

CLUSTERING POLYA DISTRIBUSI PENJUALAN MULTI-CHANNEL UNTUK IDENTIFIKASI SEGMENTASI KINERJA TOKO MENGGUNAKAN K-MEANS (STUDI KASUS CV PECI BATIK JOGOKARIYAN)

Rizkia Ayu Meisya Putri, Wahyu Widodo

Informatika, Sekolah Tinggi Manajemen Informatika dan Komputer El Rahma Yogyakarta
Jalan Sisingamangaraja No.76, Brontokusuman, Kec. Mergangsan, Kota Yogyakarta, D.I.Yogyakarta 55671
rizkiaayump05@gmail.com

ABSTRAK

Evaluasi kinerja distribusi merupakan aspek krusial bagi UMKM dalam memahami karakteristik setiap kanal penjualan pada sistem multi-channel. Saat ini, CV Peci Batik Jogokariyan masih melakukan evaluasi secara agregat, sehingga perbedaan pola penjualan antar toko belum teridentifikasi secara kuantitatif. Penelitian ini bertujuan untuk melakukan segmentasi kinerja toko, menentukan jumlah cluster yang optimal, serta memetakan profil dari masing-masing kelompok tersebut. Penelitian menggunakan algoritma K-Means Clustering terhadap data transaksi periode 2023–2025 yang mencakup 24 toko dan 1.220 transaksi. Setiap toko dianalisis berdasarkan tiga variabel utama, yaitu rata-rata penjualan bulanan (avg_monthly), koefisien variasi (CV) untuk mengukur stabilitas, dan persentase kontribusi terhadap total penjualan. Penentuan jumlah cluster dilakukan melalui Elbow Method dan divalidasi menggunakan Silhouette Score. Hasil penelitian berhasil mengidentifikasi tiga segmen toko, yaitu Small–Medium Stable, Medium Volatile, dan Dominant Performer. Temuan menunjukkan bahwa cluster Dominant Performer hanya terdiri dari dua toko namun memberikan kontribusi signifikan sebesar 58,06% terhadap total penjualan perusahaan. Hal ini mengindikasikan adanya konsentrasi kinerja yang tinggi pada sebagian kecil mitra distribusi. Hasil segmentasi ini memberikan dasar bagi manajemen untuk melakukan evaluasi kinerja distribusi secara lebih terstruktur dan objektif, serta mendukung perumusan strategi yang sesuai untuk setiap karakteristik toko.

Kata kunci : *clustering, distribusi penjualan, K-Means, koefisien variasi, segmentasi kinerja toko, multi-channel*

1. PENDAHULUAN

Perkembangan sistem distribusi pada era digital mendorong Usaha Mikro, Kecil, dan Menengah (UMKM) untuk mengadopsi strategi penjualan multi-channel, yaitu memanfaatkan kombinasi kanal offline dan online dalam mendistribusikan produk. Pendekatan ini memungkinkan perluasan jangkauan pasar serta peningkatan potensi penjualan, namun juga meningkatkan kompleksitas dalam pengelolaan distribusi, khususnya pada perusahaan yang memiliki banyak titik penjualan atau toko. Dalam kondisi tersebut, setiap toko memiliki karakteristik kinerja yang berbeda, baik dari sisi volume penjualan, stabilitas permintaan, maupun kontribusinya terhadap total penjualan perusahaan.

Namun demikian, evaluasi kinerja toko pada banyak UMKM masih dilakukan secara sederhana berdasarkan total penjualan, sehingga belum mampu merepresentasikan perbedaan karakteristik antar toko secara menyeluruh. Pemanfaatan data penjualan juga masih terbatas sebagai arsip administratif dan belum dimanfaatkan secara optimal untuk analisis pengambilan keputusan [1]. Selain itu, penelitian terdahulu terkait clustering pada data penjualan umumnya berfokus pada segmentasi produk, pelanggan, atau wilayah distribusi [2] [3] [4], sementara analisis yang secara khusus menempatkan toko sebagai unit distribusi dengan mempertimbangkan dimensi volume, stabilitas, dan kontribusi penjualan masih terbatas. Akibatnya, struktur kinerja toko tidak teridentifikasi secara

komprehensif. Permasalahan serupa juga terjadi pada CV Peci Batik Jogokariyan yang memiliki 24 toko aktif dengan sistem distribusi multi-channel, di mana evaluasi kinerja masih berfokus pada total penjualan tanpa mempertimbangkan stabilitas dan kontribusi relatif masing-masing toko.

Berdasarkan permasalahan tersebut, tujuan penelitian ini adalah untuk melakukan segmentasi kinerja toko pada CV Peci Batik Jogokariyan menggunakan pendekatan berbasis data, menentukan jumlah kelompok (cluster) optimal, serta menghasilkan profil segmentasi yang dapat digunakan sebagai dasar evaluasi kinerja toko secara lebih terstruktur.

Sebagai rencana penyelesaian masalah, penelitian ini mengusulkan penerapan teknik clustering menggunakan metode K-Means untuk mengelompokkan toko berdasarkan kemiripan karakteristik kinerjanya. Metode ini dipilih karena mampu mengelompokkan data numerik secara efisien dan mudah diinterpretasikan dalam konteks analisis bisnis [5]. Variabel yang digunakan meliputi rata-rata penjualan bulanan sebagai representasi volume, koefisien variasi (CV) sebagai indikator stabilitas penjualan, serta kontribusi terhadap total penjualan perusahaan. Koefisien variasi dihitung sebagai rasio standar deviasi terhadap rata-rata untuk menunjukkan tingkat fluktuasi penjualan [6]. Penentuan jumlah cluster optimal dilakukan menggunakan Elbow Method dan divalidasi dengan Silhouette Score [7]. Pendekatan ini diharapkan mampu menghasilkan

segmentasi kinerja toko yang lebih komprehensif dalam konteks distribusi multi-channel.

2. TINJAUAN PUSTAKA

2.1. K-Means Clustering

K-Means merupakan algoritma *clustering* berbasis partisi yang mengelompokkan data ke dalam K kelompok berdasarkan kedekatan jarak terhadap *centroid* secara iteratif. Algoritma ini bekerja melalui tiga tahap utama, yaitu inisialisasi *centroid*, penetapan setiap data ke *cluster* dengan *centroid* terdekat, dan pembaruan *centroid* berdasarkan rata-rata anggota *cluster*. Proses diulang hingga tidak terjadi perpindahan keanggotaan. Fungsi objektif yang diminimalkan adalah *Within-Cluster Sum of Squares* (WCSS) sebagaimana ditunjukkan pada Persamaan 1. Penerapan algoritma K-Means pada data penjualan mampu mengelompokkan data berdasarkan kemiripan karakteristik numerik sehingga membantu perusahaan menentukan strategi pengelolaan produk secara lebih terarah [4].

$$WCSS = \sum_i \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (1)$$

Di mana x adalah titik data, μ_i adalah *centroid cluster* ke- i , dan C_i adalah himpunan anggota *cluster* ke- i .

2.2. Elbow Method

Elbow Method merupakan teknik penentuan jumlah *cluster* optimal dengan mengamati nilai WCSS pada berbagai nilai K. Titik siku yang menunjukkan perlambatan penurunan WCSS merupakan nilai K yang direkomendasikan. Teknik ini telah digunakan secara luas pada penelitian segmentasi berbasis K-Means untuk data penjualan, dan terbukti menghasilkan pembagian kelompok yang bermakna secara bisnis [5].

2.3. Silhouette Score

Silhouette Score mengukur kualitas *clustering* dengan membandingkan kohesi intra-*cluster* dan separasi antar-*cluster*. Nilainya berkisar antara -1 hingga 1, di mana nilai mendekati 1 menunjukkan pengelompokan yang baik. Penggunaan *Silhouette Score* sebagai metrik validasi tambahan di luar *Elbow Method* menghasilkan segmentasi yang lebih valid secara statistik, terutama ketika kedua metode memberikan rekomendasi K yang berbeda [3].

2.4. Koefisien Variasi (CV)

Koefisien variasi (CV) merupakan ukuran dispersi relatif yang menunjukkan tingkat variabilitas data terhadap rata-ratanya. Nilai CV rendah mengindikasikan penjualan yang stabil, sedangkan nilai tinggi menunjukkan fluktuasi yang besar. Rumus CV ditunjukkan pada Persamaan 2.

$$CV = \frac{\text{Standar Deviasi}}{\text{Mean}} \quad (2)$$

2.5. Penelitian Terdahulu

K-Means telah diterapkan secara luas pada berbagai konteks segmentasi berbasis data penjualan. Pada segmentasi produk, penerapan K-Means berhasil membedakan produk berdasarkan tingkat penjualan sehingga membantu perencanaan stok dan strategi pemasaran [6] [8]. Pada segmentasi pelanggan, *clustering* berbasis data transaksi menghasilkan kelompok dengan karakteristik pembelian yang dapat dijadikan dasar strategi promosi yang lebih tertarget [7] [3]. Pada segmentasi wilayah, K-Means terbukti membagi area distribusi menjadi kategori berbeda berdasarkan intensitas transaksi [5]. Sementara itu, penerapan K-Means pada data penjualan berbasis cabang atau unit distribusi terbukti efektif dalam menghasilkan pengelompokan yang bermakna berdasarkan karakteristik volume dan pola transaksi, sehingga membantu perusahaan memahami struktur penjualan secara lebih sistematis [9]. Selain itu, metode tersebut juga menunjukkan potensi dalam menghasilkan model rekomendasi kunjungan berbasis pola transaksi historis [10]. Berbagai penelitian tersebut menunjukkan bahwa K-Means efektif menghasilkan segmentasi data yang mudah diinterpretasikan dalam konteks analisis bisnis [1].

Perbedaan penelitian ini dengan penelitian-penelitian tersebut terletak pada objek analisis dan kombinasi variabel yang digunakan. Penelitian ini menempatkan toko sebagai unit *clustering* dengan variabel *avg_monthly*, koefisien variasi (CV), dan kontribusi terhadap total penjualan. Kombinasi variabel tersebut secara spesifik merepresentasikan kinerja distribusi pada perusahaan dengan sistem penjualan *multi-channel*.

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *data mining* melalui penerapan algoritma K-Means *Clustering* untuk melakukan segmentasi kinerja toko berdasarkan pola distribusi penjualan *multi-channel*. Fokus penelitian terletak pada analisis dan pengelompokan data penjualan untuk menghasilkan *insight* berbasis data, bukan pada pengembangan sistem. Proses pengolahan dan analisis data dilakukan menggunakan bahasa pemrograman Python dengan bantuan pustaka Pandas, NumPy, dan Scikit-Learn. Analisis *clustering* dilakukan pada tingkat toko sebagai unit observasi, sehingga setiap toko direpresentasikan oleh sejumlah fitur yang dihitung dari data penjualan historis.

3.1. Data Penelitian

Data penelitian merupakan data penjualan CV Peci Batik Jogokariyan periode 2023–2025 yang tersimpan dalam tiga berkas *Spreadsheet* (LAPORAN PENJUALAN 2023.xlsx, 2024.xlsx, 2025.xlsx) pada *Google Drive*. Dataset awal terdiri dari 1.220 baris transaksi harian yang diolah menjadi representasi fitur pada tingkat toko. Objek *clustering* berjumlah 24 toko yang merupakan seluruh populasi toko aktif CV Peci

Batik Jogokariyan. Penggunaan seluruh populasi tanpa proses sampling memastikan bahwa hasil segmentasi merepresentasikan kondisi nyata perusahaan. Penelitian sebelumnya juga menunjukkan bahwa pendekatan serupa pada dataset dengan jumlah objek terbatas dapat menghasilkan segmentasi yang bermakna secara bisnis [4]. Data mencakup kanal distribusi *online* (Gawe becik 1, Gawe becik 2, Zaida, Dina, Dimas, Shopee, Didi, Tokopedia, Lazada, Bukalapak, Tiktokshop) dan *offline* (TOKO, Masjid JKR, PRO-U JKR, HB, Event, dan lainnya). Data tahun 2023 hanya tersedia selama dua bulan sehingga analisis dilakukan berdasarkan periode data yang tersedia.

3.2. Variabel Penelitian

Tiga fitur kuantitatif dihitung dari data penjualan historis untuk merepresentasikan karakteristik kinerja setiap toko. Variabel yang digunakan ditampilkan pada Tabel 1.

Tabel 1. Variabel fitur yang digunakan

Variabel	Keterangan	Formula
Avg_monthly	Rata-rata penjualan bulanan toko	$\frac{\text{Total penjualan}}{\text{jumlah bulan aktif}} \dots (3)$
CV	Koefisien variasi (stabilitas penjualan)	Persamaan (2)
Contribution	Kontribusi toko terhadap total penjualan perusahaan	$\frac{\text{Penjualan toko}}{\text{Total semua toko}} \dots (4)$

Ketiga variabel tersebut merepresentasikan dimensi volume penjualan, stabilitas fluktuasi, serta kontribusi relatif setiap toko terhadap total penjualan perusahaan.

3.3. Tahapan Penelitian

Tahapan penelitian dilakukan secara sistematis sebagaimana ditunjukkan pada Gambar 1, berdasarkan Gambar 1, penelitian dilakukan melalui beberapa tahapan yang dimulai dari pengolahan data hingga penerapan algoritma *clustering*. Tahap pertama adalah konsolidasi dan pembersihan data, di mana data dari tiga berkas *spreadsheet* dikonsolidasikan menjadi satu dataset terintegrasi. Proses *data cleaning* meliputi penghapusan baris agregat (TOTAL), penanganan nilai kosong pada kolom tanggal, konversi format tanggal, transformasi data ke dalam *long format*, serta penyeragaman nama toko melalui proses *mapping*. Setelah pembersihan data, diperoleh dataset akhir sebanyak 1.220 baris data yang merepresentasikan 24 toko.

Tahap berikutnya adalah pembentukan fitur, yaitu perhitungan variabel *avg_monthly* sebagai rata-rata penjualan bulanan per toko menggunakan Persamaan (3), perhitungan koefisien variasi (CV) menggunakan Persamaan (2) untuk merepresentasikan stabilitas penjualan, serta variabel *contribution*

menggunakan Persamaan (4) sebagai rasio total penjualan setiap toko terhadap total penjualan seluruh toko.

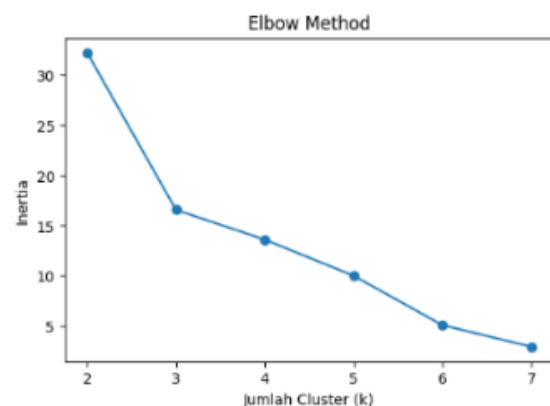


Gambar 1. Alur tahapan penelitian

Selanjutnya dilakukan standarisasi data dengan *StandardScaler* untuk menyeimbangkan skala antar variabel sehingga tidak terjadi dominasi fitur tertentu dalam proses *clustering*. Penentuan jumlah *cluster* optimal dilakukan menggunakan *Elbow Method* pada rentang nilai $K=2$ hingga $K=7$ dan divalidasi menggunakan *Silhouette Score*. Tahap terakhir adalah penerapan algoritma K-Means pada data yang telah distandarisasi menggunakan pustaka *Scikit-Learn* pada lingkungan komputasi *Google Colab*, dengan parameter *random_state = 42* untuk memastikan hasil *clustering* dapat direproduksi.

4. HASIL DAN PEMBAHASAN

4.1. Penentuan Jumlah Cluster Optimal



Gambar 2. Grafik Elbow Method penentuan nilai K optimal

Analisis *clustering* menghasilkan tiga kelompok toko dengan karakteristik penjualan yang berbeda, yang kemudian dianalisis berdasarkan nilai rata-rata fitur dan komposisi anggotanya. *Elbow Method* diterapkan pada rentang K=2 hingga K=7. Grafik pada Gambar 2 menunjukkan penurunan nilai *inertia* yang signifikan pada K=2 ke K=3, kemudian melambat setelah K=3. Titik siku yang jelas teridentifikasi pada K=3. Hasil *Silhouette Score* untuk berbagai nilai K ditampilkan pada Tabel 2.

Tabel 2. Silhouette Score pada berbagai nilai K

Nilai K	Silhouette Score
2	0,6832
3	0,5008
4	0,4616
5	0,4350
6	0,4816
7	0,4969

Dari Tabel 2 dapat dilihat bahwa *Silhouette Score* tertinggi diperoleh pada K=2 (0,6832), sedangkan K=3 menghasilkan skor 0,5008. Meski secara statistik K=2 lebih optimal, pemilihan K=3 didasarkan pada pertimbangan informatif-bisnis yang lebih kuat. Pada K=2, seluruh 22 toko di luar dua toko dominan akan digabung menjadi satu kelompok tunggal yang heterogen, yaitu mencampur toko dengan penjualan stabil namun kecil bersama toko dengan fluktuasi tinggi. Penggabungan tersebut menghilangkan perbedaan karakteristik antar toko yang penting untuk keperluan evaluasi kinerja distribusi. Sebaliknya, K=3 berhasil memisahkan toko berfluktuasi tinggi (*Cluster 1*) dari toko stabil-kecil (*Cluster 0*), menghasilkan tiga kelompok dengan profil yang dapat dibedakan secara kuantitatif dan bermakna untuk penyusunan strategi distribusi. Pendekatan memilih K berdasarkan relevansi bisnis di atas nilai statistik semata juga sejalan dengan praktik yang direkomendasikan dalam penelitian segmentasi berbasis *clustering* [7] [1].

4.2. Profil Cluster

Rata-rata nilai tiap fitur per *cluster* setelah proses K-Means ditampilkan pada Tabel 3.

Tabel 3. Profil rata-rata tiap *cluster*

Cluster	Label	Avg_ monthly	CV	Contri- bution
0	Small-Medium Stable	64,37	0,492	2,49%
1	Medium Volatile	32,11	1,173	1,33%
2	Dominant Performer	558,88	0,811	29,06%

Dari Tabel 3 dapat dilihat perbedaan karakteristik yang jelas antar *cluster*. *Cluster 2 (Dominant Performer)* memiliki *avg_monthly* sebesar 558,88 unit dan *contribution* rata-rata 29,06%, jauh melampaui *Cluster 0* dan *Cluster 1 (Medium Volatile)*

menunjukkan nilai CV tertinggi (1,173) yang mengindikasikan fluktuasi penjualan yang signifikan. *Cluster 0 (Small-Medium Stable)* memiliki nilai CV terendah (0,492) yang menunjukkan pola penjualan paling stabil meski volumenya kecil.

4.3. Anggota Tiap Cluster

Selain tiga variabel utama *clustering*, tabel berikut juga menyertakan *std_rev* yaitu standar deviasi penjualan bulanan setiap toko sebagai informasi pendukung untuk memahami tingkat fluktuasi penjualan. Rincian anggota masing-masing *cluster* beserta nilai tiap fiturnya disajikan pada Tabel 4, 5, dan 6.

Tabel 4. Anggota *cluster 0-small-medium stable*

Nama Toko	Avg_ monthly	Std_ rev	CV	Cont.
Zaida	198,04	136,48	0,689	0,103
TOKO	182,19	77,80	0,427	0,095
HB	166,50	63,13	0,379	0,013
Pro-U JKR	110,42	50,99	0,462	0,057
Kresno	20,00	N/A	0,000	0,000
Khazanah	10,83	6,45	0,595	0,003
Rafa muslim	7,00	2,83	0,404	0,000
Dimas	4,33	3,08	0,710	0,001
Pasaraya	4,00	2,05	0,514	0,001
Bukalapak	2,75	1,50	0,545	0,000
Adikusmo	2,14	1,46	0,683	0,000

Dari Tabel 4 dapat dilihat bahwa *Cluster 0* beranggotakan 11 toko dengan *avg_monthly* berkisar antara 2,14 hingga 198,04 unit dan nilai CV rata-rata 0,492. Toko Zaida, TOKO, dan PRO-U JKR menempati posisi teratas pada kelompok ini dengan volume penjualan yang lebih tinggi namun masih jauh di bawah *Cluster 2*. Nilai standar deviasi pada toko Kresno tidak tersedia karena data penjualan hanya tercatat pada satu periode sehingga variasi tidak dapat dihitung. Hal ini menunjukkan bahwa kelompok tersebut memiliki pola penjualan yang relatif stabil meskipun kontribusinya terhadap total penjualan perusahaan masih terbatas.

Tabel 5. Anggota *cluster 1 – medium volatile*

Nama Toko	Avg_ monthly	Std_ rev	CV	Cont.
Gawe becik 2	176,00	324,76	1,845	0,091
Event	49,69	59,05	1,188	0,016
Tiktokshop	25,86	23,96	0,927	0,007
Shopee	22,82	27,36	1,199	0,010
J.S Malioboro	15,11	15,19	1,005	0,003
Pro-U kg	15,00	15,56	1,037	0,001
PKG YIA	11,95	16,44	1,376	0,005
Dina	11,52	12,40	1,076	0,005
Tokopedia	11,00	13,38	1,216	0,003
Lazada	9,40	9,28	0,987	0,004
Didi	4,82	5,05	1,048	0,002

Dari Tabel 5 dapat dilihat bahwa *Cluster 1* beranggotakan 11 toko dengan nilai CV rata-rata 1,173 yang mengindikasikan fluktuasi penjualan yang tinggi.

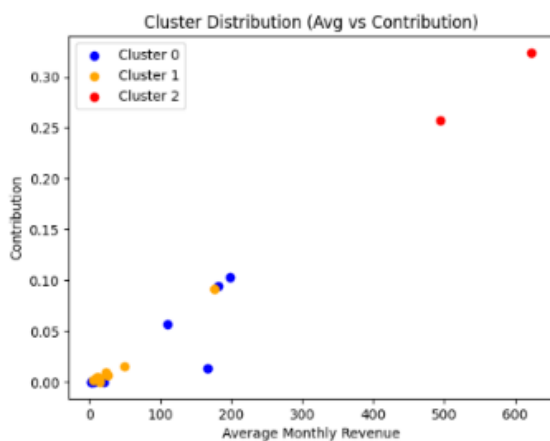
Gawe becik 2 mendominasi kelompok ini dengan *avg_monthly* 176,00 unit namun memiliki CV tertinggi (1,845), menunjukkan variasi penjualan yang sangat besar. Toko-toko *online* seperti Shopee, Tiktokshop, dan Tokopedia juga masuk dalam kelompok ini, yang mengindikasikan bahwa kanal *online* secara umum memiliki pola penjualan yang lebih fluktuatif.

Tabel 6. Anggota *cluster 2 – dominant performer*

Nama Toko	Avg_monthly	Std_rev	CV	Cont.
Gawe becik 1	623,15	777,72	1,248	0,324
Masjid JKR	494,62	185,16	0,374	0,257

Dari Tabel 6 dapat dilihat bahwa *Cluster 2* beranggotakan 2 toko, yaitu Gawe becik 1 dan Masjid JKR, yang menyumbang kontribusi gabungan sebesar 58,06% dari total penjualan perusahaan (masing-masing 32,37% dan 25,69%). Temuan ini menunjukkan bahwa hanya 2 dari 24 toko (8,3%) menyumbang lebih dari separuh total penjualan perusahaan, yang mengindikasikan adanya konsentrasi penjualan yang tinggi pada sebagian kecil toko. Struktur tersebut berhasil diidentifikasi oleh algoritma K-Means berdasarkan kedekatan jarak Euclidean pada ruang fitur yang telah distandarisasi. Gawe becik 1 memiliki *avg_monthly* tertinggi di seluruh dataset (623,15 unit/bulan), sedangkan Masjid JKR menonjol dengan nilai CV terendah di seluruh dataset (0,374) sekaligus *avg_monthly* sebesar 494,62 unit per bulan. Kombinasi volume penjualan yang tinggi dan stabilitas yang baik tersebut tidak ditemukan pada toko lain dalam dataset. Secara geometris, posisi kedua toko ini berada cukup jauh dari toko lainnya dalam ruang fitur terstandarisasi, sehingga jarak Euclidean terhadap *cluster* lain lebih besar dibandingkan jarak antar toko dalam *Cluster 0* dan *Cluster 1*. Kondisi tersebut menjelaskan mengapa kedua toko ini membentuk kelompok tersendiri dalam hasil *clustering*.

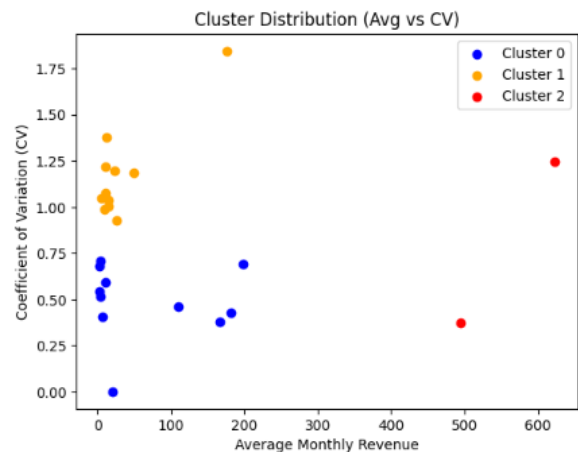
4.4. Visualisasi Distribusi Cluster



Gambar 3. Distribusi *cluster* berdasarkan *avg monthly* dan *contribution*

Untuk memperjelas pemisahan antar *cluster* secara visual, penelitian ini menyajikan dua grafik distribusi yang merepresentasikan kombinasi fitur yang berbeda sebagaimana ditunjukkan pada Gambar 3 dan Gambar 4.

Gambar 3 menampilkan visualisasi distribusi *cluster* berdasarkan dimensi *avg_monthly* dan *contribution*. Pada grafik tersebut terlihat pemisahan yang jelas antara *Cluster 2* (merah) yang berada pada area kanan atas grafik dibandingkan *Cluster 0* (biru) dan *Cluster 1* (oranye) yang terkonsentrasi di area kiri bawah dengan skala yang jauh lebih kecil.



Gambar 4. Distribusi *cluster* berdasarkan *avg monthly* dan cv

Gambar 4 menampilkan visualisasi distribusi *cluster* berdasarkan dimensi *avg_monthly* dan koefisien variasi (CV). Pada grafik ini terlihat bahwa *Cluster 1* (oranye) memiliki nilai CV yang jauh lebih tinggi dibandingkan *Cluster 0* (biru), meskipun rentang *avg_monthly* keduanya relatif berdekatan. Kondisi ini menunjukkan bahwa CV menjadi dimensi pembedaa utama antara *Cluster 0* dan *Cluster 1*, sehingga pemisahan kedua kelompok tersebut dapat dijelaskan oleh perbedaan tingkat fluktuasi penjualan.

Hasil tersebut menunjukkan bahwa kombinasi variabel *avg_monthly*, CV, dan *contribution* mampu memisahkan karakteristik toko secara jelas, terutama dalam membedakan stabilitas dan dominasi penjualan yang tidak dapat diidentifikasi hanya menggunakan satu dimensi saja.

4.5. Implikasi Manajerial

Hasil segmentasi memberikan gambaran kuantitatif yang dapat mendukung pengambilan keputusan strategis CV Peci Batik Jogokariyan. *Cluster 2* (*Dominant Performer*) dapat dipertimbangkan sebagai prioritas dalam alokasi sumber daya distribusi dan perhatian strategis karena Gawe becik 1 dan Masjid JKR secara bersama-sama menyumbang lebih dari separuh total penjualan perusahaan. Prioritas alokasi sumber daya pada kelompok dengan kontribusi dominan juga sejalan dengan hasil segmentasi berbasis *clustering* pada

penelitian terdahulu [9] [4]. Kestabilan penjualan Masjid JKR (CV=0,374) menjadikannya sebagai toko andalan yang dapat diandalkan sebagai basis distribusi yang konsisten.

Toko dalam *Cluster 1 (Medium Volatile)* memerlukan evaluasi lebih lanjut terutama Gawe becik 2 yang meskipun memiliki volume penjualan tinggi (176 unit/bulan), menunjukkan nilai CV yang sangat tinggi (1,845). Upaya stabilisasi seperti pengelolaan stok yang lebih terencana atau strategi promosi yang lebih konsisten dapat membantu menurunkan fluktuasi penjualan. Hal tersebut sejalan dengan rekomendasi penanganan kelompok *volatile* dalam segmentasi kinerja distribusi [5]. Toko-toko *online* dalam kelompok ini juga memerlukan strategi khusus untuk menstabilkan penjualan di platform digital.

Sementara itu, *Cluster 0 (Small-Medium Stable)* memerlukan strategi optimalisasi volume. Meskipun penjualannya relatif stabil, kontribusi kumulatif kelompok ini terhadap total perusahaan masih relatif kecil. Strategi pengembangan kapasitas atau penguatan pada toko-toko dengan potensi pertumbuhan seperti Zaida dan TOKO dapat menjadi prioritas, sebagaimana pendekatan serupa telah diterapkan pada segmentasi toko dengan karakteristik stabil namun bervolume rendah [6] [8].

4.6. Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan dalam menginterpretasikan hasil. Pertama, jumlah objek *clustering* terbatas pada 24 toko yang merupakan keseluruhan populasi toko aktif perusahaan, sehingga generalisasi hasil ke konteks perusahaan lain perlu dilakukan dengan hati-hati. Kedua, data penjualan tahun 2023 hanya mencakup dua bulan sehingga sebagian pola penjualan belum sepenuhnya merepresentasikan kondisi tahunan secara lengkap. Ketiga, variabel yang digunakan dalam proses *clustering* terbatas pada tiga fitur kuantitatif, yaitu rata-rata penjualan bulanan, koefisien variasi, dan kontribusi penjualan, sehingga faktor eksternal seperti lokasi toko, strategi promosi, maupun pengaruh musiman belum dipertimbangkan dalam analisis.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengelompokkan 24 toko CV Peci Batik Jogokariyan menggunakan metode K-Means berdasarkan tiga variabel, yaitu *avg_monthly*, koefisien variasi (CV), dan *contribution*. Hasil penentuan jumlah cluster menunjukkan bahwa meskipun nilai Silhouette Score tertinggi diperoleh pada K=2, pemilihan K=3 menghasilkan segmentasi yang lebih representatif. Pada K=2, sebagian besar toko tergabung dalam satu kelompok heterogen, sedangkan K=3 mampu memisahkan karakteristik toko berdasarkan stabilitas dan volume penjualan secara lebih jelas. Hasil *clustering* menghasilkan tiga kelompok, yaitu Small-Medium Stable, Medium Volatile, dan Dominant Performer. Cluster Dominant

Performer yang terdiri dari dua toko memberikan kontribusi sebesar 58,06% terhadap total penjualan. Temuan ini menunjukkan adanya konsentrasi penjualan pada sebagian kecil toko. Selain itu, perbedaan nilai *avg_monthly*, CV, dan kontribusi pada masing-masing cluster menunjukkan bahwa ketiga variabel tersebut efektif dalam merepresentasikan variasi kinerja toko.

Untuk menyempurnakan hasil analisis dan memperluas cakupan temuan pada penelitian berikutnya, disarankan beberapa pengembangan seperti integrasi variabel multichannel dengan menambahkan variabel proporsi penjualan berdasarkan tipe kanal (*online vs offline*), serta indikator pertumbuhan penjualan antar periode untuk memperkaya analisis. Selain itu, penggunaan metode *clustering* alternatif seperti K-Medoids atau Hierarchical Clustering dapat dilakukan untuk menguji konsistensi hasil segmentasi.

DAFTAR PUSTAKA

- [1] T. Wahyudi, N. Saadah, and D. Puspitasari, "Penerapan Metode K-Means Pada Data Penjualan Untuk Mendapatkan Produk Terlaris Di PT. Titian Nusantara Boga | Jurnal Sains dan Teknologi," *Jurnal Sains dan Teknologi*, vol. 5, no. 1, pp. 228–236, Aug. 2023, doi: 10.55338/saintek.v5i1.1379.
- [2] Z. Vadissa Arla, T. Sutabri, and J. Jendral Ahmad Yani No, "Analisis Produk Terlaris Menggunakan Metode K-Means Clustering Pada 'Toko Hartati,'" *Uranus: Jurnal Ilmiah Teknik Elektro, Sains dan Informatika*, vol. 2, no. 4, pp. 231–235, Nov. 2024, doi: 10.61132/uranus.v2i4.514.
- [3] Ermala Handayani and Nirwana Hendrastuty, "Klasterisasi Wilayah Penjualan Pada Toko Roni Menggunakan K-Means," *INFORMASI (Jurnal Informatika dan Sistem Informasi)*, vol. 17, no. 1, pp. 163–180, Jul. 2025, doi: 10.37424/informasi.v17i1.392.
- [4] A. M. Artiarno, P. Setiaji, and F. Nugraha, "K-Means Clustering untuk Segmentasi Pelanggan: Mengungkap Pola Pembelian Strategi Pemasaran pada Sektor Ritel," *Edumatic: Jurnal Pendidikan Informatika*, vol. 9, no. 2, pp. 442–451, Aug. 2025, doi: 10.29408/edumatic.v9i2.30336.
- [5] A. B. Almagribi, A. Bilal Almagribi, and S. Redjeki, "Clustering and Classification of Retail Sales Data: A Big Data and Data Mining Analysis," *Journal Innovations Computer Science*, vol. 4, no. 2, pp. 242–253, Nov. 2025, doi: 10.56347/jics.v4i2.303.
- [6] S. R. Sifa, Shofa Shofiah Hilabi, Bayu Priyatna, and Agustia Hananto, "Pengelompokan Penjualan Produk Dengan Menggunakan K-Means Clustering: Studi Kasus Analisis Penjualan Coffee Shop Oleh Kaggle.Com," *Jurnal Sistem Informasi (JUSIN)*, vol. 6, no. 1,

- pp. 1–10, Jun. 2025, doi: 10.32546/jusin.v6i1.3078.
- [7] D. A. Awaliyah, B. Prasetyo, R. Muzayanah, and A. D. Lestari, “Optimizing Customer Segmentation in Online Retail Transactions through the Implementation of the K-Means Clustering Algorithm,” *Scientific Journal of Informatics*, vol. 11, no. 2, pp. 539–548, Jun. 2024, doi: 10.15294/sji.v11i2.6137.
- [8] R. I. Manarung, E. Widodo, and A. M. Rifai, “Sales Data Clustering Using the K-Means Algorithm to Determine Retail Product Needs,” *International Journal Software Engineering and Computer Science (IJSECS)*, vol. 5, no. 1, pp. 226–234, Apr. 2025, doi: 10.35870/ijsecs.v5i1.4090.
- [9] V. Alvianatinova, I. Ali, N. Rahaningsih, and A. Bahtiar, “Penerapan Algoritma K-Means Clustering Dalam Pengelompokan Data Penjualan Supermarket Berdasarkan Cabang (Branch),” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 2, pp. 1529–1535, Apr. 2024, doi: 10.36040/jati.v8i2.8993.
- [10] B. Kristomoyo Kristanto, S. Widyayuningtias, P. Listio, M. Amien, and P. Iman Baskoro, “Visit Recommendation Model: Recursive K-Means Clustering Analysis of Retail Sales Data,” *Journal of Applied Informatics and Computing*, vol. 8, no. 1, pp. 221–225, Jul. 2024, doi: 10.30871/jaic.v8i1.8138