



Article

Analisis Cluster Saham LQ45 Berdasarkan Return, Volatilitas, dan Volume Perdagangan

Imanuel Migel Rumbekwan¹, Dina Jumiatsu Fitri², Radian Januari Situmeang³

¹ Universitas Cenderawasih, Jayapura, Indonesia;
email: imanuelhebat02@gmail.com

² Universitas Cenderawasih, Jayapura, Indonesia;
email: d.j.fitri@fmipa.uncen.ac.id

³ Universitas Cenderawasih, Jayapura, Indonesia;
email: r.j.situmeang@fmipa.uncen.ac.id

* Corresponding email: imanuelhebat02@gmail.com

ABSTRACT

The stocks included in the LQ45 Index exhibit diverse characteristics in terms of return, volatility, and trading activity, making stock clustering important in investment analysis. This article aims to determine the optimal cluster structure of LQ45 stocks based on return, volatility, and trading volume using the hierarchical clustering method. The similarity among stocks is measured using Ward Linkage and Euclidean Distance, while data preprocessing involves logarithmic transformation and Z-score standardization. The optimal number of clusters is evaluated using the Silhouette Coefficient, Davies-Bouldin Index (DBI), and the SW/SB ratio. The results indicate that LQ45 stocks exhibit heterogeneous characteristics and can be optimally grouped into seven clusters. The resulting clusters represent distinct stock profiles, ranging from relatively stable and liquid stocks to stocks characterized by higher levels of return, volatility, and investment risk. These findings demonstrate that hierarchical clustering is effective in identifying stock characteristics based on data similarity and can provide useful information for investment analysis and decision-making in the Indonesian capital market.

ABSTRAK

Saham-saham yang tergabung dalam indeks LQ45 memiliki karakteristik yang beragam dari sisi return, volatilitas, dan aktivitas perdagangan sehingga pengelompokan saham menjadi penting dalam analisis investasi. Artikel ini bertujuan untuk menentukan struktur cluster optimal pada saham LQ45 berdasarkan variabel return, volatilitas, dan volume perdagangan menggunakan metode hierarchical clustering.

KEYWORDS

LQ45 Stocks,
Hierarchical
Clustering, Return,
Volatility, Trading
Volume

KATA KUNCI:

Hierarchical clustering,
return, saham LQ45,
volatilitas, volume
perdagangan



Pengukuran tingkat kemiripan antar saham dilakukan menggunakan Ward Linkage dan Euclidean Distance, sedangkan tahap preprocessing data meliputi transformasi logaritma dan standarisasi Z-score. Penentuan jumlah cluster optimal dievaluasi menggunakan Silhouette Coefficient, Davies-Bouldin Index (DBI), dan rasio SW/SB. Hasil analisis menunjukkan bahwa saham-saham LQ45 memiliki karakteristik yang heterogen dan dapat dikelompokkan ke dalam tujuh cluster optimal. Cluster yang terbentuk menunjukkan profil saham yang berbeda, mulai dari saham yang relatif stabil dan likuid hingga saham dengan tingkat return, volatilitas, dan risiko investasi yang tinggi. Temuan ini menunjukkan bahwa metode hierarchical clustering efektif dalam mengidentifikasi karakteristik saham berdasarkan tingkat kemiripan data serta dapat memberikan informasi yang bermanfaat dalam analisis investasi dan pengambilan keputusan di pasar modal Indonesia.

PENDAHULUAN

Pasar modal merupakan salah satu instrumen penting dalam sistem perekonomian yang berfungsi sebagai sarana investasi dan sumber pendanaan bagi perusahaan. Perkembangan pasar modal Indonesia dalam beberapa tahun terakhir menunjukkan peningkatan jumlah investor dan aktivitas perdagangan yang semakin aktif. Saham menjadi instrumen investasi yang banyak diminati karena menawarkan peluang memperoleh *return*, namun juga mengandung risiko akibat fluktuasi harga yang terjadi di pasar (Bodie dkk., 2021). Salah satu kelompok saham yang banyak diperhatikan investor adalah saham yang tergabung dalam indeks LQ45. Indeks LQ45 merupakan salah satu indeks saham yang terdapat di Bursa Efek Indonesia (BEI) yang terdiri atas 45 saham perusahaan dengan karakteristik tertentu, karena memiliki kapitalisasi pasar besar dan tingkat likuiditas yang relatif tinggi. Meskipun berada dalam indeks yang sama, setiap saham memiliki karakteristik yang berbeda dari sisi keuntungan, risiko, dan aktivitas perdagangan sehingga diperlukan pendekatan yang dapat mengidentifikasi kemiripan karakteristik antar saham secara lebih sistematis.



Gambar 1. Perbandingan Pergerakan IHSG dan LQ45 Tahun 2025 (Data Harian)

Sumber: Bursa Efek Indonesia (2025), data diolah.

Berdasarkan Gambar 1, indeks LQ45 dan IHSG menunjukkan pola pergerakan yang relatif searah selama periode pengamatan tahun 2025. Namun demikian, terdapat perbedaan

tingkat fluktuasi pada beberapa periode yang menunjukkan bahwa saham-saham dalam indeks LQ45 tidak memiliki karakteristik yang seragam. Perbedaan tersebut mengindikasikan adanya variasi tingkat *return*, volatilitas, dan volume perdagangan antar saham. Variasi karakteristik tersebut dapat digunakan sebagai dasar untuk mengelompokkan saham ke dalam kelompok-kelompok yang memiliki tingkat kemiripan tertentu sehingga karakteristik masing-masing kelompok dapat dipahami dengan lebih baik.

Metode *clustering* telah banyak digunakan dalam penelitian pasar modal untuk mengelompokkan saham berdasarkan tingkat kemiripan karakteristik tertentu. Penelitian sebelumnya menunjukkan bahwa variabel *return*, volatilitas, dan volume perdagangan dapat memberikan informasi yang relevan dalam pembentukan kelompok saham karena merepresentasikan aspek keuntungan, risiko, dan aktivitas perdagangan saham (Nanda dkk., 2021; Li dkk., 2022). Pada pasar modal Indonesia, penggunaan metode K-Means masih banyak ditemukan dalam pengelompokan saham dan umumnya berfokus pada kombinasi variabel *return* dan volatilitas (Siregar dan Pangaribuan, 2021; Ramadhani dkk., 2023). Penggunaan ketiga variabel tersebut menjadi penting karena *return* dapat menggambarkan tingkat keuntungan saham, volatilitas mencerminkan tingkat risiko atau perubahan harga, sedangkan volume perdagangan menggambarkan tingkat aktivitas transaksi saham.

Selain pemilihan variabel dan metode *clustering*, kualitas hasil pengelompokan juga perlu dievaluasi untuk memastikan bahwa kelompok yang terbentuk memiliki tingkat kemiripan yang baik di dalam kelompok dan perbedaan yang jelas antar kelompok. Silhouette Coefficient digunakan untuk melihat keseimbangan antara kekompakan dan pemisahan kelompok, sedangkan Davies-Bouldin Index (DBI) digunakan untuk menilai tingkat kemiripan antar-*cluster*. Rasio SW/SB digunakan sebagai ukuran perbandingan variasi dalam *cluster* terhadap variasi antar-*cluster*. Wang dkk. (2022) serta Kumar dan Ravi (2023) menunjukkan pentingnya evaluasi dalam menentukan struktur *cluster* yang sesuai. Namun, penelitian pengelompokan saham masih terbatas dalam membandingkan Hierarchical Clustering dan K-Means dengan menggunakan *return*, volatilitas, dan volume perdagangan secara bersamaan serta mengevaluasi kualitas hasil pengelompokan menggunakan beberapa indikator. Oleh karena itu, penelitian ini dilakukan untuk membandingkan kedua metode tersebut sehingga dapat diketahui metode yang memberikan hasil pengelompokan saham LQ45 yang lebih baik berdasarkan karakteristik keuntungan, risiko, dan aktivitas perdagangan.

Berdasarkan kondisi tersebut, artikel ini menerapkan metode *hierarchical clustering* untuk mengelompokkan saham LQ45 berdasarkan variabel *return*, volatilitas, dan volume perdagangan. Penentuan jumlah *cluster* optimal dilakukan menggunakan *Silhouette Coefficient*, *Davies-Bouldin Index* (DBI), dan rasio SW/SB sehingga kualitas hasil pengelompokan dapat dievaluasi secara lebih komprehensif. Hasil pengelompokan diharapkan dapat memberikan gambaran mengenai karakteristik saham LQ45 berdasarkan tingkat keuntungan, risiko, dan aktivitas perdagangan, serta menjadi informasi pendukung dalam analisis investasi. Oleh karena itu, tujuan artikel ini adalah menentukan jumlah *cluster* optimal dan mengidentifikasi karakteristik masing-masing *cluster* saham LQ45 berdasarkan variabel *return*, volatilitas, dan volume perdagangan.

METODE PENELITIAN

Data Penelitian

Penelitian ini menggunakan data sekunder berupa data harian saham yang tergabung dalam indeks LQ45 selama periode tahun 2025. Populasi penelitian mencakup seluruh saham yang menjadi konstituen indeks LQ45 selama periode pengamatan. Sampel penelitian terdiri atas 27 saham yang memenuhi kriteria ketersediaan data harga penutupan dan volume perdagangan selama periode pengamatan. Pemilihan sampel dilakukan untuk memastikan bahwa data yang digunakan tersedia secara lengkap dan dapat dibandingkan secara konsisten antar saham. Saham yang mengalami perubahan konstituen selama periode pengamatan diseleksi berdasarkan ketersediaan data dan kesesuaiannya dengan kriteria sampel. Data diperoleh dari Investing.com.

Variabel yang digunakan meliputi *return*, volatilitas, dan volume perdagangan saham. *Return* digunakan untuk menggambarkan tingkat perubahan atau keuntungan harga saham, volatilitas digunakan untuk mengukur tingkat fluktuasi *return* sebagai gambaran risiko saham, sedangkan volume perdagangan digunakan untuk merepresentasikan aktivitas transaksi saham.

Perhitungan *return* saham dilakukan berdasarkan perubahan harga penutupan (*closing price*) pada periode t terhadap harga penutupan pada periode sebelumnya ($t-1$). Perhitungan *return* mengacu pada Hartono (2017) dan ditunjukkan pada Persamaan (1).

$$R_{it} = \frac{P_{it} - P_{it-1}}{P_{it-1}} \quad (1)$$

Dengan R_{it} merupakan *return* saham i pada periode t , P_{it} merupakan harga penutupan saham i pada periode t , dan P_{it-1} merupakan harga penutupan saham i pada periode sebelumnya.

Perhitungan volatilitas dilakukan menggunakan standar deviasi *return* harian masing-masing saham selama periode pengamatan. Volatilitas digunakan untuk menggambarkan tingkat fluktuasi *return*, sehingga semakin besar nilai volatilitas menunjukkan semakin tinggi tingkat perubahan *return* saham. Perhitungan volatilitas mengacu pada Bodie dkk. (2021) dan ditunjukkan pada Persamaan (2).

$$\sigma_i = \sqrt{\frac{1}{n-1} \sum_{t=1}^n (R_{it} - \bar{R}_i)^2} \quad (2)$$

Dengan σ_i merupakan volatilitas saham i , R_{it} merupakan *return* saham i pada periode t , $\bar{R}_i$ merupakan rata-rata *return* saham i , dan n merupakan jumlah periode pengamatan.

Rata-rata volume perdagangan dihitung berdasarkan volume perdagangan harian masing-masing saham selama periode pengamatan. Penggunaan rata-rata volume perdagangan bertujuan untuk menggambarkan tingkat aktivitas transaksi saham selama periode penelitian. Perhitungan rata-rata volume perdagangan ditunjukkan pada Persamaan (3).

$$V_i = \frac{1}{n} \sum_{t=1}^T Volume_{it} \quad (3)$$

Dengan V_i merupakan rata-rata volume perdagangan saham i , $Volume_{it}$ merupakan volume perdagangan saham i pada periode t , dan n merupakan jumlah periode pengamatan.

Preprocessing Data

Tahap *preprocessing* dilakukan untuk meningkatkan kualitas data sebelum proses pengelompokan saham. Tahap awal dilakukan dengan memeriksa ketersediaan data dan menangani nilai yang hilang. Apabila ditemukan nilai yang hilang pada variabel penelitian, penanganan dilakukan menggunakan nilai rata-rata variabel terkait (Little & Rubin, 2019). Selanjutnya, transformasi logaritma diterapkan pada variabel volume perdagangan untuk mengurangi kemencengan distribusi dan pengaruh nilai yang sangat besar (Gujarati & Porter, 2009).

Setelah transformasi dilakukan, seluruh variabel distandarisasi menggunakan metode *Z-score* agar setiap variabel memiliki skala yang sebanding sehingga tidak terdapat variabel yang mendominasi perhitungan jarak pada proses *clustering* (Hair dkk., 2019). Transformasi logaritma ditunjukkan pada Persamaan (4).

$$X' = \log(X) \quad (4)$$

Standarisasi Z-score ditunjukkan pada Persamaan (5).

$$Z_{ik} = \frac{X_{ik} - \bar{X}_k}{S_k} \quad (5)$$

Dengan X_{ik} merupakan nilai variabel ke- k pada saham i , $\bar{X}_k$ merupakan rata-rata variabel ke- k , dan S_k merupakan standar deviasi variabel ke- k .

Deteksi *outlier* dilakukan menggunakan kriteria $|Z| > 3$ (Aggarwal, 2017). Deteksi ini bertujuan untuk mengidentifikasi nilai yang memiliki penyimpangan relatif besar terhadap distribusi data. Nilai yang teridentifikasi sebagai *outlier* diperiksa terlebih dahulu dan tidak secara otomatis dihilangkan karena dapat merepresentasikan karakteristik aktual pergerakan saham selama periode pengamatan. Selain itu, analisis korelasi Pearson digunakan untuk mengetahui hubungan antarvariabel yang digunakan dalam penelitian (Widarjono, 2018).

Hierarchical Clustering

Metode *hierarchical clustering* digunakan untuk mengelompokkan saham berdasarkan tingkat kemiripan karakteristik data. Pengukuran kemiripan antar saham dilakukan menggunakan *Euclidean Distance* karena metode tersebut mengukur jarak antarobjek berdasarkan nilai seluruh variabel penelitian (Everitt dkk., 2011). Perhitungan *Euclidean Distance* ditunjukkan pada Persamaan (6).

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (6)$$

Dengan d_{ij} merupakan jarak antara saham i dan saham j , x_{ik} merupakan nilai variabel ke- k pada saham i , x_{jk} merupakan nilai variabel ke- k pada saham j , dan p merupakan jumlah variabel yang digunakan.

Proses penggabungan *cluster* dilakukan secara bertahap menggunakan metode *Ward Linkage*. Metode ini menggabungkan dua *cluster* yang menghasilkan peningkatan variasi dalam *cluster* paling kecil sehingga kelompok yang terbentuk memiliki karakteristik yang relatif homogen (Murtagh & Legendre, 2014). Kenaikan *within-cluster sum of squares* digunakan sebagai dasar dalam proses penggabungan dan ditunjukkan pada Persamaan (7).

$$\Delta SSE = \sum (x - \bar{x})^2 \quad (7)$$

Hasil pengelompokan menggunakan hierarchical clustering selanjutnya divisualisasikan dalam bentuk dendrogram untuk melihat struktur penggabungan dan hubungan antar saham berdasarkan tingkat kemiripannya.

Evaluasi Cluster

Penentuan jumlah *cluster* optimal dilakukan menggunakan *Silhouette Coefficient*, *Davies-Bouldin Index* (DBI), dan rasio SW/SB. *Silhouette Coefficient* digunakan untuk mengukur tingkat kekompakan data dalam *cluster* dan pemisahan antar-*cluster*, sedangkan DBI digunakan untuk mengevaluasi homogenitas dalam *cluster* dan pemisahan antar-*cluster* (Arbelaitz dkk., 2013). Rasio SW/SB digunakan untuk membandingkan variasi dalam *cluster* dengan variasi antar-*cluster* (Kaufman & Rousseeuw, 2005). Nilai *Silhouette Coefficient* yang lebih tinggi, serta nilai DBI dan rasio SW/SB yang lebih rendah menunjukkan kualitas pengelompokan yang lebih baik.

Kualitas hasil pengelompokan dinilai berdasarkan nilai *Silhouette Coefficient* yang lebih tinggi, sedangkan nilai DBI dan rasio SW/SB yang lebih rendah menunjukkan kualitas pengelompokan yang lebih baik. *Silhouette Coefficient* ditunjukkan pada Persamaan (8).

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (8)$$

Dengan $a(i)$ merupakan rata-rata jarak objek i terhadap objek lain dalam cluster yang sama, sedangkan $b(i)$ merupakan rata-rata jarak objek i terhadap objek pada cluster terdekat.

Persamaan DBI ditunjukkan pada Persamaan (9).

$$DBI = \frac{1}{k} \sum_{i=1}^k \max \left(\frac{S_i + S_j}{M_{ij}} \right) \quad (9)$$

Dengan k merupakan jumlah *cluster*, S_i dan S_j merupakan ukuran penyebaran objek

dalam *cluster i dan j*, sedangkan M_{ij} merupakan jarak antara pusat *cluster i dan j*.

Nilai SW/SB diperoleh dari perbandingan *within-cluster variance* (SW) dan *between-cluster variance* (SB). Jumlah *cluster* optimal ditentukan dengan mempertimbangkan hasil ketiga indikator secara bersama-sama. Apabila ketiga indikator memberikan rekomendasi jumlah *cluster* yang berbeda, pemilihan jumlah *cluster* dilakukan berdasarkan hasil evaluasi secara keseluruhan dengan mempertimbangkan indikator yang memberikan hasil terbaik dan karakteristik *cluster* yang terbentuk. Dengan demikian, jumlah *cluster* yang dipilih merupakan jumlah yang memberikan kualitas pengelompokan yang paling sesuai berdasarkan ketiga indikator tersebut.

HASIL DAN PEMBAHASAN

Analisis Variabel Penelitian

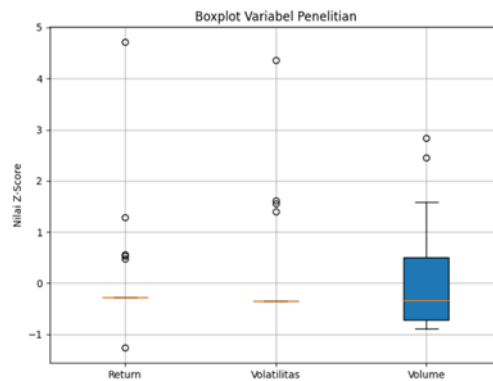
Tabel 1. Deskriptif Variabel Penelitian

Variabel	Min	Max	Mean	Std Deviasi	Kurtosis
X_1 (Return)	-4.9283	24.8673	1.3249	4.9089	19.3591
X_2 (Volatilitas)	0.0140	154.4620	11.3186	32.8938	12.5914
X_3 (Volume)	6.3502	260.3808	64.9405	68.4186	1.7562

Sumber: Hasil pengolahan data menggunakan Excel, 2026

Berdasarkan Tabel 1, variabel *return* (X_1) memiliki rentang nilai yang cukup besar dengan nilai minimum sebesar -4.9283 dan maksimum sebesar 24.8673. Perbedaan rentang tersebut menunjukkan bahwa tingkat keuntungan antar saham LQ45 selama periode pengamatan tahun 2025 cenderung beragam. Nilai standar deviasi sebesar 4.9089 memperlihatkan bahwa sebagian saham mengalami perubahan *return* yang lebih tinggi dibandingkan saham lainnya. Pada variabel volatilitas (X_2), nilai minimum sebesar 0.0140 dan maksimum sebesar 154.4620 menunjukkan adanya perbedaan tingkat risiko antar saham yang cukup besar. Nilai standar deviasi sebesar 32.8938 menunjukkan bahwa beberapa saham memiliki pergerakan harga yang lebih aktif selama periode pengamatan. Sementara itu, variabel volume perdagangan (X_3) memiliki rata-rata sebesar 64.9405 dengan standar deviasi sebesar 68.4186 yang menunjukkan bahwa aktivitas transaksi saham LQ45 juga cukup beragam. Perbedaan karakteristik pada variabel *return*, volatilitas, dan volume perdagangan menunjukkan bahwa saham-saham dalam indeks LQ45 memiliki pola keuntungan, risiko, dan aktivitas perdagangan yang tidak seragam sehingga mendukung penggunaan metode *clustering* dalam proses pengelompokan saham.

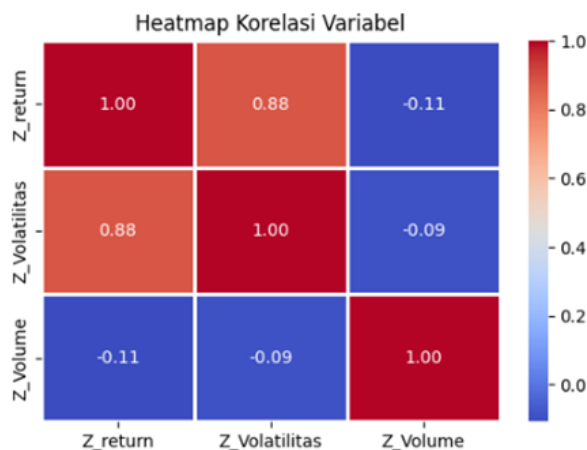
Preprocessing Data



Gambar 2. BoxPlot Variabel Penelitian

Sumber: Hasil pengolahan data menggunakan Python, 2026.

Gambar 2 menunjukkan sebaran data variabel *return*, volatilitas, dan volume perdagangan setelah proses transformasi logaritma dan standarisasi menggunakan metode *Z-score*. Transformasi logaritma dilakukan pada variabel volume perdagangan untuk mengurangi perbedaan nilai yang terlalu besar antar saham, sedangkan standarisasi *Z-score* digunakan agar seluruh variabel memiliki skala yang sama dalam proses *clustering*. Variabel *return* dan volatilitas memiliki rentang data yang relatif sempit pada sebagian besar saham, namun terdapat beberapa *outlier* yang menunjukkan adanya saham dengan karakteristik berbeda dibandingkan mayoritas saham lainnya. *Outlier* tertinggi pada variabel *return* berasal dari saham CTRA, sedangkan *outlier* pada volatilitas berasal dari saham CTRA, BRPT, BBTN, dan AKRA yang memiliki perubahan harga lebih tinggi dibandingkan saham lainnya selama periode pengamatan. Variabel volume perdagangan memiliki rentang data yang lebih lebar dibandingkan *return* dan volatilitas yang menunjukkan bahwa aktivitas perdagangan saham LQ45 lebih beragam. *Outlier* pada volume perdagangan berasal dari saham BBRI, BBKA, dan BMRI yang memiliki aktivitas transaksi jauh lebih tinggi dibandingkan saham lainnya. Perbedaan sebaran data dan keberadaan *outlier* pada setiap variabel menunjukkan bahwa saham-saham LQ45 memiliki karakteristik keuntungan, risiko, dan aktivitas perdagangan yang tidak seragam sehingga mendukung penggunaan metode *clustering* dalam proses pengelompokan saham.

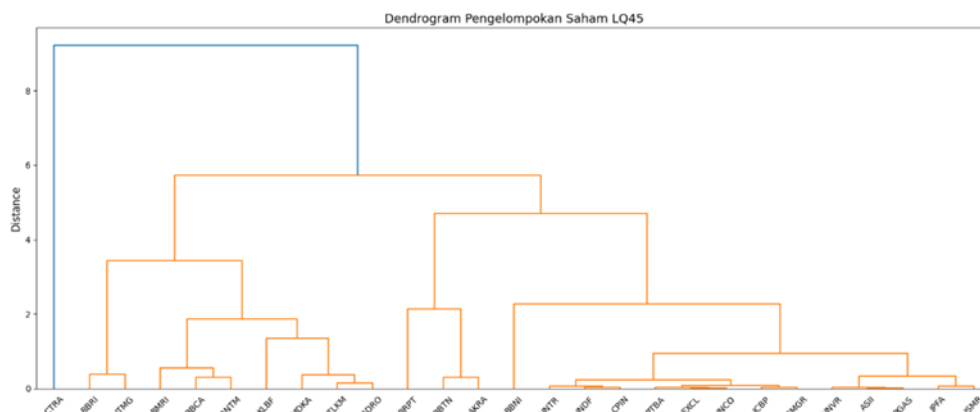


Gambar 3. Heatmap Korelasi Antar Variabel

Sumber: Hasil pengolahan data menggunakan Python, 2026.

Gambar 3 menunjukkan hubungan antar variabel penelitian menggunakan *heatmap* korelasi. Variabel *return* dan volatilitas memiliki korelasi positif yang cukup tinggi sebesar 0.88, yang menunjukkan bahwa saham dengan tingkat keuntungan tinggi cenderung memiliki tingkat risiko yang lebih tinggi. Sementara itu, variabel volume perdagangan memiliki korelasi negatif yang rendah terhadap *return* dan volatilitas. Nilai korelasi tersebut menunjukkan bahwa aktivitas perdagangan saham tidak memiliki hubungan linear yang kuat terhadap tingkat keuntungan maupun risiko saham. Secara umum, hubungan antar variabel masih berada pada tingkat yang dapat diterima sehingga setiap variabel tetap memberikan informasi yang berbeda dalam proses *clustering* saham LQ45.

Hierarchical Clustering dan Evaluasi Cluster



Gambar 4. Dendrogram Pengelompokan Saham LQ45

Sumber: Hasil pengolahan data menggunakan Python, 2026.

Gambar 4 menunjukkan hasil pengelompokan saham LQ45 menggunakan metode *hierarchical clustering* dengan Ward Linkage dan ukuran jarak *Euclidean Distance*. Dendrogram memperlihatkan bahwa saham-saham LQ45 membentuk beberapa kelompok berdasarkan tingkat kemiripan karakteristik *return*, volatilitas, dan volume perdagangan. Saham yang bergabung pada jarak (*distance*) yang rendah menunjukkan karakteristik yang lebih mirip dibandingkan saham yang bergabung pada jarak yang lebih tinggi. Beberapa pasangan saham seperti BBRI–ITMG, BBKA–BMRI, dan PGAS–JPFA bergabung pada jarak yang relatif dekat sehingga menunjukkan pola keuntungan, risiko, dan aktivitas perdagangan yang hampir serupa selama periode pengamatan. Sebaliknya, saham CTRA berada pada cabang yang terpisah dengan jarak penggabungan paling tinggi dibandingkan saham lainnya. Perbedaan tersebut menunjukkan bahwa saham CTRA memiliki karakteristik yang paling berbeda, terutama dari sisi *return* dan volatilitas yang cenderung lebih tinggi dibandingkan mayoritas saham LQ45 lainnya. Secara umum, struktur dendrogram menunjukkan bahwa saham-saham LQ45 memiliki karakteristik yang heterogen sehingga metode *hierarchical clustering* mampu menggambarkan hubungan kedekatan antar saham secara bertahap dan lebih jelas.

Berdasarkan struktur pengelompokan pada Gambar 4, penentuan jumlah *cluster* optimal selanjutnya dievaluasi secara numerik menggunakan *Silhouette Coefficient*, Davies-Bouldin Index (DBI), dan rasio SW/SB. Nilai *Silhouette Coefficient* yang lebih tinggi, sedangkan nilai DBI dan rasio SW/SB yang lebih rendah menunjukkan kualitas pengelompokan yang lebih baik.

Tabel 2. Hasil Evaluasi Jumlah Cluster

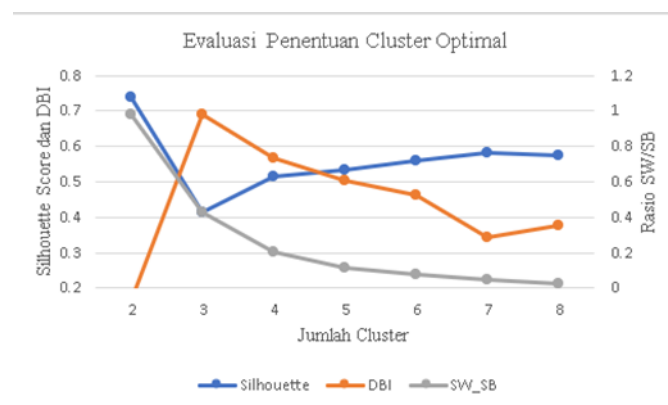
Jumlah Cluster	Silhouette	DBI	SW/SB
2	0.741	0.165	0.983
3	0.415	0.690	0.430
4	0.514	0.568	0.205
5	0.534	0.504	0.111
6	0.560	0.464	0.075
7	0.580	0.344	0.044
8	0.574	0.375	0.022

Sumber: Hasil pengolahan data menggunakan Excel, 2026

Berdasarkan Tabel 3, nilai *Silhouette Coefficient* tertinggi diperoleh pada dua *cluster* sebesar 0,740644. Namun, setelah mengalami penurunan pada tiga *cluster* sebesar 0,415452, nilai tersebut meningkat secara bertahap hingga mencapai 0,580399 pada tujuh *cluster*. Pada delapan *cluster*, nilai *Silhouette Coefficient* sedikit menurun menjadi 0,573593. Pola tersebut menunjukkan bahwa kualitas kekompakan dan pemisahan *cluster* kembali meningkat hingga tujuh *cluster*, tetapi tidak mengalami peningkatan ketika jumlah *cluster* ditambah menjadi delapan.

Pada indikator DBI, nilai terendah diperoleh pada dua *cluster*, yaitu sebesar 0,16492. Nilai tersebut meningkat menjadi 0,68992 pada tiga *cluster*, kemudian menurun secara bertahap hingga mencapai 0,34410 pada tujuh *cluster*. Pada delapan *cluster*, nilai DBI meningkat menjadi 0,37512. Penurunan nilai DBI dari tiga hingga tujuh *cluster* menunjukkan bahwa kualitas pemisahan dan kekompakan *cluster* semakin baik, sedangkan peningkatan pada delapan *cluster* menunjukkan bahwa penambahan kelompok tidak lagi memberikan perbaikan berdasarkan indikator DBI.

Rasio SW/SB menunjukkan pola yang berbeda, yaitu terus menurun dari 0,98303 pada dua *cluster* hingga mencapai 0,02224 pada delapan *cluster*. Pada tujuh *cluster*, rasio SW/SB sebesar 0,04428 menunjukkan bahwa variasi dalam *cluster* relatif kecil dibandingkan variasi antar-*cluster*. Meskipun nilai terendah diperoleh pada delapan *cluster*, hasil tersebut perlu dipertimbangkan bersama dengan *Silhouette Coefficient* dan DBI.



Gambar 5. Hasil Evaluasi Cluster

Sumber: Hasil pengolahan data menggunakan Excel, 2026

Gambar 5 menunjukkan hasil evaluasi jumlah *cluster* menggunakan *Silhouette Coefficient*, *Davies-Bouldin Index* (DBI), dan rasio SW/SB pada jumlah *cluster* 2 hingga 8. Nilai

Silhouette Coefficient mengalami penurunan dari *cluster* 2 ke *cluster* 3, kemudian meningkat secara bertahap hingga *cluster* 7 sebelum kembali relatif stabil pada *cluster* 8. Pola tersebut menunjukkan bahwa kualitas pemisahan antar *cluster* menjadi lebih baik seiring bertambahnya jumlah *cluster*. Sementara itu, nilai DBI mengalami penurunan mulai dari *cluster* 3 hingga *cluster* 7, yang menunjukkan bahwa jarak antar *cluster* semakin baik dan tingkat kemiripan data dalam setiap *cluster* menjadi lebih tinggi. Rasio SW/SB juga terus menurun hingga *cluster* 7 yang menunjukkan bahwa homogenitas data dalam *cluster* semakin baik dibandingkan pemisahan antar *cluster*. Berdasarkan ketiga indikator evaluasi tersebut, jumlah *cluster* terbaik diperoleh pada *cluster* 7 karena memiliki kombinasi nilai *Silhouette Coefficient* yang relatif tinggi, nilai DBI yang lebih rendah, serta rasio SW/SB yang paling kecil dibandingkan jumlah *cluster* lainnya. Hasil tersebut menunjukkan bahwa *cluster* 7 mampu menghasilkan struktur pengelompokan saham yang lebih baik berdasarkan karakteristik *return*, volatilitas, dan volume perdagangan saham LQ45.

Karakteristik Cluster Saham LQ45

Tabel 2. Karakteristik Cluster

Cluster	Jumlah Saham	Mean Return	Mean Volatilitas	Mean Volume	Karakteristik	Anggota Saham
Cluster 1	2	-0.27482	-0.34353	2.64425	Likuiditas Tinggi dengan Risiko Rendah	BBRI, ITMG
Cluster 2	7	-0.41568	-0.34323	0.69995	Stabil Moderat	BBCA, BMRI, TLKM, KLBF, ADRO, ANTM, MDKA
Cluster 3	2	0.52199	1.51006	-0.51057	Risiko Tinggi	BBTN, AKRA
Cluster 4	1	0.53953	1.55449	1.33969	Aktif dan Agresif	BRPT
Cluster 5	13	-0.27459	-0.34336	-0.63748	Stabil Konservatif	ASII, EXCL, ICBP, INDF, UNVR, CPIN, JPFA, SMGR, JSMR, PGAS, PTBA, INCO, UNTR
Cluster 6	1	1.28552	-0.34336	-0.04192	Return Tinggi Risiko Rendah	BBNI
Cluster 7	1	4.70944	4.35168	-0.56452	Sangat Agresif	CTRA

Sumber: Hasil pengolahan data menggunakan Python, 2026.

Hasil *hierarchical clustering* menghasilkan tujuh kelompok saham dengan karakteristik yang berbeda berdasarkan nilai rata-rata *return*, volatilitas, dan volume perdagangan. Perbedaan karakteristik tersebut menunjukkan adanya heterogenitas profil saham LQ45, baik dari sisi tingkat keuntungan, fluktuasi harga, maupun aktivitas

perdagangan. Cluster 1 terdiri atas saham BBRI dan ITMG yang memiliki tingkat likuiditas tinggi dengan volatilitas relatif rendah, sehingga mencerminkan kelompok saham yang aktif diperdagangkan dengan tingkat fluktuasi harga yang relatif lebih terkendali. Cluster 2 merupakan kelompok terbesar kedua yang didominasi saham berkapitalisasi besar, seperti BBKA, BMRI, dan TLKM, dengan karakteristik yang relatif stabil. Sebaliknya, Cluster 3 dan Cluster 4 menunjukkan tingkat volatilitas yang lebih tinggi sehingga mencerminkan kelompok saham dengan fluktuasi harga yang relatif lebih besar dibandingkan kelompok lainnya. Cluster 5 terdiri atas mayoritas saham dengan nilai *return*, volatilitas, dan volume perdagangan yang cenderung berada di bawah rata-rata. Cluster 6 hanya terdiri atas saham BBNI yang menunjukkan *return* relatif tinggi dengan volatilitas yang masih rendah. Sementara itu, Cluster 7 yang hanya beranggotakan saham CTRA menunjukkan nilai *return* dan volatilitas tertinggi sehingga merepresentasikan kelompok dengan karakteristik yang paling agresif selama periode pengamatan.

Keragaman karakteristik tersebut menunjukkan bahwa pengelompokan berdasarkan sektor atau kapitalisasi saja belum tentu mampu menggambarkan kesamaan karakteristik saham secara empiris. Bagi investor, hasil *clustering* dapat digunakan sebagai dasar awal untuk melakukan segmentasi saham berdasarkan karakteristik *return*, volatilitas, dan aktivitas perdagangan. Kelompok dengan volatilitas relatif rendah menunjukkan karakteristik fluktuasi harga yang lebih terbatas, sedangkan kelompok dengan *return* dan volatilitas tinggi menunjukkan karakteristik perubahan nilai yang lebih besar. Sementara itu, kelompok dengan volume perdagangan relatif tinggi mencerminkan aktivitas transaksi yang lebih aktif. Informasi tersebut dapat menjadi bahan pertimbangan dalam membentuk dan mengevaluasi portofolio berdasarkan karakteristik saham yang diinginkan. Meskipun demikian, hasil *clustering* tidak dimaksudkan sebagai rekomendasi langsung untuk membeli atau menjual saham tertentu, karena keputusan investasi tetap perlu mempertimbangkan faktor fundamental, kondisi pasar, tingkat toleransi risiko, dan tujuan investasi.

Temuan tersebut selanjutnya memiliki relevansi dalam konteks manajemen portofolio berbasis Modern Portfolio Theory (MPT). Markowitz (1952) menjelaskan bahwa risiko portofolio dapat dikurangi melalui diversifikasi antar-aset yang tidak berkorelasi sempurna untuk membentuk *efficient frontier* (Bodie, Kane, & Marcus, 2014; Elton, Gruber, Brown, & Goetzmann, 2009). Dalam kerangka tersebut, *clustering* dapat berfungsi sebagai tahap penyaringan awal dalam proses pemilihan aset (Tola, Lillo, Gallegati, & Mantegna, 2008; Walker, 2014). Hasil penelitian menunjukkan bahwa saham perbankan besar (*Big Four*) tidak terkonsentrasi dalam satu kelompok, melainkan tersebar ke beberapa *cluster*: BBRI berada pada Cluster 1 dengan likuiditas tinggi dan volatilitas relatif rendah, BBKA dan BMRI berada pada Cluster 2 dengan karakteristik relatif stabil, BBNI berada pada Cluster 6 dengan *return* relatif tinggi dan volatilitas yang masih rendah, sedangkan BBTN berada pada Cluster 3 dengan volatilitas yang lebih tinggi. Pola tersebut memperlihatkan bahwa saham dalam sektor yang sama tidak selalu memiliki karakteristik statistik yang seragam. Dengan demikian, klasifikasi berdasarkan sektor tidak cukup apabila digunakan sebagai satu-satunya dasar dalam menentukan strategi diversifikasi (Mantegna, 1999).

Meskipun demikian, hasil *clustering* perlu diposisikan sebagai informasi pendukung, bukan sebagai pengganti analisis risiko portofolio secara keseluruhan. Penggunaan saham dari *cluster* yang berbeda dapat menjadi salah satu pendekatan awal dalam membangun

portofolio *mean-variance*. Namun, karena pengelompokan dalam penelitian ini tidak mengukur hubungan korelasi antar-saham secara langsung, estimasi matriks kovarians tetap diperlukan untuk mengevaluasi manfaat diversifikasi secara lebih akurat. Selain itu, penggunaan data historis dalam mengestimasi *expected return* memiliki potensi *estimation error*, terutama ketika periode observasi relatif pendek (DeMiguel, Garlappi, & Uppal, 2009).

Berdasarkan karakteristik masing-masing kelompok, hasil segmentasi juga dapat menjadi dasar dalam menyusun alternatif alokasi aset sesuai dengan karakteristik risiko investor. Investor dengan karakteristik konservatif dapat mempertimbangkan saham dengan volatilitas relatif rendah, seperti saham pada Cluster 5 (ICBP, INDF, JSMR) dan Cluster 2 (BBCA, BMRI). Investor dengan karakteristik moderat dapat mempertimbangkan kombinasi saham dari beberapa *cluster*, misalnya Cluster 2 dan Cluster 5 sebagai kelompok utama, kemudian melengkapinya dengan Cluster 1 (BBRI, ITMG) yang memiliki aktivitas perdagangan relatif tinggi dan Cluster 6 (BBNI) yang memiliki *return* relatif tinggi. Sementara itu, investor dengan karakteristik agresif dapat mempertimbangkan saham dengan volatilitas lebih tinggi dari Cluster 3 (BBTN, AKRA), Cluster 4 (BRPT), dan Cluster 7 (CTRA) sebagai bagian *satellite* dalam pendekatan *core-satellite* (Amenc, Malaise, & Martellini, 2004). Namun, pembagian tersebut sebaiknya dipahami sebagai ilustrasi pemanfaatan hasil segmentasi, bukan sebagai rekomendasi investasi individual. Penentuan alokasi akhir tetap memerlukan pertimbangan terhadap korelasi antar-aset, kondisi fundamental perusahaan, kondisi pasar, serta tujuan dan batasan risiko investor.

KESIMPULAN

Penelitian ini bertujuan untuk menentukan struktur *cluster* optimal saham yang tergabung dalam indeks LQ45 berdasarkan *return*, volatilitas, dan volume perdagangan menggunakan metode *hierarchical clustering*. Hasil penelitian menunjukkan bahwa saham-saham LQ45 memiliki karakteristik yang heterogen dalam aspek keuntungan, risiko, dan aktivitas perdagangan. Penerapan *hierarchical clustering* dengan *Ward Linkage* dan *Euclidean Distance* menghasilkan struktur pengelompokan yang dapat membedakan saham berdasarkan tingkat kemiripan karakteristiknya. Evaluasi menggunakan *Silhouette Coefficient*, *Davies-Bouldin Index* (DBI), dan rasio SW/SB menunjukkan kecenderungan jumlah *cluster* yang berbeda. Setelah mempertimbangkan hasil evaluasi secara keseluruhan serta interpretabilitas struktur kelompok, tujuh *cluster* dipilih sebagai struktur yang paling sesuai. Kelompok yang terbentuk menunjukkan karakteristik yang beragam, mulai dari saham yang relatif stabil dan likuid hingga saham dengan *return* dan volatilitas yang lebih tinggi, termasuk BRPT dan CTRA. BBNI juga membentuk kelompok tersendiri karena memiliki *return* yang relatif tinggi dengan volatilitas yang lebih rendah dibandingkan kelompok saham dengan karakteristik risiko tinggi. Penelitian ini memiliki keterbatasan pada penggunaan tiga variabel utama, yaitu *return*, volatilitas, dan volume perdagangan, sehingga karakteristik saham yang terbentuk belum mencakup aspek fundamental dan indikator pasar lainnya. Selain itu, hasil pengelompokan sangat bergantung pada periode pengamatan, variabel yang digunakan, serta metode *clustering* dan ukuran jarak yang diterapkan. Oleh karena itu, hasil *cluster* dalam penelitian ini perlu dipahami sebagai representasi karakteristik saham pada konteks data dan periode penelitian, bukan sebagai klasifikasi yang bersifat permanen. Penelitian selanjutnya dapat memperluas analisis dengan

memasukkan variabel fundamental, seperti profitabilitas, leverage, dan ukuran perusahaan, maupun indikator teknikal lainnya. Penelitian berikutnya juga dapat membandingkan *hierarchical clustering* dengan metode *clustering* lainnya untuk menguji konsistensi struktur kelompok yang dihasilkan. Pendekatan tersebut diharapkan dapat memberikan pemahaman yang lebih komprehensif mengenai karakteristik dan segmentasi saham di pasar modal Indonesia.

DAFTAR PUSTAKA

- Aggarwal, C. C. (2017). *Outlier analysis* (2nd ed.). Cham, Switzerland: Springer.
<https://doi.org/10.1007/978-3-319-47578-3>
- Amenc, N., Malaise, P., & Martellini, L. (2004). Revisiting core-satellite investing. *The Journal of Portfolio Management*, 31(1), 64–75. <https://doi.org/10.3905/jpm.2004.443322>
- Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., & Perona, I. (2013). An extensive comparative study of cluster validity indices. *Pattern Recognition*, 46(1), 243–256.
<https://doi.org/10.1016/j.patcog.2012.07.021>
- Bodie, Z., Kane, A., & Marcus, A. J. (2014). *Investments*. Singapore: McGraw-Hill Education Asia.
- Bodie, Z., Kane, A., & Marcus, A. J. (2021). *Investments* (12th ed.). New York, NY: McGraw-Hill Education.
- DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, 22(5), 1915–1953. <https://doi.org/10.1093/rfs/hhm075>
- Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzmann, W. N. (2009). *Modern portfolio theory and investment analysis* (8th ed.). Hoboken, NJ: John Wiley & Sons.
- Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). Chichester, UK: John Wiley & Sons.
- Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics* (5th ed.). New York, NY: McGraw-Hill Education.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Hartono, J. (2017). *Teori portofolio dan analisis investasi* (edisi ke-11). Yogyakarta: BPFE Yogyakarta.
- Jones, C. P. (2014). *Investments: Principles and concepts* (12th ed.). Hoboken, NJ: John Wiley & Sons.
- Kaufman, L., & Rousseeuw, P. J. (2005). *Finding groups in data: An introduction to cluster analysis*. Hoboken, NJ: John Wiley & Sons.
- Kumar, R., & Ravi, V. (2023). Hierarchical clustering framework for financial market segmentation using stock market indicators. *Expert Systems with Applications*, 213, 118986. <https://doi.org/10.1016/j.eswa.2022.118986>
- Li, Y., Zhang, X., & Chen, H. (2022). Stock market clustering analysis using return, volatility, and trading volume indicators. *Physica A: Statistical Mechanics and Its Applications*, 603, 127746. <https://doi.org/10.1016/j.physa.2022.127746>
- Little, R. J. A., & Rubin, D. B. (2019). *Statistical analysis with missing data* (3rd ed.). Hoboken, NJ: John Wiley & Sons.
-

- Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B - Condensed Matter and Complex Systems*, 11(1), 193–197.
<https://doi.org/10.1007/s100510050929>
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.
<https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
- Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classification*, 31(3), 274–295. <https://doi.org/10.1007/s00357-014-9161-z>
- Nanda, S. R., Mahapatra, D., & Swain, S. N. (2021). Clustering financial assets based on return and volatility for portfolio diversification. *Journal of King Saud University - Computer and Information Sciences*, 33(7), 876–884.
<https://doi.org/10.1016/j.jksuci.2019.04.003>
- Ramadhani, F., Nugroho, A., & Prasetyo, E. (2023). Analisis pengelompokan saham sektor unggulan menggunakan metode clustering pada pasar modal Indonesia. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 7(2), 341–349.
<https://doi.org/10.29207/resti.v7i2.4820>
- Reilly, F. K., & Brown, K. C. (2011). *Investment analysis and portfolio management* (10th ed.). Mason, OH: Cengage Learning.
- Siregar, H., & Pangaribuan, R. (2021). Portfolio optimization based on clustering of Indonesia Stock Exchange: A case study of LQ45. *International Journal of Advanced Computer Science and Applications*, 12(6), 667–674.
<https://doi.org/10.14569/IJACSA.2021.0120677>
- Tola, V., Lillo, F., Gallegati, M., & Mantegna, R. N. (2008). Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, 32(1), 235–258.
<https://doi.org/10.1016/j.jedc.2007.01.034>
- Walker, S. T. (2014). *Understanding alternative investments: Creating diversified portfolios that ride the wave of investment success*. New York, NY: Palgrave Macmillan US.
<https://doi.org/10.1057/9781137370198>
- Wang, J., Liu, Y., & Zhao, L. (2022). Hierarchical clustering analysis of global stock market structure under financial uncertainty. *Physica A: Statistical Mechanics and Its Applications*, 586, 126475. <https://doi.org/10.1016/j.physa.2021.126475>
- Widarjono, A. (2018). *Analisis regresi dengan SPSS*. Yogyakarta: UPP STIM YKPN.
-