Jurnal Ekonomi Bisnis, Manajemen dan Akuntansi (Jebma)
Volume : 06 | Nomor 03 | November 2026 | E-ISSN : 2797-7161 | DOI: doi.org/jebma.v6n3.9797

Analisis *Clustering* Saham IDX 30 Menggunakan Algoritma K-*Means* Dalam Pembentukan Porotfolio Optimal Model Markowitz

Penulis:
Nurfadillah[1]
Rita Raya[2]
Nicea Roona Paranoan[3]
Yacob Ruru[4]
Caecilia Bintang Girik Allo[5]
Sitti Rosnafi'an Sumardi[6]

Afiliasi:
Universitas Cenderawasih[1,2,3,4,5]

Korespondensi:
nurfadillah240812@email.com

Histori Naskah:
Submit: 15-10-2026
Accepted: 26-10-2026
Published: 01-11-2026

Abstrak:
Penelitian ini bertujuan untuk mengelompokkan saham IDX30 berdasarkan karakteristik *return*, *Price to Book Value* (PBV), *Price to Earnings Ratio* (PER), dan *Return on Equity* (ROE) menggunakan metode K-*Means* serta membentuk portofolio optimal menggunakan Model Markowitz. Data yang digunakan terdiri atas 26 saham IDX30 yang dianalisis berdasarkan variabel *return*, PBV, PER, dan ROE. Sebelum proses *clustering*, data distandarisasi menggunakan metode Z-*score* untuk menyamakan skala antarvariabel. Penentuan jumlah *cluster* dilakukan menggunakan *Silhouette Coefficient*, sedangkan pemilihan saham kandidat pada masing-masing *cluster* dilakukan berdasarkan *Sharpe Ratio*. Selanjutnya, saham kandidat digunakan dalam pembentukan portofolio optimal menggunakan Model Markowitz dengan mempertimbangkan *expected return*, risiko, dan kovarians antar saham. Hasil penelitian menunjukkan bahwa jumlah optimal yang terbentuk adalah dua *cluster* dengan nilai *Silhouette Coefficient* sebesar 0,5251. *Cluster* 1 terdiri atas 23 saham, sedangkan *cluster* 2 terdiri atas 3 saham dan hasil *clustering* mencapai kondisi konvergen. Berdasarkan *Sharpe Ratio*, ANTM dipilih sebagai saham kandidat dari *cluster* 1 dengan nilai 0,10089, sedangkan UNVR dipilih dari *cluster* 2 dengan nilai 0,05292. Hasil optimasi Model Markowitz menghasilkan bobot portofolio sebesar 70,93% untuk ANTM dan 29,07% untuk UNVR, dengan *expected return* sebesar 0,316% dan risiko sebesar 2,699% pada setiap periode pengamatan.

Kata kunci: *Clustering*, IDX30, K-*Means*, Model Markowitz, *Silhouette Coefficient*.

Pendahuluan

Pasar modal merupakan salah satu sarana investasi yang mempertemukan pihak yang memiliki kelebihan dana (investor) dengan pihak yang membutuhkan dana (emiten) melalui perdagangan instrumen keuangan (Lusiana dkk., 2025). Saham menjadi salah satu instrumen investasi yang banyak digunakan karena memberikan peluang bagi investor untuk memperoleh *return* melalui dividen maupun *capital gain* (Febriansyah & Adlan, 2025). Peningkatan aktivitas investasi menyebabkan pertimbangan dalam pengambilan keputusan investasi tidak hanya berorientasi pada tingkat *return*, tetapi juga mencakup risiko dan hubungan antar-aset dalam penyusunan portofolio. Diversifikasi menjadi salah satu pendekatan dalam pengelolaan risiko melalui kombinasi beberapa saham yang memiliki karakteristik berbeda dalam suatu portofolio.

Pembentukan portofolio menghadapi kompleksitas yang berasal dari banyaknya alternatif saham dengan karakteristik yang berbeda, baik dari sisi tingkat *return*, risiko, kinerja keuangan, maupun valuasi. Perbedaan karakteristik tersebut menyebabkan proses pemilihan saham memerlukan pendekatan yang sistematis agar aset yang digunakan dalam pembentukan portofolio sesuai dengan karakteristik investasi yang dipertimbangkan (Griastana, 2025). Model Markowitz menentukan proporsi investasi berdasarkan sekumpulan aset yang telah ditetapkan sebagai kandidat dengan mempertimbangkan *return*, risiko, dan

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Volume : 06 | Nomor 03 | November 2026 | E-ISSN : 2797-7161 | DOI: doi.org/jebma.v6n3.9797

kovarians antar-aset. Oleh karena itu, proses seleksi saham sebelum tahap optimasi menjadi bagian penting dalam pembentukan portofolio.

Seiring perkembangan teknologi, analisis investasi kini tidak hanya dilakukan secara konvensional, tetapi juga memanfaatkan teknik *data mining* (Maulana dkk., 2025). Pendekatan yang dapat digunakan adalah *clustering*, yaitu pengelompokan objek berdasarkan kemiripan karakteristik sehingga pola dalam data dapat diidentifikasi secara sistematis (Pratama dkk., 2024). Salah satu metode yang dapat digunakan adalah K-*Means*, yaitu metode *clustering* yang mengelompokkan objek berdasarkan kedekatan terhadap *centroid* (Tohendry & Jollyta, 2023). Penelitian ini menggunakan empat variabel dalam proses pengelompokan, yaitu *return*, *Price to Earnings Ratio* (PER), *Price to Book Value* (PBV), dan *Return on Equity* (ROE). *Return* merepresentasikan tingkat pengembalian investasi, PER dan PBV merepresentasikan karakteristik valuasi saham, sedangkan ROE merepresentasikan kemampuan perusahaan dalam menghasilkan laba berdasarkan ekuitas yang dimiliki. Hasil *clustering* digunakan sebagai dasar dalam proses seleksi saham dengan *Sharpe Ratio* untuk menentukan saham kandidat dari masing-masing *cluster*. Saham kandidat selanjutnya digunakan dalam Model Markowitz untuk memperoleh proporsi investasi dengan mempertimbangkan *expected return*, risiko, dan kovarians antar-saham.

Penelitian Nasution (2025) memiliki kesamaan dengan penelitian ini dalam penggunaan metode K-*Means* untuk pengelompokan saham dan Model Markowitz untuk pembentukan portofolio optimal. Namun, terdapat perbedaan pada variabel yang digunakan, yaitu Nasution menggunakan *closing price*, PER, PBV, dan ROE, sedangkan penelitian ini menggunakan *return*, PER, PBV, dan ROE. Selain itu, penelitian ini menggunakan *Silhouette Coefficient* untuk menentukan jumlah *cluster* optimal dan *Sharpe Ratio* untuk memilih saham representatif dari setiap *cluster* sebelum dilakukan optimasi portofolio. Dengan demikian, penelitian ini mengembangkan tahapan analisis yang mengintegrasikan pengelompokan saham, pemilihan saham berdasarkan *Sharpe Ratio*, dan optimasi portofolio menggunakan Model Markowitz.

Meskipun penelitian Nasution (2025) telah mengintegrasikan metode K-*Means* dan Model Markowitz dalam pembentukan portofolio saham, masih terdapat ruang pengembangan dalam tahapan analisis yang digunakan. Perbedaan penelitian ini terletak pada penggunaan *return* sebagai variabel pengelompokan untuk merepresentasikan karakteristik pengembalian saham, sedangkan penelitian Nasution menggunakan *closing price* bersama variabel PER, PBV, dan ROE. Selain itu, penelitian ini tidak hanya melakukan pengelompokan saham, tetapi juga menerapkan *Silhouette Coefficient* untuk menentukan jumlah *cluster* optimal dan *Sharpe Ratio* untuk menyeleksi saham kandidat dari setiap *cluster* sebelum dilakukan optimasi portofolio menggunakan Model Markowitz. Integrasi tahapan tersebut menjadi fokus pengembangan penelitian ini karena menghubungkan proses identifikasi karakteristik saham, seleksi kandidat berdasarkan kinerja yang telah mempertimbangkan risiko, dan penentuan proporsi investasi berdasarkan *expected return*, risiko, serta kovarians antar-saham. Dengan demikian, kontribusi penelitian ini terletak pada penerapan tahapan analisis yang terintegrasi untuk membentuk portofolio optimal dari saham-saham IDX30.

Penggunaan Model Markowitz dalam penelitian ini didasarkan pada kemampuannya dalam menentukan kombinasi investasi dengan mempertimbangkan tingkat pengembalian dan risiko portofolio. Pendekatan tersebut memungkinkan pemilihan proporsi investasi dilakukan secara terukur dengan memperhitungkan hubungan pergerakan *return* antar-saham. Integrasi metode K-*Means* dan Model Markowitz diharapkan dapat mendukung proses pembentukan portofolio yang lebih sistematis, mulai dari pengelompokan saham hingga penentuan proporsi investasi yang sesuai dengan fungsi tujuan optimasi (Nisa'a dkk., 2024).

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk mengelompokkan 26 saham IDX30 menggunakan metode K-*Means* berdasarkan variabel *return*, PER, PBV, dan ROE, menentukan saham kandidat berdasarkan *Sharpe Ratio* dari masing-masing *cluster*, serta membentuk portofolio optimal menggunakan Model Markowitz. Penelitian ini diharapkan memberikan informasi mengenai karakteristik kelompok saham dan menghasilkan komposisi portofolio berdasarkan pertimbangan *return* dan risiko.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Volume : 06 | Nomor 03 | November 2026 | E-ISSN : 2797-7161 | DOI: doi.org/jebma.v6n3.9797

Studi Literatur

Clustering dan K-*Means*

Analisis *clustering* merupakan salah satu metode dalam analisis multivariat yang digunakan untuk mengelompokkan sejumlah objek berdasarkan tingkat kemiripan antar data. Proses *clustering* dilakukan dengan menempatkan objek yang memiliki karakteristik relatif sama ke dalam *cluster* yang sama, sedangkan objek dengan karakteristik yang berbeda ditempatkan pada *cluster* yang berbeda. Dengan demikian, analisis *clustering* dapat digunakan untuk melihat pola pengelompokan berdasarkan karakteristik dari objek yang diamati (Chatfield & Collins, 1980). Dalam proses *clustering*, tingkat kedekatan antarobjek dapat ditentukan berdasarkan ukuran jarak atau ketidakmiripan. Salah satu ukuran yang digunakan adalah jarak *Euclidean*, yang menunjukkan tingkat kedekatan antarobjek berdasarkan karakteristik yang diamati. Semakin kecil jarak antarobjek, semakin dekat karakteristik objek tersebut. Oleh karena itu, pengukuran jarak menjadi salah satu dasar dalam proses penempatan objek ke dalam *cluster* yang sama (Chatfield & Collins, 1980).

K-*Means* merupakan algoritma *clustering* paling populer dan sering digunakan, karena cara kerja dan perhitungannya mudah dilakukan. Algoritma K-*Means* merupakan salah satu metode data *clustering* non hirarki yang mempartisi data ke dalam bentuk satu atau lebih *cluster* atau kelompok. Pada *cluster,* setiap data yang mempunyai karakteristik yang mirip atau sama akan mengelompokkan ke dalam satu *cluster* dan data yang memiliki karakteristik berbeda akan dikelompokkan ke dalam *cluster* yang lain (Saputra & Kristiyanti, 2022).

Variabel Fundamental

Return merupakan keuntungan atau kerugian yang diperoleh dari investasi dalam periode tertentu. *Return* menjadi indikator utama bagi investor untuk menilai kinerja investasi dan membandingkannya dengan alternatif investasi lainnya (Harahap, 2024). *Price Earning Ratio* (PER) merupakan rasio perbandingan antara harga saham dan laba per lembar saham. PBV digunakan untuk menilai apakah harga saham berada pada kondisi *overvalued* atau *undervalued* berdasarkan nilai buku per lembar saham perusahaan (Jumairi dkk., 2025). *Return on Equity* (ROE) merupakan indikator untuk mengukur laba yang diperoleh pemegang saham atas modal yang diinvestasikan. ROE menunjukkan kemampuan perusahaan menghasilkan laba setelah pajak melalui ekuitas yang dimiliki. Semakin tinggi ROE, semakin efisien manajemen dalam mengelola modal sendiri untuk menghasilkan keuntungan (Rinto dkk., 2024). *Price to Book Value* (PBV) merupakan rasio yang membandingkan harga pasar saham dengan nilai buku perusahaan. PBV mencerminkan penilaian pasar terhadap kekayaan bersih perusahaan serta kepercayaan investor terhadap prospek dan potensi pertumbuhannya. Semakin tinggi PBV, umumnya semakin positif ekspektasi pasar terhadap kinerja perusahaan (Indriyani dkk., 2024).

Model Markowitz

Teori portofolio modern diperkenalkan oleh Harry Markowitz pada tahun 1952 dan dikenal sebagai *Mean-Variance Optimization* (MVO). Menurut Markowitz (dalam Nasution, 2025), model ini bertujuan memilih portofolio dengan return maksimum pada tingkat risiko tertentu atau risiko minimum pada tingkat return tertentu. Risiko diukur menggunakan varians atau standar deviasi, sedangkan *return* portofolio dihitung berdasarkan rata-rata tertimbang return setiap aset. Keunggulan Model Markowitz terletak pada kemampuannya menentukan kombinasi aset yang efisien dengan mempertimbangkan return, risiko, serta hubungan antar-aset melalui varians dan kovarians. Model ini juga mendukung diversifikasi melalui pemilihan aset dengan korelasi rendah atau negatif sehingga berpotensi mengurangi risiko portofolio secara keseluruhan (Harahap, 2024).

Penelitian Terdahulu

Penelitian terdahulu yang menunjukkan bahwa kombinasi variabel tersebut mampu memberikan gambaran yang lebih komprehensif terhadap karakteristik saham. Penelitian oleh Nasution (2025)

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Volume : 06 | Nomor 03 | November 2026 | E-ISSN : 2797-7161 | DOI: doi.org/jebma.v6n3.9797

menggunakan variabel *closing price*, PER, PBV, dan ROE dalam pengelompokan saham menggunakan metode K-*Means clustering*, kemudian menentukan bobot investasi optimal melalui Model Markowitz. Hasil penelitian menunjukkan bahwa pendekatan tersebut menghasilkan portofolio dengan tingkat return yang lebih optimal, risiko yang lebih rendah, serta meningkatkan efektivitas diversifikasi.

Selanjutnya, penelitian oleh Fathurachman dan Saepudin (2025) juga menunjukkan hasil yang sejalan, di mana metode K-*Means Clustering* digunakan untuk mengelompokkan saham berdasarkan *return* dan risiko, kemudian dikombinasikan dengan *Convolutional Neural Network* (CNN) untuk memprediksi *return* saham. Hasil penelitian menunjukkan bahwa pendekatan tersebut mampu menghasilkan prediksi yang akurat serta membentuk portofolio dengan kinerja yang lebih baik dibandingkan indeks LQ45, dengan rata-rata *return* sebesar 0,0009. Hal ini menunjukkan bahwa metode *clustering* tidak hanya efektif untuk pengelompokan saham, tetapi juga dapat mendukung pengambilan keputusan investasi yang lebih optimal.

Kedua penelitian tersebut menunjukkan bahwa metode *clustering* dapat dimanfaatkan dalam analisis saham, baik sebagai dasar pembentukan portofolio maupun sebagai bagian dari pendekatan prediksi *return*. Namun, penelitian Nasution (2025) menggunakan variabel *closing price*, PER, PBV, dan ROE, sedangkan penelitian ini menggunakan *return*, PER, PBV, dan ROE sebagai variabel pengelompokan. Sementara itu, Fathurachman dan Saepudin (2025) mengombinasikan K-Means dengan CNN untuk memprediksi *return* saham. Berbeda dengan kedua penelitian tersebut, penelitian ini mengintegrasikan K-Means, *Silhouette Coefficient*, dan *Sharpe Ratio* sebelum melakukan optimasi portofolio menggunakan Model Markowitz. Perbedaan tersebut terletak pada tahapan analisis dan penggunaan metode untuk menentukan saham kandidat serta proporsi investasi dalam portofolio optimal.

Metode Penelitian

Penelitian ini menggunakan data sekunder kuantitatif saham yang tergabung dalam indeks IDX30 selama periode Januari–Desember 2025. Sampel penelitian terdiri atas saham-saham yang konsisten berada dalam periode penelitian yaitu sebanyak 26 saham. Data harga penutupan (*closing price*) diperoleh dari website *Yahoo Finance*, sedangkan data *Price to Earnings Ratio* (PER), *Price to Book Value* (PBV), dan *Return on Equity* (ROE) diperoleh dari laporan keuangan perusahaan melalui situs resmi Bursa Efek Indonesia (BEI) dan sumber data keuangan yang relevan. Data yang telah diperoleh kemudian diolah dan disusun berdasarkan masing-masing saham IDX30 untuk membentuk variabel penelitian berupa *Return*, PER, PBV, dan ROE. Selanjutnya, data tersebut digunakan sebagai variabel dalam proses pengelompokkan (*clustering*) saham. Pengolahan data dilakukan menggunakan Microsoft Excel dan Google Colab. Tahapan analisis yang dilakukan dalam penelitian ini adalah sebagai berikut.

Tahapan penelitian diawali dengan analisis statistik deskriptif untuk menggambarkan karakteristik data, kemudian dilanjutkan dengan standarisasi variabel menggunakan metode *z-score normalization* untuk menyamakan skala pengukuran. Selanjutnya, pengelompokan saham dilakukan menggunakan algoritma K-Means dengan menentukan jumlah *cluster* optimal berdasarkan nilai *Silhouette Coefficient*. Proses K-Means meliputi penentuan *centroid* awal, perhitungan jarak setiap objek terhadap *centroid*, pengelompokan objek berdasarkan jarak terdekat, serta pembaruan *centroid* berdasarkan rata-rata anggota *cluster*. Proses tersebut diulang hingga mencapai kondisi konvergen. Hasil pengelompokan kemudian diinterpretasikan berdasarkan karakteristik masing-masing *cluster*, dilanjutkan dengan pemilihan saham kandidat berdasarkan nilai *Sharpe Ratio* tertinggi dari setiap *cluster*.

Tahap berikutnya adalah pembentukan portofolio optimal menggunakan Model Markowitz, yang meliputi perhitungan *expected return* dan standar deviasi masing-masing saham kandidat, penyusunan matriks kovarians antar-saham, serta penentuan bobot investasi melalui optimasi dengan tujuan memaksimalkan *Sharpe Ratio*. Optimasi dilakukan dengan kendala total bobot investasi sebesar 100% dan bobot setiap saham tidak negatif. Berdasarkan bobot optimal yang diperoleh, dilakukan perhitungan *expected return* dan risiko portofolio untuk mengevaluasi komposisi investasi yang dihasilkan.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Volume : 06 | Nomor 03 | November 2026 | E-ISSN : 2797-7161 | DOI: doi.org/jebma.v6n3.9797

Hasil

Statistik Deskriptif

Statistik deskriptif digunakan untuk menggambarkan karakteristik variabel yang digunakan dalam pengelompokan saham IDX30, yaitu return, Price to Book Value (PBV), Price Earning Ratio (PER), dan Return on Equity (ROE). Ringkasan statistik masing-masing variabel disajikan pada Tabel berikut.

Tabel 1. Analisis Deskriptif

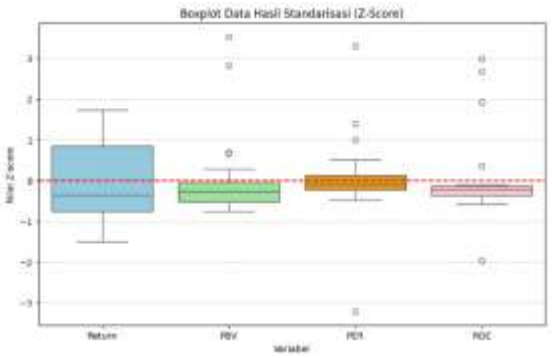
Variabel	*Mean*	Standar Deviasi	Nilai Minimum	Nilai Maksimum
Return	4,8265	23,186	-29,960	44,960
PBV	2,7471	3,5757	0,0000	15,4031
PER	0,2737	0,5343	-0,7841	1,8669
ROE	11,6075	24,7891	-68,3152	93,7478

Sumber: Hasil Pengelohan Data Google Colab (2026)

Berdasarkan hasil statistik deskriptif, menunjukkan adanya keberagaman karakteristik data pada variabel *return*, PBV, PER, dan ROE. Rentang nilai yang cukup beragam menunjukkan bahwa karakteristik perusahaan dalam sampel tidak homogen, sehingga diperlukan proses standarisasi sebelum dilakukan *clustering* menggunakan metode K-*Means*.

Standarisasi Data

Variabel yang digunakan dalam penelitian ini memiliki skala pengukuran dan rentang nilai yang berbeda. Perbedaan skala tersebut dapat memengaruhi proses pengelompokan menggunakan metode K-*Means*, hasil Z-*score* divisualisasikan dalam bentuk boxplot untuk melihat sebaran data setelah proses standardisasi, dapat dilihat pada gambar berikut.



Gambar 1. Boxplot Data Standarisasi

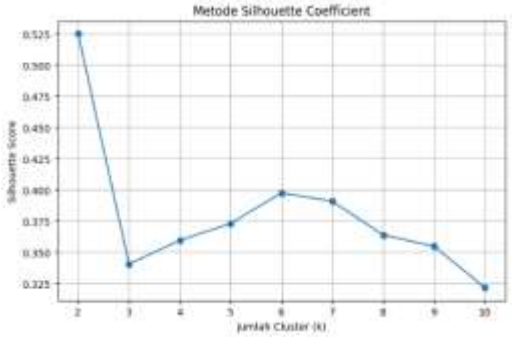
Berdasarkan Gambar 1, standardisasi menggunakan metode Z-*score* mengubah skala setiap variabel sehingga memiliki rata-rata 0. Proses ini bertujuan agar perbedaan skala antarvariabel tidak menyebabkan variabel tertentu lebih dominan dalam perhitungan jarak *Euclidean*. Boxplot menunjukkan adanya nilai ekstrem (*outlier*), terutama pada variabel PBV, PER, dan ROE. Nilai ekstrem tersebut menunjukkan bahwa beberapa saham memiliki nilai variabel yang jauh berbeda dari sebagian besar saham lainnya. Outlier tetap dipertahankan karena merupakan bagian dari data penelitian.

Penentuan Jumlah *Cluster*

Penentuan jumlah k *cluster* dilakukan dengan mengevaluasi nilai *Silhouette Coefficient* pada beberapa alternatif jumlah *cluster*, yaitu $k = 2$ sampai $k = 10$. Pada setiap nilai k, dilakukan proses K-

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

means dan perhitungan *Silhouette Coefficient* untuk masing-masing saham. Selanjutnya, nilai tersebut dirata-ratakan untuk memperoleh nilai *Silhouette Coefficient* pada setiap k. Hasil perhitungan tersebut kemudian disajikan dalam bentuk grafik sebagai berikut.



Gambar 2. Nilai *Silhouette Coefficient*

Berdasarkan Gambar 2, nilai *Silhouette Coefficient* tertinggi sebesar 0,5251 diperoleh pada $k = 2$. Nilai tersebut diperoleh dari rata-rata nilai *Silhouette Coefficient* seluruh saham. Oleh karena itu, jumlah cluster optimal yang digunakan dalam proses K-*Means* adalah 2 *cluster*.

Clustering Menggunakan Metode K-*Means*

Centroid awal ditentukan secara acak menggunakan data hasil standarisasi Z-*score*. Penentuan *centroid* bertujuan untuk menghasilkan titik awal dalam proses iterasi K-*Means*. Nilai *centroid* awal digunakan dalam penelitian ini dapat dilihat pada tabel berikut.

Tabel 2. Nilai *Centroid* Awal

Cluster	*Return*	PBV	ROE	PER
C_1	-1,4356	-0,5593	-0,1783	-0,3255
C_2	1,5899	3,5395	0,0514	2,6836

Sumber: Hasil Pengelohan Data Google Colab (2026)

Tahap selanjutnya dalam proses *clustering* adalah menghitung jarak setiap saham terhadap kedua *centroid* awal menggunakan metode *Euclidean Distance*. Setiap saham ditempatkan pada *cluster* yang memiliki nilai jarak *Euclidean* paling kecil karena dianggap memiliki karakteristik yang paling mendekati *centroid* tersebut. Hasil perhitungan jarak antara saham ke *centroid* seperti pada *tabel* berikut:

Tabel 3. Jarak Antara Data terhadap *Centoid* Pada Iterasi Pertama

Saham	d_{i1}	d_{i2}
ADRO	0,0000	5,9212
AMRT	1,4310	5,1004
ANTM	2,2861	4,7957
ASII	2,7542	5,0057
BBCA	1,0468	5,0296
BBNI	1,0230	5,4140

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Volume : 06 | Nomor 03 | November 2026 | E-ISSN : 2797-7161 | DOI: doi.org/jebma.v6n3.9797

Saham	d_{i1}	d_{i2}
BBRI	0,7639	5,7076
BMRI	0,7333	5,3489
BRPT	2,4243	4,3133
CPIN	1,0805	5,1149
GOTO	3,1994	6,2678
ICBP	0,4667	5,5220
INCO	3,5381	5,3282
INDF	0,8319	5,4687
INKP	2,2901	5,2773
ISAT	1,1884	5,1606
KLBF	0,9915	5,1421
MBMA	4,0400	3,9088
MDKA	3,6066	5,7153
MEDC	2,2238	4,8086
PGAS	2,0705	5,4797
PTBA	4,1169	2,6137
SMGR	3,5189	6,7336
TLKM	2,5264	4,9487
UNTR	1,8684	5,1295
UNVR	5,9212	0,0000

Sumber: Hasil Pengelohan Data Google Colab (2026)

Berdasarkan Tabel 3, terlihat bahwa saham ADRO memiliki jarak ke *centroid* pertama ($d_{1,1}$) 0,0000 dan jarak ke *centroid* kedua ($d_{1,2}$) sebesar 5,9212. Karena nilai jarak terkecil adalah 0,0000 maka saham ADRO masuk ke dalam *cluster* 1. Selanjutnya menghitung nilai *centroid* baru yang diperoleh dari nilai rata-rata anggota *cluster* tersebut. Hasil perhitungan *centroid* pada iterasi pertama dapat dilihat pada Tabel berikut:

Tabel 4. *Centroid* Baru pada Iterasi Pertama

Centroid	***Return***	**PBV**	**ROE**	**PER**
C_1	-0,0710	-0,2659	0,0221	-0,3301
C_2	0,5442	2,0385	-0,1695	2,5307

Sumber: Hasil Pengelohan Data Google Colab (2026)

Berdasarkan Tabel 4, diperoleh *centroid* baru yang menunjukkan karakteristik rata-rata masing-masing *cluster* berdasarkan variabel *return*, PBV, PER, dan ROE yang telah distandarisasi. *Centroid* baru tersebut selanjutnya digunakan sebagai pusat *cluster* pada iterasi berikutnya untuk menghitung kembali jarak *Euclidean* dan memastikan kestabilan keanggotaan setiap *cluster*. Hasil perhitungan jarak antara saham ke *centroid* seperti pada Tabel berikut:

Tabel 5. Jarak Antar Saham terhadap *Centroid* pada Iterasi Kedua

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Volume : 06 | Nomor 03 | November 2026 | E-ISSN : 2797-7161 | DOI: doi.org/jebma.v6n3.9797

Saham	d_{i1}	d_{i2}	*Cluster*
ADRO	1,4101	4,3389	1
AMRT	1,7915	3,7043	1
ANTM	0,9109	3,5340	1
ASII	1,4189	3,8446	1
BBCA	1,1178	3,5672	1
BBNI	0,4565	3,9225	1
BBRI	0,9970	4,1276	1
BMRI	0,7315	3,8129	1
BRPT	1,4152	3,1993	1
CPIN	0,4456	3,6509	1
GOTO	3,3107	4,9806	1
ICBP	1,3169	3,9691	1
INCO	2,2856	4,4131	1
INDF	0,6572	3,9406	1
INKP	1,0376	4,0004	1
ISAT	0,2728	3,7079	1
KLBF	0,6261	3,6754	1
MBMA	3,4708	2,3719	2
MDKA	2,4903	4,9747	1
MEDC	1,1522	3,3794	1
PGAS	0,9846	4,1507	1
PTBA	3,9005	1,6763	2
SMGR	3,4599	5,5787	1
TLKM	1,1315	3,8204	1
UNTR	0,5993	3,7611	1
UNVR	5,1306	1,8490	2

Sumber: Hasil Pengelohan Data Google Colab (2026)

Berdasarkan hasil perhitungan jarak pada iterasi kedua, terlihat bahwa tidak terjadi perubahan keanggotaan pada masing-masing *cluster.* Hal ini menunjukkan bahwa hasil *clustering* telah konvergen, sehingga anggota pada masing-masing *cluster* ditetapkan sebagai hasil akhir proses *clustering* menggunakan metode K-*means.* Anggota tiap *cluster* disajikan pada Tabel berikut.

Tabel 6. Hasil *Clustering* Menggunakan Metode K-*means*

Cluster	Banyak Anggota	Kode Saham
		ADRO, AMRT, ANTM, ASII, BBCA, BBNI, BBRI, BMRI, BRPT,

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

1	23	CPIN, GOTO, ICBP, INCO, INDF, INKP, ISAT, KLBF, MDKA, MEDC, PGAS, SMGR, TLKM, dan UNTR.
2	3	MBMA, PTBA, dan UNVR.

Sumber: Hasil Pengelohan Data Google Colab (2026)

Pemilihan Saham

Setelah proses *clustering* menggunakan metode K-*Means* selesai, tahap selanjutnya adalah menentukan saham kandidat yang akan digunakan dalam pembentukan portofolio menggunakan Model Markowitz. Pemilihan saham kandidat dilakukan pada masing-masing *cluster* dengan menggunakan *Sharpe Ratio* sebagai kriteria seleksi. *Sharpe Ratio* digunakan sebagai kriteria seleksi karena mempertimbangkan tingkat *return* terhadap risiko dengan memperhitungkan *risk-free rate*. maka, pemilihan saham tidak hanya memperhatikan besarnya *expected return*, tetapi juga risiko yang menyertainya. Nilai *Sharpe Ratio* yang lebih tinggi menunjukkan tingkat *return* yang lebih besar terhadap risiko setelah memperhitungkan *risk-free rate*. Maka didapat saham kandidat portofolio untuk masing-masing *cluster* dapat dilihat sebagai berikut.

Tabel 7. *Sharpe Ratio* pada *Cluster* 2

Kode Saham	*Expected Return*	Risiko	*Sharpe Ratio*
ANTM	0,00359	0,03346	0,10089
UNVR	0,00211	0,03573	0,05292

Sumber: Hasil Pengelohan Data Google Colab (2026)

Berdasarkan Tabel 7, saham ANTM memiliki nilai *Sharpe Ratio* tertinggi pada *cluster* 1, yaitu sebesar 0,10089. Maka, saham ANTM dipilih sebagai saham kandidat dari *cluster* 1 dalam pembentukan portofolio. Sedangkan, saham UNVR memiliki nilai *Sharpe Ratio* tertinggi pada *cluster* 2, yaitu sebesar 0,05292. Oleh karena itu, saham UNVR dipilih sebagai saham kandidat dari *cluster* 2 dalam pembentukan portofolio.

Hasil Optimasi Portofolio Model Markowitz

Berdasarkan hasil pemilihan saham kandidat diperoleh saham ANTM sebagai perwakilan *cluster* 1 dan UNVR sebagai perwakilan *cluster* 2. Perhitungan *expected return* dan risiko masing-masing saham dapat dilihat pada Tabel 7. Nilai *expected return* dan risiko saham kandidat selanjutnya digunakan sebagai nilai dalam pembentukan portofolio menggunakan Model Markowitz.

Selanjutnya dihitung Matriks kovarians untuk mengukur hubungan pergerakan *return* antar pasangan saham dalam suatu portofolio. Hasil perhitungan varians dan kovarians antar saham disajikan dalam bentuk matriks kovarians seperti pada Tabel berikut:

Tabel 8. Matriks Kovarians Antar Saham

Saham	ANTM	UNVR
ANTM	0,00112	0,00014
UNVR	0,00014	0,00127

Sumber: Hasil Pengelohan Data Google Colab (2026)

Berdasarkan Tabel 8, matriks kovarians terdiri atas saham ANTM dan UNVR. Nilai kovarians antara ANTM dan UNVR sebesar 0,00014 dan bernilai positif. Hal ini menunjukkan bahwa *return* saham ANTM dan UNVR selama periode pengamatan cenderung bergerak searah, sehingga ketika *return* salah

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Volume : 06 | Nomor 03 | November 2026 | E-ISSN : 2797-7161 | DOI: doi.org/jebma.v6n3.9797

satu saham meningkat, *return* saham lainnya cenderung ikut meningkat, dan sebaliknya. Nilai diagonal matriks sebesar 0,00112 pada saham ANTM dan 0,00127 pada saham UNVR merupakan varians masing-masing saham yang menunjukkan tingkat penyebaran *return* terhadap nilai rata-ratanya.

Penentuan bobot portofolio dilakukan untuk memperoleh proporsi dana yang dialokasikan pada masing-masing saham. Pada penelitian ini, pembobotan dilakukan menggunakan Model Markowitz melalui fungsi tujuan memaksimalkan *Sharpe Ratio*. perhitungan dilakukan dengan menggunakan nilai *expected return, risk-free rate,* dan matriks kovarians saham ANTM dan UNVR yang telah dihitung pada tahap sebelumnya. Berdasarkan komponen tersebut, bobot optimal ditentukan dengan mencari kombinasi proporsi investasi yang menghasilkan nilai *Sharpe Ratio* maksimum, dengan kendala jumlah seluruh bobot investasi sama dengan 1 atau 100% dan setiap bobot tidak bernilai negatif. hasil optimasi bobot portofolio disajikan pada Tabel berikut.

Tabel 8. Hasil optimasi Portofolio Model Markowitz

Komponen	Hasil
Bobot ANTM	70,93%
Bobot UNVR	29,07%
Expected return portofolio	0,316%
Risiko portofolio	2,699%
Sharpe Ratio portofolio	0,10909

Sumber: Hasil Pengelohan Data Google Colab (2026)

Berdasarkan Tabel 8, hasil optimasi Model Markowitz menunjukkan bahwa saham ANTM memperoleh bobot portofolio sebesar 70,93%, sedangkan saham UNVR sebesar 29,07%. Angka tersebut menunjukkan besarnya proporsi dana yang dialokasikan untuk masing-masing saham dari total dana investasi. Dengan setiap Rp1.000.000 dana yang diinvestasikan, sekitar Rp709.300 dialokasikan untuk membeli saham ANTM dan Rp290.700 untuk membeli saham UNVR. Perbedaan proporsi tersebut menunjukkan bahwa ANTM memperoleh alokasi dana lebih besar dibandingkan UNVR dalam portofolio hasil optimasi.

Pembahasan

Hasil *clustering* menggunakan metode K-Means menghasilkan dua *cluster* optimal berdasarkan nilai *Silhouette Coefficient* sebesar 0,5251, dengan 23 saham pada *cluster* 1 dan 3 saham pada *cluster* 2. Berdasarkan nilai *centroid* akhir, *cluster* 1 memiliki nilai *return*, PBV, dan PER di bawah rata-rata, sedangkan ROE mendekati rata-rata. Sebaliknya, *cluster* 2 memiliki nilai *return*, PBV, dan PER di atas rata-rata, sedangkan ROE di bawah rata-rata. Keanggotaan kedua *cluster* tidak berubah pada iterasi berikutnya, sehingga proses *clustering* dinyatakan konvergen.

Selanjutnya, saham ANTM dan UNVR dipilih sebagai kandidat portofolio berdasarkan nilai *Sharpe Ratio* tertinggi pada masing-masing *cluster*. Pendekatan ini sejalan dengan penelitian Devi dkk. (2023), yang menerapkan K-*Means* sebagai tahap pengelompokan saham sebelum pembentukan portofolio optimal menggunakan metode *Mean-Variance*. Perbedaannya, penelitian ini menggunakan variabel *return*, PER, PBV, dan ROE, serta mengintegrasikan *Silhouette Coefficient* dan *Sharpe Ratio* dalam tahapan seleksi saham sebelum optimasi portofolio.

Hasil optimasi Model Markowitz menunjukkan bahwa ANTM memperoleh bobot 70,93% dan UNVR 29,07%. Bobot ANTM yang lebih besar berkaitan dengan *expected return* yang lebih tinggi (0,00359) dan risiko individual yang lebih rendah (0,03346) dibandingkan UNVR, dengan *expected return* 0,00211 dan risiko 0,03573. Portofolio yang terbentuk menghasilkan *expected return* sebesar 0,00316, risiko sebesar 0,02699 atau 2,699% per periode, serta *Sharpe Ratio* sebesar 0,10909. Risiko portofolio tersebut lebih rendah daripada rata-rata risiko individual kedua saham, yaitu sekitar 3,460%. Hasil ini menunjukkan manfaat diversifikasi melalui penggabungan saham dengan mempertimbangkan bobot

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Volume : 06 | Nomor 03 | November 2026 | E-ISSN : 2797-7161 | DOI: doi.org/jebma.v6n3.9797

investasi, varians, dan kovarians *return*. Optimasi dilakukan dengan tujuan memaksimalkan *Sharpe Ratio*, dengan kendala total bobot sebesar 100% dan bobot setiap saham tidak negatif.

Hasil penelitian ini sejalan dengan penelitian Nasution (2025) yang mengintegrasikan K-Means dan Model Markowitz dalam pembentukan portofolio saham. Kesamaan tersebut menunjukkan bahwa pengelompokan saham dapat digunakan sebagai tahap awal sebelum penentuan proporsi investasi. Namun, penelitian ini menambahkan seleksi saham berdasarkan *Sharpe Ratio* tertinggi dari setiap *cluster*, sehingga menghasilkan ANTM dan UNVR sebagai kandidat portofolio. Portofolio yang terbentuk menghasilkan *expected return* sebesar 0,3163%, risiko sebesar 2,699%, dan *Sharpe Ratio* sebesar 0,10909. Hasil tersebut menunjukkan bahwa pendekatan yang digunakan dapat menghasilkan komposisi investasi berdasarkan pertimbangan tingkat pengembalian dan risiko, meskipun kinerjanya tetap bergantung pada karakteristik saham dan periode pengamatan yang digunakan.

Secara keseluruhan, integrasi K-*Means* dan Model Markowitz memberikan pendekatan sistematis dalam pengelompokan, seleksi, dan penentuan bobot investasi saham. Namun, hasil penelitian ini terbatas pada 26 saham IDX30 dan periode pengamatan yang digunakan, sehingga komposisi portofolio yang diperoleh belum tentu optimal pada periode berbeda. Penelitian selanjutnya dapat mempertimbangkan penambahan saham kandidat serta pengujian portofolio pada periode di luar sampel untuk mengevaluasi kestabilan dan kinerja hasil optimasi.

Kesimpulan

Berdasarkan hasil penelitian, metode K-*Means* berhasil mengelompokkan 26 saham IDX30 menjadi dua *cluster* dengan nilai *Silhouette Coefficient* sebesar 0,5251, yang menunjukkan bahwa jumlah tersebut merupakan pengelompokan terbaik dari jumlah *cluster* yang diuji. *Cluster* 1 terdiri atas 23 saham dengan karakteristik *return*, PBV, dan PER yang relatif lebih rendah serta ROE mendekati rata-rata, sedangkan *cluster* 2 terdiri atas tiga saham dengan karakteristik *return*, PBV, dan PER yang relatif lebih tinggi serta ROE lebih rendah. Berdasarkan hasil *clustering*, saham ANTM dan UNVR dipilih sebagai kandidat portofolio karena memiliki *Sharpe Ratio* tertinggi pada masing-masing *cluster*. Optimasi menggunakan Model Markowitz menghasilkan bobot investasi sebesar 70,93% pada ANTM dan 29,07% pada UNVR, dengan *expected return* sebesar 0,3163%, risiko sebesar 2,699%, dan *Sharpe Ratio* sebesar 0,10909. Hasil ini menunjukkan bahwa kombinasi metode K-*Means* dan Model Markowitz dapat digunakan untuk mengelompokkan saham berdasarkan kemiripan karakteristik serta menentukan komposisi portofolio optimal berdasarkan fungsi tujuan dan kendala optimasi yang ditetapkan.

Hasil penelitian ini dapat menjadi bahan pertimbangan bagi investor dalam menyusun portofolio saham secara lebih sistematis dengan mempertimbangkan karakteristik fundamental, tingkat pengembalian, dan risiko investasi. Namun, komposisi portofolio yang dihasilkan perlu disesuaikan dengan profil risiko, tujuan investasi, serta kondisi pasar karena hasil optimasi didasarkan pada data historis dan tidak menjamin pengembalian pada masa mendatang. Penelitian selanjutnya disarankan untuk memperluas periode pengamatan, menguji kestabilan komposisi portofolio pada periode berbeda, serta mempertimbangkan penambahan saham kandidat dan biaya transaksi agar hasil optimasi lebih representatif terhadap kondisi investasi aktual.

Referensi

Chatfield, C., & Collins, A. J. (1980). *Introduction to Multivariate Analysis*. London: Chapman and Hall.

Fathurachman, R. M., & Saepudin, D. (2025). Prediksi Return Saham Berbasis Clustering Menggunakan K-Means Clustering dan Convolutional Neural Network. *Jurnal Penelitian Informatika*, *3*(1), 33–39. https://doi.org/10.25124/logic.v3i1.9463

Febriansyah, F., & Adlan, M. A. (2025). Saham Sebagai Instrumen Investasi: Peluang, Risiko, dan Strategi Investasi Di Tengah Dinamika Ekonomi Global 2024 – 2025. *Journal Of Economis and Business*, *3*(2), 207–220.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Griastana, P. A. G. (2025). Pengambilan Keputusan Investasi Melalui Model Markowitz: Analisis Portofolio Optimal Saham Indeks IDX30. *Journal of Economic, Business and Accounting*, *8*(5), 318–327.

Harahap, K. (2024). *Manajemen Investasi dan Portofolio*. Medan: PT Media Penerbit Indonesia.

Indriyani, Lokano, R., & Dewi, E. S. (2024). Analisis Pengaruh Return on Equity terhadap Harga Saham dengan Price to Book Value sebagai Variabel Intervening pada Perusahaan yang Terdaftar di Bursa Efek Indonesia. *Prosiding Seminar Nasional Ilmu Ekonomi dan Akuntansi*, *1*(2), 365–376.

Jumairi, N., Jasmir, J., & Purnama, B. (2025). Penerapan Data Mining Dalam Mengelompokkan Saham Melalui Rasio Keuangan Pada Saham Papan Utama Bursa Efek Indonesia Menggunakan Metode K-Means Clustering. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *9*(3), 4863–4869. https://doi.org/10.36040/jati.v9i3.13670

Maulana, A., Albani, R., Susanti, E., Rahmaya, E. S., & Haris, I. A. (2025). Penerapan Algoritma K-Means Clustering Untuk Segmentasi Wilayah Berdasarkan Indikator Ekonomi di Indonesia. *RIGGS: Journal of Artificial Intelligence and Digital Business*, *4*(2), 4055–4062. https://doi.org/10.31004/riggs.v4i2.1169

Nasution, M. A. (2025). *Optimasi K-Means dan Markowitz untuk Meningkatkan Portofolio Saham pada Bursa Efek Indonesia*. (Skripsi, Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara Medan).

Nisa'a, K., Salzabila, N. A., & Amila, A. A. F. (2024). Penerapan Model Markowitz dalam Pembentukan Portofolio Optimal pada Saham Teraktif di BEI. *Lontara Journal of Mathematics, Statistics And Application*, *1*(1), 7–17

Pratama, Y., Sulistianingsih, E., Debataraja, N. N., & Imro'ah, N. (2024). K-Means clustering dan mean variance efficient portofolio dalam portofolio saham. *journal of probability and statistics*, *5*(1), 24–30. https://doi.org/10.34312/jjps.v5i1.20298

Rinto, M., Susilowati, D., & Hernadi Moorcy, N. (2024). Pengaruh Return On Equity, Earning Per Share, Dan Debt To Equity Ratio Terhadap Harga Saham Pada Perusahaan Subsektor Food And Beverage Yang Terdaftar Di Bursa Efek Indonesia. *Media Riset Ekonomi [Mr.Eko]*, *3*(1), 49–57. https://doi.org/10.36277/mreko.v3i1.374

Saputra, I., & Kristiyanti, D. A. (2022). *Machine Learning untuk Pemula* (1 ed.). Yogyakarta: PT Anak Hebat Indonesia.

Sindi Ayu Lusiana, Lukita Tripermata, & Andini Utari Putri. (2025). Analisis Fundamental dan Teknikal terhadap Return Saham pada Perusahaan Multisektor IDX 30 yang Terdaftar di BEI. *AKUA: Jurnal Akuntansi dan Keuangan*, *4*(4), 873–879. https://doi.org/10.54259/akua.v4i4.5646

Tohendry, D., & Jollyta, D. (2023). Penerapan Algoritma K-Means Clustering Untuk Pengelompokkan Saham Berdasarkan Price Earning Ratio dan Price To Book Value. *Jurnal Mahasiswa Aplikasi Teknologi Komputer dan Informasi*, *5*(1), 1–7.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License