.433
BBCA	0.414	0.718	BBCA	0.435	0.862
ERAA	0.291	0.165	ERAA	0.674	0.252
EXCL	0.368	0.867	EXCL	0.272	0.848
GGRM	0.427	0.846	GGRM	0.456	0.546
ICBP	0.551	0.133	ICBP	0.761	0.267
INCO	0.560	0.440	INCO	0.579	0.160
INDF	0.579	0.497	INDF	0.319	0.101
INKP	0.583	0.552	INKP	0.912	0.287
INTP	0.584	0.654	INTP	0.909	0.174
ITMG	0.599	0.037	ITMG	0.283	0.886
JPFA	0.603	0.698	JPFA	0.316	0.232
JSMR	0.610	0.604	JSMR	0.875	0.116
KLBF	0.637	0.344	KLBF	0.685	0.419
MDKA	0.650	0.685	MDKA	0.256	0.788
TPIA	0.926	0.018	TPIA	0.436	0.761
UNTR	0.952	0.092	UNTR	0.669	0.366
UNVR	0.972	0.111	UNVR	0.828	0.518
WIKA	0.978	0.034	WIKA	0.964	0.169

By examining table 4, it was clear that the rotated component results in greater differentiation than the unrotated result. For example, prior to rotation, it was unclear whether INKP belonged to group 1 or group 2, but after rotation, it was clear that it belonged to group 1, with a correlation of 0.9 to group 1 versus 0,287 to group 2. INTP loaded 0.582 to group 1 and 0.654 to group 2 prior to rotation and 0.909 to group 1 and 0.174 to group 2 following rotation (distinct). MDKA before rotation loaded 0.650 and 0.685 for group 1 and 2, after rotation loaded 0.256 and 0.788 for group 1 and 2. ERAA before rotation loaded 0.291 and 0.165, after rotation loaded 0.674 and 0.252.

Additionally, we can see from the rotated components that some variables were in ambiguous groups and that they correlated almost equally with both components (GGRM). Thus, to further divide the group in order to avoid the issue, a cut-off value is chosen arbitrarily (0,6). Variables with rotated loadings less than 0.6 were omitted. And the following result was obtained:

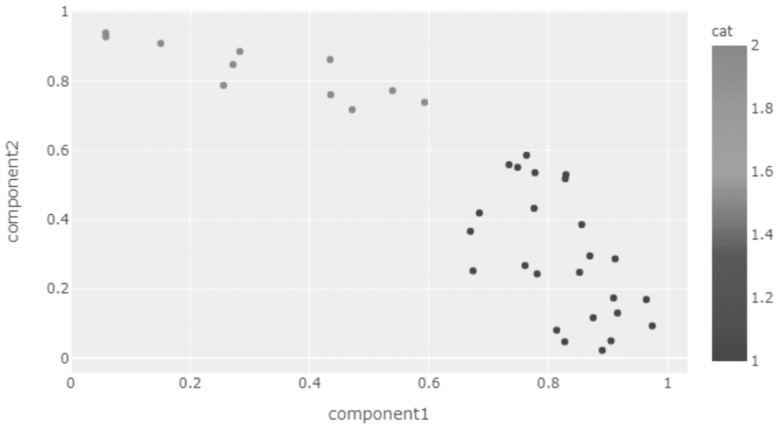
Table 5. Final Components

	comp.1	comp.2		comp1	comp2		comp.1	comp.2
PTPP	0.974	0.093	UNVR	0.828	0.518	ADRO	0.059	0.939
WIKA	0.964	0.169	TOWR	0.828	0.047	TLKM	0.059	0.928
MNCN	0.916	0.130	PWON	0.814	0.080	AKRA	0.151	0.909
INKP	0.912	0.287	TINS	0.781	0.243	ITMG	0.283	0.886
INTP	0.909	0.174	ASII	0.778	0.536	BBCA	0.435	0.862
BSDE	0.905	0.050	ACES	0.776	0.433	EXCL	0.272	0.848
BBRI	0.890	0.022	HMSP	0.764	0.586	MDKA	0.256	0.788
JSMR	0.875	0.116	ICBP	0.761	0.267	BMRI	0.539	0.773
TBIG	0.870	0.295	SMGR	0.749	0.551	TPIA	0.436	0.761
TKIM	0.856	0.386	PTBA	0.734	0.559	BBNI	0.593	0.739
BBTN	0.852	0.248	KLBF	0.685	0.419	MIKA	0.472	0.718
PGAS	0.830	0.530	ERAA	0.674	0.252			

UNTR	0.669	0.366

We can see from the new group that ambiguity has been reduced further and that both components now load satisfactorily in each group. From the identified group, a plot of both components was created to visually inspect whether groups were clearly distinct from one another. Given the fact that the plot indicated that the components do not overlap, suffice it to say that the components were distinct.

Figure 3. Plot of Components



Following the identification of components, the daily rate of return was calculated. Additionally, we used a trading simulation in which we buy at the end of the first trading day of 2021 and sell at the end of the last trading day. The table showed the daily return on each stock, as well as the return on the buy and hold strategy.

Table 6. Simulated Trading Return of Component 1 (in percentage)

Component 1	BnH total return	BnH average daily return	BnH daily stdev	Daily RoR	Daily stdev
ACES	-25.59	-14.52	7.46	0.33	1.96
ASII	-7.63	-10.89	7.89	0.32	1.83
BBRI	-5.34	-3.20	7.86	0.32	1.80
BBTN	-3.85	-9.99	11.53	0.20	2.20
BSDE	-19.76	-13.69	8.74	0.34	2.10
ERAA	25.00	22.59	10.17	0.17	3.00
HMSP	-35.97	-21.60	11.44	0.35	1.70
ICBP	-9.40	-9.36	4.18	0.21	1.41
INKP	-25.75	-14.04	20.97	0.39	3.19
INTP	-19.83	-17.72	9.61	0.27	2.13
JSMR	-16.56	-12.32	6.45	0.24	1.95
KLBF	9.49	0.67	7.03	0.16	2.31
MNCN	-23.28	-17.52	9.12	0.35	2.16
PGAS	-9.42	-15.76	13.81	0.32	2.47
PTBA	0.00	-10.97	9.22	0.35	2.02
PTPP	-46.68	-34.08	10.26	0.47	3.22
PWON	-8.46	-4.04	8.47	0.25	2.25
SMGR	-42.54	-23.53	10.26	0.47	2.40
TBIG	73.76	58.07	26.35	-0.13	3.02
TINS	-8.67	3.43	13.96	0.30	3.87
TKIM	-27.63	-10.05	25.15	0.32	3.48
TOWR	18.04	23.11	12.20	0.09	2.25
TPIA	-22.52	-5.15	15.36	0.17	2.11

p-ISSN : 2774-9576
e-ISSN : 2747-2981

UNTR	-14.45	-16.44	7.23	0.22	2.22
UNVR	-44.88	-28.22	14.95	0.37	2.10
WIKA	-44.61	-33.67	17.02	0.45	3.33
Average	-12.94	-8.42	11.86	0.28	2.40

The following was for the second component stock group:

Table 7. Simulated Trading Return of Component 2 (in percentage)

Component 2	BnH total return	BnH average daily return	BnH daily stdev	Daily RoR	Daily stdev
ADRO	58.76	-2.64	18.83	2.32	0.14
AKRA	25.08	10.54	14.22	2.20	0.14
BBCA	6.80	-1.60	7.07	1.37	0.20
BBNI	5.49	-7.60	11.39	1.84	0.18
BMRI	9.23	-0.95	7.78	1.68	0.28
EXCL	10.21	-8.72	12.97	2.27	0.10
ITMG	54.76	18.39	29.95	2.63	0.08
MDKA	52.12	9.56	16.62	3.04	-0.01
MIKA	-19.93	-9.16	8.70	2.25	0.32
TLKM	16.91	-0.44	8.12	1.68	-0.03
TPIA	-22.52	-5.15	15.36	2.11	0.17
Average	17.90	0.20	13.73	0.14	2.13

A hypothesis test was conducted to further validate daily return and risk. H_{a1} was accepted and H_{a2} rejected in 95% significance level. Result in the table:

Table 8. ANOVA Results

ANOVA	F	Sig.
Daily RoR between component 1 and 2	9.767	0.004
Daily Stdev between component 1 and 2	1.773	0.192

There was a significant difference in stock returns between the PCA identified groups while the risk associated with the return was not significantly different.

V. DISCUSSION

The previous research [5][9][12] attempted to forecast market prices or direction of it and then develop strategies based on the predicted prices. Due to the chaotic nature of the market, prediction is most often less accurate in real trading situation, as it is based on historical data. Uncertainty, on the other hand, is a reality in the market. This research filled a necessary gap in complementing previous research by identifying two distinct groups in LQ45 with distinct price movements and significantly different returns with similar risk. As a result, the findings of this study can be used to supplement previous attempts at price or direction prediction.

PCA has been shown to be beneficial in LQ45 stocks. The tables 6 and 7 illustrated the trading simulation using a buy-and-hold strategy in great detail. For the year 2021, the first component generated an average negative return on investment (-12.94%). PTPP, SMGR, UNVR, and WIKA were the worst-performing stocks while TBIG and ERAA were the best-performing stocks. Overall the majority of component 1 stocks performed poorly when used with buy and hold strategy. However, an intriguing contrast occured when the average daily rate of return on group 1 stocks was positive 0.28%. This implied that group 1 stocks were preferable for daily short term trading than buy and hold strategies.

The average buy and hold strategy produced a positive result in the second component (17.9%), and the average daily return was also positive (0.14%), albeit lower than in the first

p-ISSN : 2774-9576
e-ISSN : 2747-2981

component. This implied that the second group identified provided a higher rate of return for buy and hold strategies but a lower rate of return for daily trading setups than group 1. The risk appeared to be comparable between the two groups, with a difference of only 0.27 percent and thus no statistically significant difference at the 95 significance percent level.

PCA successfully explained 70.91% of price movement and identified two distinct groups of stocks to invest in. For investors with a high tolerance for risk and a shorter time horizon. Group one was the superior alternative. For investors with a low risk tolerance and a longer time horizon. Group two was a superior alternative. Due to the stochastic nature of the market, dynamic adjustments are required because price movement may change over time. The next research can then focus on developing a dynamic pca with the desired time frame to accommodate the stochastic world of stock investing.

VI. CONCLUSION

- The LQ45 index provides valuable insight into low stock risk groups in 2021, with an average risk of less than 3% for average daily rate of returns in all LQ45 stock members.
- PCA successfully identified two distinct price movement groups / components in LQ45, accounting for 70.91 percent of the variance.
- Component 1 stocks had a higher average daily rate of return with an average of 0.28 percent while component 2 stocks had a lower average daily rate of return with an average of 0.14 percent.
- There was no difference in the average risk of daily average return between the two components identified.
- PCA successfully identified higher rates of return groups with comparable risk for daily trading strategy and buy and hold strategies - Both in which may be suitable for investors with varying levels of risk tolerance.

REFERENCES

[1] Stock Exchange, Indonesia., (2022). Indonesia Stock Exchange: Stocks., https://www.idx.co.id/produk/saham/
[2] Mandiri, Bank. (2022). Risks in Mutual Fund Investment. https://mandiri-investasi.co.id/id/belajar-investasi/pusat-info/risiko-berinvestasi-di-reksa-dana/
[3] Stock Exchange, Indonesia., (2021). IDX Stock Index Handbook v1.2. Jakarta: IDX.
[4] Grimes, A. (2012). The Art and Science of Technical Analysis. New Jersey: John Wiley & Sons.
[5] Faurina, R., Winduratna, B. and Nugroho P.(2018). Predicting Stock Movement Using Unidirectional LSTM and Feature Reduction: The Case of An Indonesia Stock. Int. Conference of Electrical Engineering and Computer Science (ICEECS). 180–185.
[6] Xie, H., Fan, K., and Wang, S. (2021). Candlestick Forecasting for Investments. New york: Routledge.
[7] Zhong, X., & Enke, D. (2017). Forecasting daily stock market return using dimensionality reduction. Expert Systems with Applications. 67, 1–33.
[8] Hansun S., & Young, J. C. (2021). Predicting LQ45 financial sector indices using RNN-LSTM. Journal of Big Data, 8(104).
[9] Darma, Y. D. (2021), Empirical Test of Apt Model to Predicting Portofolio's Stock Return Incorporated with Lq45 from 2014 until 2018 in Indonesia, INCEESS 2020. 199–214.
[10] Jollife, I.T (2002). Principal components analysis. New York: Springer.
[11] Jollife, I. T. & Cadima, J. (2016). Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065).
[12] Ansari, R., Al Hashfi R. U., & Setiyono, B. (2020). Examining Causality Effects On Stock Returns, Foreign Equity Inflow, and Investor Sentiment: Evidence From Indonesian Islamic Stocks, Indonesian. Capital. Market. Review., 12(2), 120–136.

p-ISSN : 2774-9576
e-ISSN : 2747-2981

[13] Kontan. (2021). LQ45 Members. https://www.kontan.co.id/indeks-lq45.
[14] Yahoo. (2021). Yahoo Finance, https://finance.yahoo.com/
[15] PythonTeam. (2021). Python. https://www.python.org/doc/
[16] Acal, C., Aguilera, A. M., & Escabias, M., (2020), New modeling approaches based on varimax rotation of functional principal components," Mathematics, 8(11), 1–15
[17] Ross, S. A., et al. (2019). Corporate Finance, 12th ed. New York: McGraw-Hill.