Jurnal Publikasi Teknik Informatika
Volume 5 Nomor 3, September 2026
e-ISSN: 2808-8972 p-ISSN: 2808-9367, Hal. 41-53
DOI: https://doi.org/10.55606/jupti.v5i3.7945
Tersedia: https://journalcenter.org/index.php/jupti

Pengelompokan Kinerja Supir Menggunakan Metode K-Means Clustering Berdasarkan Data Historis Jumlah Pengiriman

Andre Hartanto
Teknik Informatika, Universitas Katolik Darma Cendika, Indonesia

Penulis Korespondensi: andre.hartanto@ukdc.ac.id

Abstract. *Driver performance assessment is a critical aspect of transportation company operations; however, manual evaluation tends to be inefficient and prone to subjectivity. This study aims to classify driver performance at a transportation company in Surabaya using the K-Means Clustering algorithm based on historical delivery data from 2020 to 2025. The dataset comprises 31 drivers with 135 data records. The research workflow includes data collection, preprocessing through aggregation and Min-Max Scaling normalization, optimal cluster determination using the Elbow Method, clustering process, and evaluation using Silhouette Score and Davies-Bouldin Index. Results indicate that k=3 is the optimal number of clusters, categorizing drivers into three performance groups: High (7 drivers, 22.6%), Moderate (15 drivers, 48.4%), and Low (9 drivers, 29.0%). A Silhouette Score of 0.5948 and Davies-Bouldin Index of 0.5191 confirm that the clustering results are valid and of good quality. These findings are expected to serve as an objective, data-driven foundation for management in driver performance evaluation and development.*

Keywords: *Davies-Bouldin Index; Driver Performance; Elbow Method; K-Means Clustering; Silhouette Score.*

Abstrak. Penilaian kinerja supir menjadi elemen penting dalam operasional perusahaan transportasi, namun evaluasi secara manual dinilai kurang efisien dan berpotensi subjektif. Penelitian ini bertujuan mengelompokkan kinerja supir pada sebuah perusahaan transportasi di Surabaya menggunakan algoritma K-Means Clustering berbasis data historis jumlah pengiriman periode 2020–2025. Dataset terdiri dari 31 supir dengan 135 record data. Alur penelitian mencakup pengumpulan data, preprocessing melalui agregasi dan normalisasi Min-Max Scaling, penentuan jumlah cluster optimal menggunakan Elbow Method, proses clustering, serta evaluasi menggunakan Silhouette Score dan Davies-Bouldin Index. Hasil pengujian menunjukkan k=3 sebagai jumlah cluster optimal, yang membagi supir ke dalam tiga kategori kinerja yaitu Tinggi (7 supir, 22,6%), Sedang (15 supir, 48,4%), dan Rendah (9 supir, 29,0%). Nilai Silhouette Score sebesar 0,5948 dan Davies-Bouldin Index sebesar 0,5191 mengindikasikan hasil pengelompokan yang valid dan berkualitas baik. Temuan ini diharapkan menjadi landasan pengambilan keputusan yang objektif bagi manajemen dalam evaluasi dan pembinaan kinerja supir.

Kata Kunci: Elbow Method; Davies-Bouldin Index; K-Means Clustering; Kinerja Supir; Silhouette Score.

1. LATAR BELAKANG

Kemajuan di bidang informatika telah memberikan dampak yang signifikan terhadap berbagai sektor kehidupan, termasuk dunia bisnis dan industri. Salah satu sektor yang paling merasakan dampak tersebut adalah industri transportasi dan logistik, yang kini menjelma menjadi sektor strategis dalam menggerakkan roda perekonomian nasional. Seiring dengan pertumbuhan pasar logistik Indonesia yang terus meningkat dari tahun ke tahun, perusahaan transportasi dituntut untuk mengelola sumber daya operasionalnya secara lebih efisien dan berbasis data (Mauludin et al., 2025; Nurjanah, 2025).

Dalam ekosistem perusahaan transportasi dan logistik, supir memegang peran yang tidak bisa dipandang sebelah mata. Mereka adalah garda terdepan layanan pengiriman yang secara langsung menentukan kelancaran distribusi barang, ketepatan waktu, dan kepuasan pelanggan. Di sisi lain, rendahnya kinerja supir berpotensi memicu keterlambatan pengiriman

Naskah Masuk: 30 Juni 2026; Revisi: 07 Agustus 2026; Diterima: 28 Agustus 2026; Terbit: 01 September 2026

sekaligus membengkaknya biaya operasional. Kondisi ini sejalan dengan tuntutan nyata di lapangan, di mana perusahaan logistik tidak hanya dituntut mengangkut barang, tetapi juga berperan sebagai mitra terpercaya bagi pelanggan yang pada akhirnya berdampak pada loyalitas dan kebiasaan pembelian ulang (Anikah et al., 2022; Olivia, 2023).

Salah satu persoalan yang kerap dihadapi perusahaan di sektor ini adalah sulitnya melakukan penilaian kinerja supir secara objektif, khususnya ketika data yang tersedia mencakup rentang waktu panjang dan melibatkan banyak supir sekaligus. Penilaian yang masih dilakukan secara manual cenderung membutuhkan waktu yang tidak sedikit, mudah dipengaruhi unsur subjektivitas, dan kurang efisien. Oleh karena itu, dibutuhkan pendekatan berbasis data yang dapat mengelompokkan supir secara otomatis berdasarkan pola kinerjanya (Mauludin et al., 2025).

K-Means Clustering merupakan salah satu algoritma *data mining* yang mengandalkan teknik *unsupervised learning* dalam prosesnya (Anikah et al., 2022). Algoritma ini bekerja dengan cara menghitung kedekatan setiap data terhadap pusat *cluster* (*centroid*) dan mengalokasikannya ke *cluster* terdekat secara iteratif hingga diperoleh hasil yang optimal (Komputer et al., 2025). Penerapan *K-Means* sendiri telah banyak terbukti efektif dalam berbagai penelitian yang berkaitan dengan pengelompokan kinerja karyawan maupun analisis produktivitas tenaga kerja (Clarissa et al., 2025; Ramadhani & Athalina, 2025).

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengelompokkan kinerja supir pada sebuah perusahaan transportasi dan logistik di Surabaya menggunakan algoritma *K-Means Clustering*, dengan memanfaatkan data historis jumlah pengiriman dari tahun 2020 hingga 2025. Hasil pengelompokan diklasifikasikan ke dalam tiga kategori kinerja yaitu Tinggi, Sedang, dan Rendah. Penentuan jumlah *cluster* optimal dilakukan melalui *Elbow Method*, sedangkan validasi kualitas *cluster* menggunakan *Silhouette Score* dan *Davies-Bouldin Index* untuk memastikan hasil pengelompokan yang valid dan dapat dipertanggungjawabkan

2. KAJIAN TEORITIS

K-Means Clustering adalah algoritma data mining yang mengelompokkan data secara otomatis tanpa supervisi (*unsupervised learning*) menggunakan sistem partisi. Algoritma ini mengelompokkan data ke dalam beberapa *cluster*, di mana data dalam satu *cluster* memiliki karakteristik yang serupa satu sama lain, namun berbeda dengan data yang berada di *cluster* lain (No et al., 2019). Tahapan algoritma K-Means sebagai berikut (No et al., 2019):

a. Menentukan jumlah *cluster* k.
b. Menginisialisasi pusat *cluster* (*centroid*) secara acak.
c. Menghitung jarak setiap data ke masing-masing *centroid* menggunakan jarak *Euclidean* dengan rumus:

$$D(i,j) = \sqrt{\sum_k (X_{ik} - C_{jk})^2} \quad \text{.....2)}$$

Di mana D(i,j) adalah jarak data ke-i ke pusat *cluster* ke-j, X adalah nilai data, dan C adalah nilai *centroid*.

d. Mengelompokkan setiap data ke *cluster* dengan *centroid* terdekat.
e. Memperbarui posisi *centroid berdasarkan* rata-rata seluruh anggota cluster.
f. Mengulangi langkah 3–5 hingga posisi *centroid* tidak berubah lagi.

Elbow Method

Elbow Method adalah salah satu metode yang digunakan untuk menentukan jumlah *cluster* optimal dengan cara melihat persentase hasil perhitungan setiap *cluster* yang akan membentuk pola menyerupai siku pada titik tertentu. Metode ini biasanya disajikan dalam bentuk grafik agar pola siku yang terbentuk dapat terlihat lebih jelas. ujuan utama *Elbow Method* adalah mencari nilai k terkecil yang masih menghasilkan nilai withinss rendah. Untuk mengevaluasi jumlah cluster k ini, digunakan teknik SSE (*Sum of Squared Error*) yang mengukur jarak atau selisih antara data asli dengan pusat cluster terkait (Maori, 2023).

Silhouette Score

Silhouette score adalah metode evaluasi yang digunakan untuk mengukur kualitas hasil proses clustering berdasarkan tingkat kesamaan setiap data dengan cluster tempatnya berada serta perbedaannya terhadap cluster lain. Nilai *silhouette* dihitung untuk setiap titik data sehingga dapat menggambarkan seberapa tepat data tersebut ditempatkan dalam clusternya. Pengujian dilakukan dengan membandingkan nilai silhouette score pada beberapa jumlah *cluster*. Nilai *silhouette score* yang lebih tinggi menunjukkan kualitas *clustering* yang lebih baik, sedangkan nilai yang rendah mengindikasikan bahwa pemisahan antar *cluster* belum optimal. Dengan demikian, *Silhouette Score* berperan penting dalam membantu menentukan jumlah *cluster* yang optimal untuk suatu dataset (Hendrastuty, 2024).

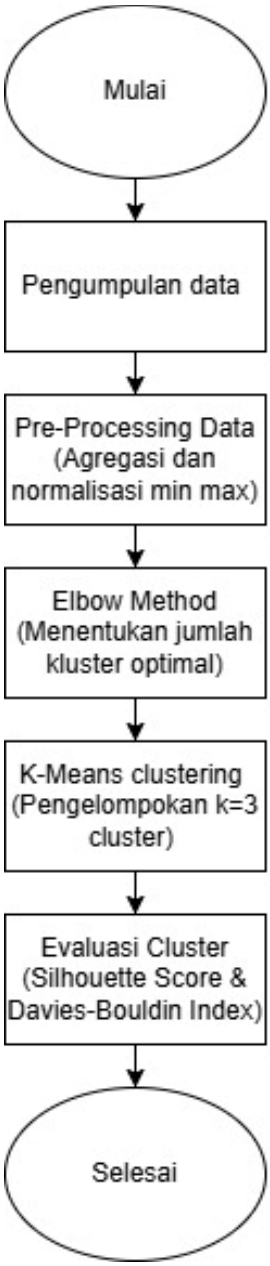
Davies-Bouldin Index (DBI)

Davies-Bouldin Index (DBI) adalah metrik evaluasi yang digunakan untuk mengukur kualitas hasil *clustering*. DBI mengukur seberapa jauh suatu cluster terpisah dari *cluster* lain berdasarkan varian data dan jarak antar pusat *cluster*. Nilai DBI dihitung sebagai rata-rata kemiripan setiap cluster dengan cluster lain yang paling mirip dengannya. Jumlah *cluster* yang

menghasilkan nilai DBI paling minimal dianggap sebagai jumlah cluster yang optimal. Semakin rendah nilai DBI yang dihasilkan, maka semakin baik kualitas cluster yang terbentuk (Menggunakan & Autoencoder, 2025).

3. METODE PENELITIAN

Tahapan pada penelitian ini dirancang secara sistematis mulai dari pengumpulan data hingga visualisasi hasil, sebagaimana ditunjukkan pada Gambar 1. Alur Penelitian.



Gambar 1. Alur Penelitian.

Pengumpulan Data

Dalam penelitian ini diperoleh dari sebuah perusahaan transportasi yang berlokasi di Surabaya, data yang dikumpulkan berisi historis jumlah pengiriman barang per supir. Data ini mencakup periode tahun 2020 hingga 2025 dengan total 31 supir dan 135 record data. Atribut data yang digunakan meliputi kode supir, nama supir, tahun, dan jumlah pengiriman.

Preprocessing Data

Dalam Tahap preprocessing dilakukan dalam dua langkah. Pertama, agregasi data dilakukan dengan menghitung rata-rata jumlah pengiriman per supir selama periode 2020–2025, sehingga setiap supir direpresentasikan oleh satu nilai rata-rata yang mencerminkan produktivitasnya secara keseluruhan. Kedua, nilai rata-rata tersebut dinormalisasi menggunakan metode *Min-Max Scaling* untuk mengubah seluruh nilai ke rentang 0 hingga 1, sehingga perbedaan skala data tidak memengaruhi hasil perhitungan jarak pada proses *clustering*.

Penentuan Jumlah Cluster (Elbow Method)

Dalam menentukan jumlah *cluster* yang optimal, penelitian ini menggunakan *Elbow Method* sebagai pendekatannya. Metode ini bekerja dengan menghitung nilai *inertia* pada setiap nilai k, kemudian hasilnya diplot dalam bentuk grafik sehingga pola penurunan yang menyerupai siku dapat teridentifikasi dengan lebih mudah. Titik siku tersebut menandakan nilai k yang paling optimal, yakni nilai k terkecil yang masih menghasilkan nilai *within-cluster sum of squares* (*withinss*) rendah. Untuk mengukur seberapa besar perbedaan antara data aktual dengan pusat *cluster*-nya, digunakan teknik *SSE (Sum of Square Error)* sebagai dasar evaluasinya (Maori, 2023).

Pengujian *Elbow Method* pada penelitian ini dilakukan dengan menghitung nilai *inertia* untuk k=1 hingga k=7. Berdasarkan grafik yang dihasilkan pada Gambar 2, terlihat bahwa penurunan nilai *inertia* mulai melambat secara signifikan pada k=3, membentuk pola menyerupai siku. Oleh karena itu, k=3 dipilih sebagai jumlah *cluster* optimal yang digunakan dalam proses *clustering*.

Proses K-Means Clustering

Pada penelitian ini, proses pengelompokan kinerja supir dilakukan menggunakan algoritma *K-Means Clustering*. Algoritma ini termasuk ke dalam kategori *unsupervised learning* yang mampu membagi data ke dalam sejumlah kelompok secara otomatis tanpa memerlukan label kelas sebelumnya. Prinsip kerjanya adalah memastikan bahwa data yang berada dalam satu *cluster* memiliki kemiripan yang tinggi satu sama lain, sekaligus memiliki perbedaan yang jelas terhadap data di *cluster* lainnya (No et al., 2019). Proses *clustering* dijalankan secara iteratif melalui tahapan berikut:

a. Menentukan banyak k.
b. Menginisialisasi posisi *centroid* awal dengan acak sebanyak banyak titik k.
c. Menghitung seberapa dekat antar data pada masing-masing *centroid* dengan menggunakan jarak *Euclidean* dengan rumus::

$$D(i,j) = \sqrt{\sum_k (X_{ik} - C_{jk})^2} \quad \text{.....(2)}$$

Di mana D(i,j) adalah jarak data ke-i ke pusat *cluster* ke-j, X adalah nilai data, dan C adalah nilai *centroid*.

d. Setiap dari data dialokasikan ke *cluster* yang memiliki posisi paling dekat dengan *centroid*.
e. Posisi dari *centroid* diperbarui berdasarkan nilai rata-rata dari seluruh anggota *cluster* yang baru terbentuk.
f. Langkah 3–5 diulang secara terus-menerus hingga posisi *centroid* tidak mengalami perubahan lagi (*konvergen*).

Pada penelitian ini, proses *clustering* dilakukan dengan k=3 yang menghasilkan pengelompokan supir ke dalam tiga kategori kinerja yaitu Tinggi, Sedang, dan Rendah berdasarkan jumlah rata-rata pengiriman dari tiap masing-masing supir.

Evaluasi Cluster

Pada penelitian ini, kualitas hasil clustering dievaluasi menggunakan dua metrik, yaitu *Silhouette Score* dan *Davies-Bouldin Index* (DBI). *Silhouette Score* merupakan metrik evaluasi yang mengukur sejauh mana setiap titik data benar-benar sesuai berada di dalam *cluster*-nya sendiri dibandingkan dengan *cluster* lain yang ada. Nilai *silhouette* dihitung untuk setiap titik data secara individual, sehingga mencerminkan tingkat kohesi dan separasi hasil pengelompokan secara keseluruhan. Semakin tinggi nilai *Silhouette Score* yang dihasilkan, semakin baik pula kualitas pemisahan antar *cluster* yang terbentuk. Sebaliknya, nilai yang rendah mengindikasikan bahwa sebagian data tidak terdistribusi secara homogen dalam *cluster*-nya(Hendrastuty, 2024) .

Davies-Bouldin Index (DBI) bekerja dengan cara mengukur tingkat kemiripan setiap *cluster* terhadap *cluster* lain yang paling menyerupainya, berdasarkan perbandingan antara sebaran data di dalam *cluster* (*intra-cluster*) dengan jarak antar pusat *cluster* (*inter-cluster*). Nilai DBI diperoleh dari rata-rata seluruh perbandingan tersebut, sehingga semakin rendah nilai yang dihasilkan, semakin kompak dan terpisah *cluster* yang terbentuk, yang berarti kualitas *clustering* semakin baik (Menggunakan & Autoencoder, 2025). Selain itu, dilakukan perbandingan nilai evaluasi untuk k=2, k=3, dan k=4 sebagai justifikasi pemilihan k=3.

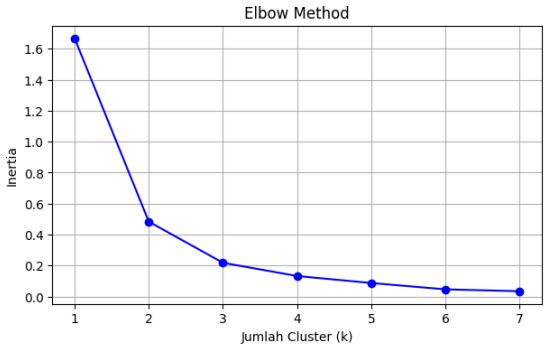
4. HASIL DAN PEMBAHASAN

Hasil Preprocessing Data

Sebelum dilakukan clustering, data historis jumlah pengiriman per supir diagregasi dengan menghitung rata-rata pengiriman selama periode 2020–2025. Hasil agregasi menunjukkan bahwa rata-rata jumlah pengiriman per supir berkisar antara 27.00 hingga 138.33 pengiriman per tahun. Selanjutnya data dinormalisasi menggunakan metode Min-Max Scaling sehingga seluruh nilai berada pada rentang 0 hingga 1.

Hasil Penentuan Jumlah Cluster (Elbow Method)

Pengujian *Elbow Method* pada penelitian ini dilakukan dengan menghitung nilai *inertia* pada setiap nilai k mulai dari k=1 hingga k=7. Grafik hasil pengujian ditampilkan pada Gambar 2. Grafik Elbow Method, di mana terlihat bahwa kurva mengalami perlambatan penurunan yang cukup signifikan pada titik k=3 hingga membentuk pola menyerupai siku. Berdasarkan pola tersebut, k=3 ditetapkan sebagai jumlah *cluster* optimal yang digunakan dalam proses *clustering*.



Gambar 2. Grafik Elbow Method.

Hasil Perbandingan Nilai k

Untuk memperkuat justifikasi pemilihan k=3, dilakukan perbandingan nilai evaluasi untuk k=2, k=3, dan k=4 sebagaimana ditampilkan pada.

Tabel 1. Perbandingan Nilai Evaluasi Clustering.

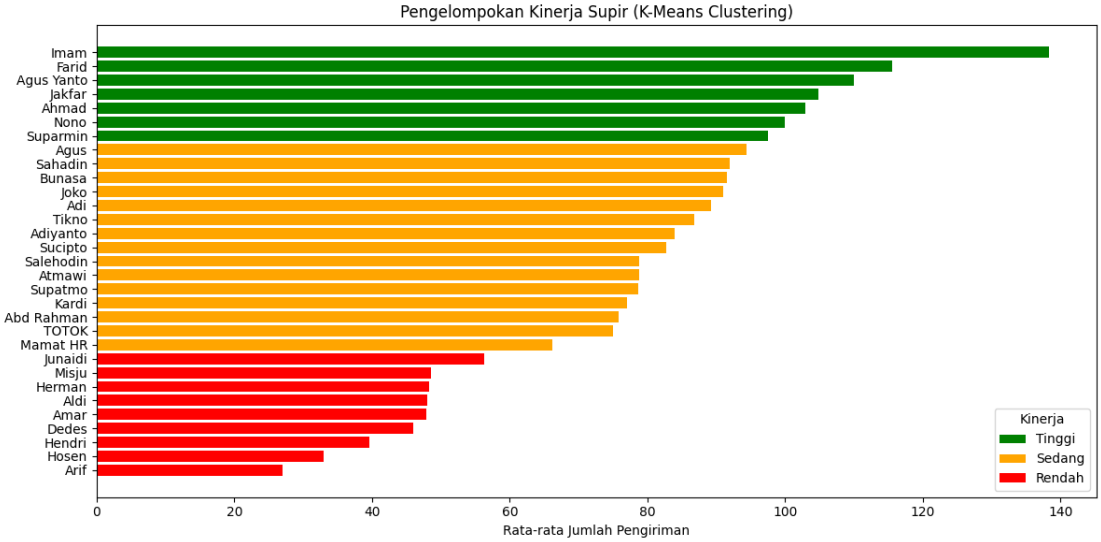
k	Silhouette Score	Davies-Bouldin Index
2	0.6490	0.4169
3	0.5948	0.5191
4	0.5734	0.5132

Meskipun secara metrik nilai *Silhouette Score* pada k=2 menghasilkan angka yang lebih tinggi dibandingkan k=3, pemilihan k=3 tetap dipertahankan dalam penelitian ini. Hal ini didasarkan pada pertimbangan bahwa pembagian ke dalam tiga kategori kinerja yaitu Tinggi,

Sedang, dan Rendah dinilai lebih representatif secara kontekstual dan lebih mudah ditindaklanjuti oleh manajemen perusahaan sebagai acuan pengambilan keputusan.

Hasil Proses K-Means Clustering

Penerapan *K-Means Clustering* dengan k=3 pada penelitian ini berhasil membagi 31 supir ke dalam tiga kelompok kinerja yaitu Tinggi, Sedang, dan Rendah. Dasar pengelompokan yang digunakan adalah nilai rata-rata jumlah pengiriman masing-masing supir sepanjang periode 2020–2025. Distribusi hasil pengelompokan seluruh supir dapat diamati pada Gambar 3. Bar Chart Pengelompokan Kinerja Supir, sementara ringkasan hasil *clustering* secara keseluruhan disajikan pada **Error! Reference source not found.**.



Gambar 3. Bar Chart Pengelompokan Kinerja Supir.

Tabel 2. Hasil Clustering Berdasarkan Kinerja.

Kinerja	Jumlah Supir	Rata-rata Trip	Min Trip	Max Trip
Rendah	9	43.83	27.00	56.25
Sedang	15	82.79	66.20	94.33
Tinggi	7	109.88	97.50	138.33

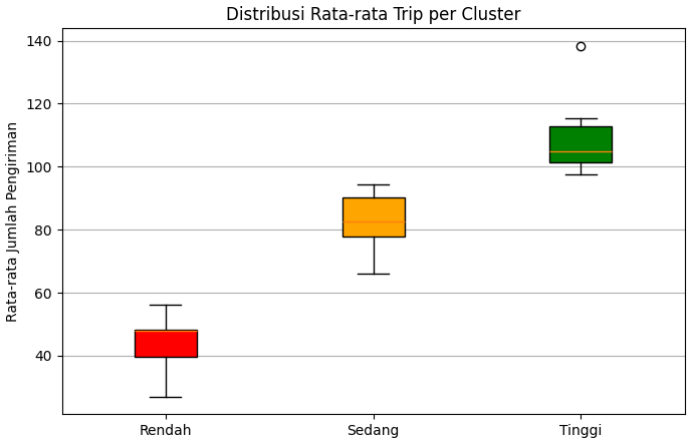
Kelompok pertama adalah kategori kinerja rendah yang beranggotakan 9 supir dengan rata-rata jumlah pengiriman sebesar 43.83 *trip* per tahun. Supir-supir dalam kelompok ini memiliki rentang pengiriman antara 27.00 hingga 56.25 *trip* per tahun, dengan rincian sebagai berikut: Arif (27.00), Hosen (33.00), Hendri (39.67), Dedes (46.00), Amar (47.83), Aldi (48.00), Herman (48.25), Misju (48.50), dan Junaidi (56.25).

Kelompok kedua adalah kategori kinerja sedang yang sekaligus menjadi kelompok dengan jumlah anggota terbanyak, yakni 15 supir dengan rata-rata pengiriman 82.79 *trip* per tahun. Rentang pengiriman pada kelompok ini berkisar antara 66.20 hingga 94.33 *trip* per

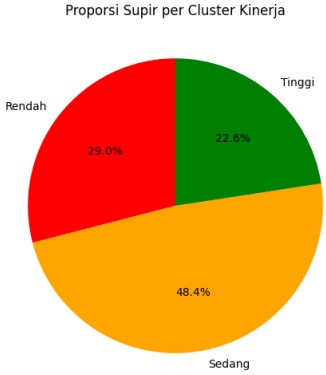
tahun, yang terdiri dari: Mamat HR (66.20), TOTOK (75.00), Abd Rahman (75.75), Kardi (77.00), Supatmo (78.67), Salehodin (78.83), Atmawi (78.83), Sucipto (82.67), Adiyanto (84.00), Tikno (86.83), Adi (89.20), Joko (91.00), Bunasa (91.50), Sahadin (92.00), dan Agus (94.33).

Kelompok ketiga adalah kategori kinerja tinggi yang beranggotakan 7 supir dengan rata-rata pengiriman tertinggi yaitu 109.88 *trip* per tahun. Rentang pengiriman kelompok ini berada di antara 97.50 hingga 138.33 *trip* per tahun, terdiri dari: Suparmin (97.50), Nono (100.00), Ahmad (103.00), Jakfar (104.83), Agus Yanto (110.00), Farid (115.50), dan Imam (138.33).

Sebaran data masing-masing kelompok dapat dicermati lebih lanjut melalui *boxplot* pada Gambar 4. Boxplot Distribusi Rata-rata Trip per Cluster. Dari visualisasi tersebut, tampak bahwa ketiga kelompok memiliki distribusi data yang cukup terpisah satu sama lain, yang mengonfirmasi bahwa hasil pengelompokan yang diperoleh cukup baik. Selain itu, ditemukan satu *outlier* pada kelompok Tinggi, yaitu supir Imam dengan rata-rata 138.33 *trip* per tahun yang jauh di atas anggota kelompok lainnya. Adapun proporsi jumlah supir per kelompok disajikan secara visual pada Gambar 5. Pie Chart Proporsi Supir per Cluster.



Gambar 4. Boxplot Distribusi Rata-rata Trip per Cluster.



Gambar 5. Pie Chart Proporsi Supir per Cluster

Hasil *clustering* menunjukkan bahwa sebagian besar supir berada pada kelompok kinerja Sedang dengan jumlah 15 supir (48.4%), diikuti kelompok Rendah dengan 9 supir (29.0%), dan kelompok Tinggi dengan 7 supir (22.6%). Hasil akhir pengelompokan seluruh supir pada Tabel 3. Clustering Kinerja Supir.

Tabel 3. Clustering Kinerja Supir.

Kode Supir	Nama Supir	Rata-rata Trip	Kinerja
S-014	Arif	27	Rendah
S-036	Hosen	33	Rendah
S-046	Hendri	39,67	Rendah
S-015	Dedes	46	Rendah
S-016	Amar	47,83	Rendah
S-033	Aldi	48	Rendah
S-003	Herman	48,25	Rendah
S-009	Misju	48,5	Rendah
S-013	Junaidi	56,25	Rendah
S-010	Mamat HR	66,2	Sedang
S-029	TOTOK	75	Sedang
S-027	Abd Rahman	75,75	Sedang
S-018	Kardi	77	Sedang
S-043	Supatmo	78,67	Sedang
S-008	Salehodin	78,83	Sedang
S-019	Atmawi	78,83	Sedang
S-001	Sucipto	82,67	Sedang
S-037	Adiyanto	84	Sedang
S-035	Tikno	86,83	Sedang
S-045	Adi	89,2	Sedang
S-012	Joko	91	Sedang
S-006	Bunasa	91,5	Sedang
S-005	Sahadin	92	Sedang
S-007	Agus	94,33	Sedang
S-050	Suparmin	97,5	Tinggi
S-004	Nono	100	Tinggi
S-002	Ahmad	103	Tinggi
S-011	Jakfar	104,83	Tinggi
S-026	Agus Yanto	110	Tinggi
S-034	Farid	115,5	Tinggi
S-025	Imam	138,33	Tinggi

Hasil Evaluasi Cluster

Untuk mengukur kualitas hasil pengelompokan yang diperoleh, digunakan dua metrik evaluasi yaitu *Silhouette Score* dan *Davies-Bouldin Index* (DBI) dengan hasil sebagai berikut:

a. *Silhouette Score*: 0.5948 — nilai ini mencerminkan bahwa antar kelompok yang terbentuk sudah cukup kohesif dan memiliki batas pemisahan yang jelas.
b. *Davies-Bouldin Index*: 0.5191 — nilai ini menunjukkan bahwa setiap kelompok yang terbentuk bersifat kompak dengan tingkat separasi antar kelompok yang memadai.

Secara keseluruhan, kedua metrik tersebut memberikan gambaran bahwa hasil pengelompokan yang dihasilkan pada penelitian ini dapat dinyatakan valid dan layak digunakan sebagai dasar analisis lebih lanjut.

5. KESIMPULAN DAN SARAN

Kesimpulan

Penelitian ini berhasil mengklasifikasikan kinerja supir pada sebuah perusahaan transportasi di Surabaya melalui penerapan algoritma *K-Means Clustering* yang didasarkan pada rekam jejak historis jumlah pengiriman selama periode 2020–2025. Adapun poin-poin kesimpulan yang dapat dirumuskan adalah sebagai berikut.

Pertama, algoritma *K-Means Clustering* yang diterapkan pada data 31 supir berhasil menghasilkan tiga segmen kinerja yaitu Tinggi, Sedang, dan Rendah. Segmen Sedang mendominasi dengan 15 supir (48,4%), disusul segmen Rendah sebanyak 9 supir (29,0%), dan segmen Tinggi sebanyak 7 supir (22,6%).

Kedua, *Elbow Method* ditetapkan bahwa k=3 merupakan jumlah *cluster* yang paling optimal, yang ditandai dengan melandainya penurunan nilai *inertia* secara signifikan pada titik tersebut. Walaupun nilai *Silhouette Score* pada k=2 lebih tinggi (0.6490), k=3 tetap dipertahankan karena pembagian ke dalam tiga kategori dinilai lebih kontekstual dan mudah ditindaklanjuti oleh manajemen perusahaan.

Ketiga, pengukuran kualitas *clustering* menggunakan *Silhouette Score* menghasilkan angka 0.5948 yang mencerminkan bahwa batas antar kelompok terbentuk dengan cukup jelas. Sementara itu, *Davies-Bouldin Index* sebesar 0.5191 mempertegas bahwa setiap kelompok bersifat kompak dan memiliki separasi yang memadai, sehingga hasil pengelompokan secara keseluruhan dapat dinyatakan valid.

Keempat, Kelompok Tinggi mencatatkan rata-rata 109.88 *trip* per tahun dengan rentang 97.50–138.33 *trip*. Supir Imam yang membukukan 138.33 *trip* per tahun teridentifikasi sebagai *outlier* karena nilainya jauh melampaui rekan-rekan satu kelompoknya. Kelompok Sedang dan Rendah masing-masing mencatatkan rata-rata 82.79 dan 43.83 *trip* per tahun, dengan supir Arif (27.00 *trip*) sebagai yang terendah.

Kelima, temuan ini diharapkan dapat dijadikan rujukan oleh pihak manajemen dalam merancang kebijakan yang lebih terukur dan berbasis data, baik dalam hal penilaian kinerja, pemberian penghargaan bagi supir berprestasi, maupun pelaksanaan program pembinaan bagi supir yang masih perlu ditingkatkan kinerjanya.

Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dipertimbangkan untuk penelitian maupun pengembangan lebih lanjut.

Pada penelitian berikutnya, disarankan untuk memperkaya atribut data yang digunakan dalam proses *clustering* dengan menambahkan variabel-variabel seperti ketepatan waktu pengiriman, jarak tempuh, tingkat komplain pelanggan, dan data kehadiran supir. Dengan cakupan variabel yang lebih luas, hasil pengelompokan diharapkan mampu merepresentasikan kinerja supir secara lebih menyeluruh dan komprehensif.

Untuk memperkuat validitas pemilihan metode, perlu dilakukan perbandingan antara algoritma *K-Means Clustering* dengan pendekatan *clustering* lainnya seperti *K-Medoids*, *DBSCAN*, atau *Fuzzy C-Means*. Perbandingan tersebut dapat memberikan gambaran yang lebih objektif mengenai metode mana yang paling optimal untuk diterapkan pada data kinerja supir.

Ke depannya, hasil penelitian ini dapat dikembangkan lebih lanjut menjadi sebuah sistem informasi berbasis *web* atau aplikasi yang mengintegrasikan algoritma *K-Means Clustering* secara otomatis. Dengan adanya sistem tersebut, manajemen perusahaan dapat memantau dan mengevaluasi kinerja supir secara *real-time* tanpa perlu bergantung pada analisis manual yang memakan waktu.

Disarankan kepada manajemen perusahaan untuk menjadikan hasil pengelompokan ini sebagai landasan dalam merancang program pengembangan sumber daya manusia yang lebih terstruktur, seperti pemberian *reward* atau insentif bagi supir berkinerja Tinggi, pelaksanaan pelatihan dan pendampingan bagi supir berkinerja Sedang agar dapat berkembang ke kelompok Tinggi, serta evaluasi mendalam terhadap supir berkinerja Rendah guna mengidentifikasi faktor-faktor yang menghambat produktivitas mereka. Guna meningkatkan akurasi dan stabilitas hasil *clustering*, penelitian selanjutnya disarankan untuk menggunakan data dengan rentang waktu yang lebih panjang dan jumlah sampel supir yang lebih besar, sehingga pola kinerja yang dihasilkan lebih stabil dan dapat digeneralisasi dengan lebih baik.

DAFTAR REFERENSI

Anikah, I., Surip, A., Rahayu, N. P., Musa, M. H. A., & Tohidi, E. (2022). Pengelompokan data barang dengan menggunakan metode K-Means untuk menentukan stok persediaan barang. *Jurnal Informatika dan Rekayasa Perangkat Lunak*, *4*(2), 58–64.

Clarissa, N., Suhartanto, A., Wijoyo, S. H., & Purnomo, W. (2025). Strategi peningkatan SDM berdasarkan pengelompokan kualitas kinerja pegawai CV Mediatama Perkasa Bogor menggunakan K-Means clustering. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, *9*(5), 1–7.

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.

Hendrastuty, N. (2024). Penerapan data mining menggunakan algoritma K-Means clustering dalam evaluasi hasil pembelajaran siswa. *Jurnal Komputasi dan Sistem Informasi*, *3*, 46–56.

Kurniawan, A., & Pratama, R. A. (2025). Penerapan algoritma K-Means clustering untuk pemetaan kinerja kurir pengiriman barang berdasarkan riwayat operasional. *Jurnal Teknologi Informasi dan Ilmu Komputer*, *12*(1), 55–66. https://doi.org/10.25126/jtiik.20251216120

Maori, N. A. (2023). Metode Elbow dalam optimasi jumlah cluster pada K-Means clustering. *Jurnal Ilmiah Matematika dan Sains*, *14*(2), 277–287.

Mauludin, M. R., Nurdiawan, O., & Basysyar, F. M. (2025). Penerapan algoritma K-Means clustering untuk analisis kinerja pengiriman paket Shopee Express di Hub Transit Kedawung. *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, *13*(1), 1188–1192.

Nugroho, F. A., & Handayani, S. (2024). Analisis klastering K-Means dalam penentuan tingkat efisiensi armada distribusi dan pengemudi pada perusahaan logistik. *Jurnal Rekayasa Sistem dan Industri*, *11*(2), 112–123.

Nurjanah, N. (2025). *Masa depan kurir Indonesia: Inovasi layanan logistik sebagai kunci sukses*. Penerbit Widina.

Olivia, L. (2023). Optimalisasi pengiriman barang dengan mengelompokkan titik pengiriman menggunakan metode K-Means clustering dan Hierarchical clustering. *Jurnal Komputer dan Informatika*, *18*(2), 116–123.

Pratama, M. R., & Wibowo, A. (2023). Pengelompokan data kinerja pegawai menggunakan metode K-Means berbasis Silhouette Coefficient. *Jurnal Media Informatika Budidarma*, *7*(3), 1230–1239. https://doi.org/10.30865/mib.v7i3.6120

Ramadhani, A. N., & Athalina, G. (2025). Optimalisasi kinerja karyawan berbasis HR analytics dengan K-Means clustering dan analisis faktor demografi. *Jurnal Manajemen dan Analisis Data*, *15*(1), 1–14.

Santi, I. H. (2020). *Analisis data menggunakan clustering K-Means*. Penerbit Deepublish.

Setiawan, H., & Rahmawati, D. (2024). Evaluasi penentuan jumlah cluster optimum pada algoritma K-Means menggunakan metode Elbow. *Jurnal Ilmiah Intech: Information Technology Journal*, *6*(1), 45–54.

Tamba, S. P., & Kesuma, F. T. (2019). Penerapan data mining untuk menentukan penjualan sparepart Toyota dengan metode K-Means clustering. *Jurnal Pengolahan Data Informatika*, *2*(2).

Utami, W. S., & Febrianto, E. (2025). Pemanfaatan data historis operasional untuk pengelompokan produktivitas tenaga kerja pada industri jasa pengiriman. *Jurnal Riset Sistem Informasi dan Teknik Komputer*, *10*(1), 88–98.