



Department of Digital Business

**Journal of Artificial Intelligence and Digital Business (RIGGS)**

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 5 No. 2 (2026) pp: 3381-3389

P-ISSN: 2963-9298, e-ISSN: 2963-914X

---

## Analisis Pola Kepadatan Penumpang Menggunakan *K-Means Clustering* Pada Pelabuhan Penyeberangan Torobulu–Tampo

Muhammad Chaerur Rijal<sup>1</sup>, Muhamad Faza Almaliki<sup>2</sup>, Isnawaty<sup>3</sup>, Maudhy Satyadharna<sup>4</sup>

<sup>1</sup>Jurusan Teknik Elektro, Politeknik Negeri Ujung Pandang

<sup>2</sup>Program Studi Doktor Informatika, Universitas Hasanuddin

<sup>3</sup>Jurusan Informatika, Universitas Halu Oleo

<sup>4</sup>Dinas Perhubungan Provinsi Sulawesi Tenggara

<sup>1</sup>[m.ch\\_rijal@poliupg.ac.id](mailto:m.ch_rijal@poliupg.ac.id), <sup>2</sup>[muhamadfazaalmaliki@gmail.com](mailto:muhamadfazaalmaliki@gmail.com), <sup>3</sup>[isnawaty@uho.ac.id](mailto:isnawaty@uho.ac.id), <sup>4</sup>[maudhymaudhy@gmail.com](mailto:maudhymaudhy@gmail.com)

### Abstrak

Penelitian ini bertujuan untuk menganalisis pola kepadatan penumpang pada lintasan penyeberangan Torobulu–Tampo menggunakan metode *K-Means Clustering* sebagai pendekatan analisis operasional transportasi berbasis data. Data yang digunakan dalam penelitian ini berupa data harian jumlah penumpang Pelabuhan Torobulu selama periode Januari 2023 hingga Oktober 2025. Penelitian dilakukan melalui beberapa tahapan, yaitu *preprocessing data* untuk membersihkan data yang tidak konsisten, proses normalisasi untuk menyamakan skala data, penentuan jumlah cluster optimal menggunakan metode evaluasi clustering, serta proses pengelompokan data menggunakan algoritma *K-Means*. Hasil penelitian menunjukkan bahwa pola jumlah penumpang dapat dikelompokkan menjadi tiga kategori utama, yaitu kondisi sepi (*low*), normal (*medium*), dan padat (*peak season*). Cluster padat didominasi oleh periode hari besar nasional, musim mudik Idulfitri, dan libur panjang yang menyebabkan peningkatan mobilitas masyarakat secara signifikan. Sementara itu, cluster normal dan sepi lebih banyak terjadi pada periode operasional reguler di luar musim liburan. Hasil clustering juga menunjukkan adanya pola musiman yang memengaruhi tingkat kepadatan penumpang pada lintasan penyeberangan Torobulu–Tampo. Informasi tersebut dapat dimanfaatkan sebagai dasar dalam optimalisasi jadwal operasional kapal, pengaturan kapasitas penumpang, serta peningkatan kualitas pelayanan transportasi penyeberangan. Selain itu, hasil penelitian ini dapat membantu pihak pengelola pelabuhan dalam melakukan perencanaan operasional yang lebih efektif dan adaptif terhadap perubahan pola mobilitas penumpang. Penelitian ini diharapkan menjadi dasar pengembangan sistem analisis transportasi penyeberangan berbasis data yang lebih cerdas, efisien, terintegrasi, dan berkelanjutan di masa mendatang.

**Kata kunci:** *K-Means Clustering*, Pola Penumpang, Transportasi Penyeberangan, Optimalisasi Jadwal Kapal

### 1. Latar Belakang

Transportasi penyeberangan memiliki peran penting dalam mendukung konektivitas wilayah, mobilitas masyarakat, serta distribusi barang dan jasa antarwilayah kepulauan di Indonesia [1]. Salah satu lintasan penyeberangan yang memiliki aktivitas mobilitas cukup tinggi adalah penyeberangan Torobulu–Tampo yang menghubungkan wilayah daratan dan wilayah kepulauan di Sulawesi Tenggara. Aktivitas penumpang pada penyeberangan ini cenderung mengalami fluktuasi yang dipengaruhi oleh faktor musiman, hari besar nasional, serta periode mudik dan libur panjang. Kondisi tersebut menyebabkan terjadinya lonjakan penumpang pada waktu tertentu yang berpotensi memengaruhi efektivitas operasional kapal dan kualitas pelayanan transportasi penyeberangan [2].

Perkembangan konsep *Intelligent Transportation Systems* mendorong pemanfaatan teknologi berbasis data dalam mendukung pengambilan keputusan operasional transportasi secara lebih adaptif dan efisien [3]. Salah satu pendekatan yang banyak digunakan adalah penerapan teknik *data mining* dan *machine learning* untuk mengidentifikasi pola mobilitas penumpang [4]. Penelitian [5] menggunakan metode ARIMA untuk memprediksi produksi barang angkutan kereta api nasional periode 2019–2026 berdasarkan data deret waktu dari BPS. Hasil penelitian menunjukkan bahwa model ARIMA mampu menghasilkan prediksi yang stabil dan andal sehingga dapat mendukung perencanaan transportasi logistik berbasis kereta api. Sementara itu, penelitian [6] menggunakan metode *Long Short-Term Memory* (LSTM) untuk memprediksi jumlah penumpang *Busway* DKI Jakarta berdasarkan data historis *time series* tahun 2015–2016. Hasil penelitian menunjukkan bahwa metode LSTM

---

Analisis Pola Kepadatan Penumpang Menggunakan *K-Means Clustering* Pada Pelabuhan Penyeberangan Torobulu–Tampo

memberikan tingkat akurasi terbaik dibandingkan *Exponential Smoothing* dengan nilai rata-rata RMSE sebesar 2640,53 dan MAPE sebesar 9,14%, sehingga dinilai efektif dalam melakukan *forecasting* jumlah penumpang transportasi publik.

Selanjutnya, penelitian [7] menerapkan metode *clustering* dengan algoritma *K-Means Clustering* untuk mengelompokkan karakteristik wilayah transportasi umum di Jawa Barat berdasarkan indikator kepadatan penduduk, jumlah kendaraan bermotor, ketersediaan transportasi umum, panjang jalan, anggaran transportasi, dan pendapatan per kapita. Hasil penelitian membentuk dua klaster utama, yaitu wilayah metropolitan dan wilayah pinggiran, yang kemudian digunakan sebagai dasar penyusunan strategi kebijakan transportasi umum yang lebih adaptif sesuai karakteristik masing-masing wilayah. Berdasarkan beberapa penelitian tersebut, metode *forecasting* seperti ARIMA dan *Long Short-Term Memory (LSTM)* banyak digunakan untuk memprediksi jumlah penumpang transportasi umum, sedangkan metode *clustering* digunakan untuk mengelompokkan pola mobilitas pada sektor transportasi perkotaan dan sistem transportasi publik berbasis cerdas. Namun, sebagian besar penelitian tersebut masih berfokus pada transportasi darat dan perkotaan, sementara penelitian terkait transportasi penyeberangan masih relatif terbatas, khususnya pada lintasan penyeberangan regional di Indonesia.

Data historis jumlah penumpang merupakan salah satu indikator utama dalam menggambarkan dinamika mobilitas transportasi penyeberangan [8]. Pola perubahan jumlah penumpang pada periode tertentu dapat digunakan untuk mengidentifikasi kecenderungan kepadatan operasional, terutama pada musim mudik, hari besar nasional, dan periode libur panjang [9]. Oleh karena itu, analisis berbasis data historis penumpang menjadi penting sebagai dasar pengambilan keputusan operasional transportasi secara lebih adaptif [10]. Berdasarkan kondisi tersebut, terdapat kesenjangan penelitian pada pemanfaatan metode *clustering* untuk menganalisis pola kepadatan penumpang transportasi penyeberangan sebagai dasar pengambilan keputusan operasional kapal.

Selain itu, sebagian penelitian sebelumnya lebih menitikberatkan pada prediksi jumlah penumpang tanpa mengidentifikasi karakteristik kepadatan penumpang berdasarkan pola historis harian. Oleh karena itu, penelitian ini dilakukan untuk mengidentifikasi pola kepadatan penumpang pada lintasan penyeberangan Torobulu-Tampo menggunakan metode *K-Means Clustering*. Metode ini dipilih karena mampu mengelompokkan data penumpang berdasarkan tingkat kemiripan karakteristik kepadatan sehingga dapat digunakan untuk mendukung optimalisasi layanan transportasi penyeberangan [11]. Pendekatan *clustering* digunakan karena mampu melakukan pengelompokan pola data secara *unsupervised* tanpa memerlukan label kelas sebelumnya [12]. Dengan demikian, pola kepadatan penumpang dapat diidentifikasi berdasarkan kemiripan karakteristik distribusi data historis penumpang harian.

Kebaruan penelitian ini terletak pada penerapan analisis *clustering* berbasis data historis penumpang harian pada transportasi penyeberangan regional Torobulu-Tampo yang masih relatif jarang diteliti, khususnya dalam mengidentifikasi pola kepadatan operasional sebagai dasar pendukung pengambilan keputusan jadwal kapal berbasis data historis. Selain itu, sebagian besar penelitian sebelumnya berfokus pada pendekatan prediktif jumlah penumpang, sementara identifikasi karakteristik kepadatan operasional berbasis *clustering* pada transportasi penyeberangan regional masih terbatas. Penelitian ini juga memberikan kontribusi dalam bentuk analisis pola musiman penumpang yang dapat dimanfaatkan sebagai pendukung pengambilan keputusan operasional kapal secara lebih efektif. Berdasarkan uraian tersebut, tujuan penelitian ini adalah menganalisis pola kepadatan penumpang menggunakan metode *K-Means Clustering* serta mengidentifikasi karakteristik periode kepadatan penumpang pada lintasan penyeberangan Torobulu-Tampo sebagai pendukung optimalisasi operasional transportasi penyeberangan berbasis data.

## 2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *data mining* untuk menganalisis pola kepadatan penumpang pada lintasan penyeberangan Torobulu-Tampo. Teknik penelitian yang digunakan adalah *clustering* menggunakan algoritma *K-Means Clustering* untuk mengelompokkan data penumpang berdasarkan tingkat kepadatan penumpang harian.

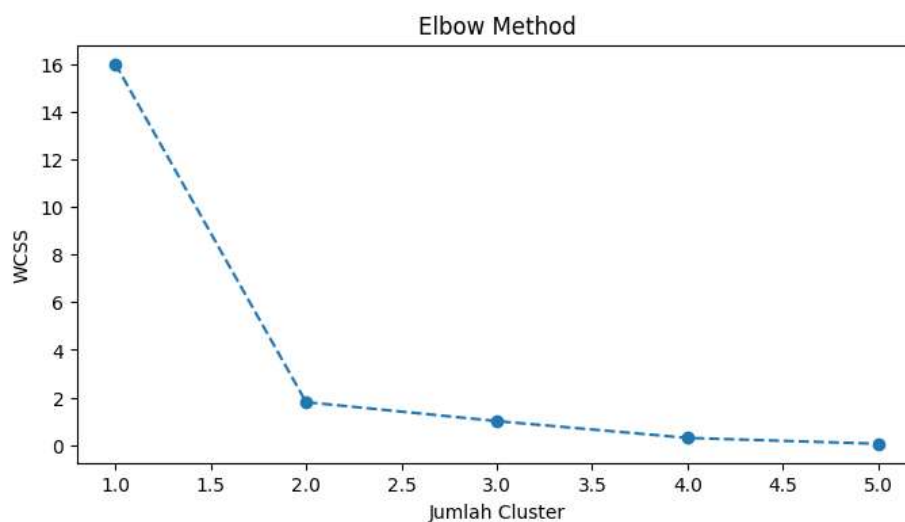
Data penelitian berupa data historis jumlah penumpang harian Pelabuhan Torobulu periode Januari 2023 hingga Oktober 2025. *Dataset* terdiri dari data tanggal dan jumlah penumpang harian yang diperoleh dari data operasional pelabuhan. Selain itu, penelitian juga menggunakan data jadwal keberangkatan kapal sebagai data pendukung untuk analisis operasional transportasi penyeberangan. Total data yang digunakan sebanyak 1.034 data harian.

Selain menggunakan variabel jumlah penumpang harian, penelitian ini juga melakukan ekstraksi fitur temporal dari data tanggal untuk memperkuat analisis pola mobilitas penumpang. *Feature engineering* dilakukan dengan menurunkan atribut waktu seperti hari operasional, bulan, kategori *weekday* dan *weekend*, serta periode hari besar nasional dan musim mudik. Pendekatan ini digunakan untuk membantu identifikasi pola musiman dan karakteristik kepadatan penumpang berdasarkan periode operasional tertentu.

Tahapan penelitian dimulai dengan proses *preprocessing* data yang meliputi pemeriksaan data kosong, pembersihan data duplikat, transformasi format tanggal, serta normalisasi data jumlah penumpang menggunakan metode *min-max scaling* untuk mengurangi perbedaan rentang nilai antar data. Selanjutnya dilakukan eksplorasi pola data untuk mengidentifikasi tren dan lonjakan penumpang pada periode tertentu. Adapun rumus *min-max scaling* yang digunakan sebagaimana pada Persamaan 2.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

Dimana  $x$  adalah data asli dan  $x'$  adalah data normalisasi. Kemudian proses *clustering* dilakukan menggunakan algoritma *K-Means Clustering* karena metode ini mampu mengelompokkan data berdasarkan tingkat kemiripan karakteristik jumlah penumpang. Penentuan jumlah *cluster* dilakukan menggunakan pendekatan analisis kebutuhan operasional transportasi penyeberangan dengan membagi data menjadi tiga kelompok utama, yaitu kategori sepi (*low*), normal (*medium*), dan padat (*peak season*). Penentuan jumlah *cluster* juga didukung melalui analisis visual menggunakan pendekatan *Elbow Method* untuk melihat titik penurunan *inertia* yang paling optimal pada proses *clustering* [13]. Adapun pendekatan tersebut dapat ditunjukkan pada Gambar 1.



Gambar 1. Pendekatan *Elbow Method*

Berdasarkan hasil *Elbow Method*, terlihat bahwa penurunan nilai *inertia* mulai melandai pada nilai  $k=3$  sehingga jumlah *cluster* optimal ditentukan sebanyak tiga *cluster* yang merepresentasikan kategori *low*, *medium*, dan *peak season*. Algoritma *K-Means Clustering* bekerja dengan menentukan *centroid* awal, menghitung jarak antar data menggunakan *Euclidean Distance*, kemudian melakukan iterasi pembaruan *centroid* hingga diperoleh *cluster* yang stabil [14]. Persamaan 2 merupakan persamaan jarak *Euclidean* yang digunakan :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

Dengan  $d(x, y)$  merupakan jarak antara data dan *centroid*,  $x_i$  adalah nilai data penumpang, dan  $y_i$  adalah nilai *centroid* pada setiap *cluster*. Proses *clustering* dilakukan menggunakan inisialisasi *centroid K-Means++* dengan jumlah iterasi maksimum sebanyak 300 iterasi dan *random state* sebesar 42 untuk menjaga konsistensi hasil

*clustering*. Kemudian secara matematis, algoritma *K-Means Clustering* bertujuan untuk meminimalkan total jarak antar data terhadap pusat *cluster* (*centroid*) yang direpresentasikan melalui *objective function*. Fungsi objektif tersebut digunakan untuk menghasilkan pengelompokan data dengan tingkat kemiripan yang tinggi dalam setiap *cluster* dan perbedaan yang jelas antar *cluster* [15]. Persamaan *objective function K-Means Clustering* ditunjukkan pada Persamaan 3.

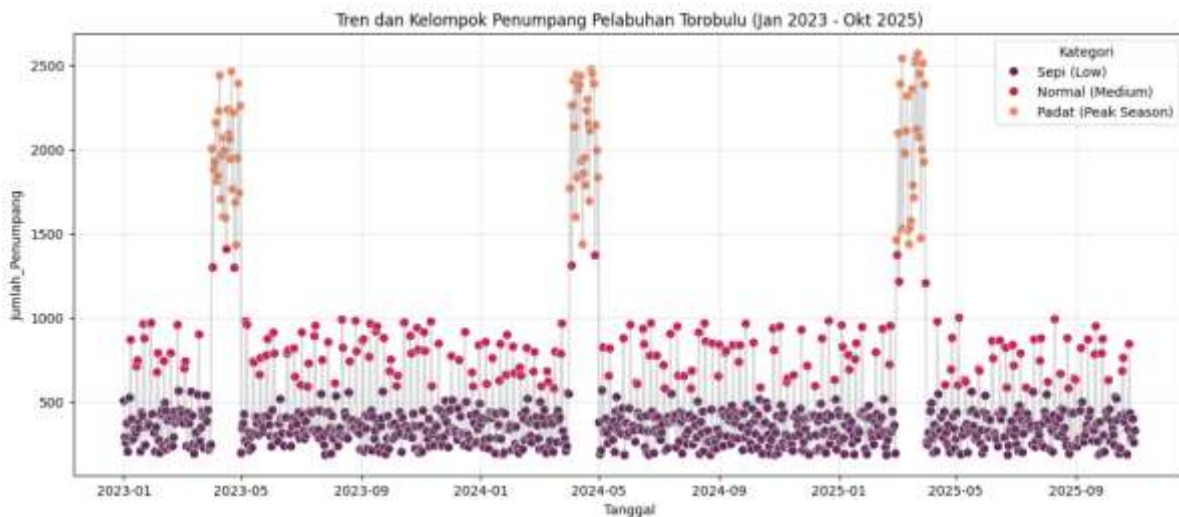
$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i - c_j\|^2 \quad (3)$$

Hasil *clustering* divisualisasikan dalam bentuk grafik tren penumpang untuk mengidentifikasi pola kepadatan berdasarkan periode waktu tertentu. Selanjutnya dilakukan analisis terhadap karakteristik masing-masing *cluster* untuk mengetahui pola musiman, periode lonjakan penumpang, serta keterkaitannya dengan jadwal operasional kapal. Seluruh proses pengolahan dan analisis data dilakukan menggunakan bahasa pemrograman Python pada platform Google Colab dengan bantuan pustaka Pandas, NumPy, Scikit-learn, dan Matplotlib.

### 3. Hasil dan Diskusi

Hasil penelitian diperoleh melalui proses *clustering* menggunakan algoritma *K-Means Clustering* terhadap data historis jumlah penumpang harian Pelabuhan Torobulu periode Januari 2023 hingga Oktober 2025. Proses *clustering* dilakukan untuk mengidentifikasi pola kepadatan penumpang berdasarkan tingkat kemiripan jumlah penumpang pada setiap periode operasional. Berdasarkan hasil pengelompokan, data penumpang terbagi menjadi tiga kategori utama, yaitu kategori sepi (*low*), normal (*medium*), dan padat (*peak season*).

Visualisasi hasil *clustering* digunakan untuk memperlihatkan pola distribusi penumpang serta perubahan tingkat kepadatan penumpang pada setiap periode waktu. Selain itu, visualisasi juga menunjukkan adanya lonjakan penumpang pada waktu tertentu yang berkaitan dengan musim mudik, hari besar nasional, dan periode libur panjang. Hasil tersebut menjadi dasar dalam menganalisis karakteristik mobilitas penumpang serta potensi penerapan teknologi berbasis data untuk mendukung optimalisasi operasional kapal penyeberangan pada lintasan Torobulu–Tampo. Gambar 2 menunjukkan hasil visualisasi *clustering* jumlah penumpang menggunakan metode *K-Means Clustering*.



Gambar 2. Hasil Visualisasi *Clustering* Jumlah Penumpang

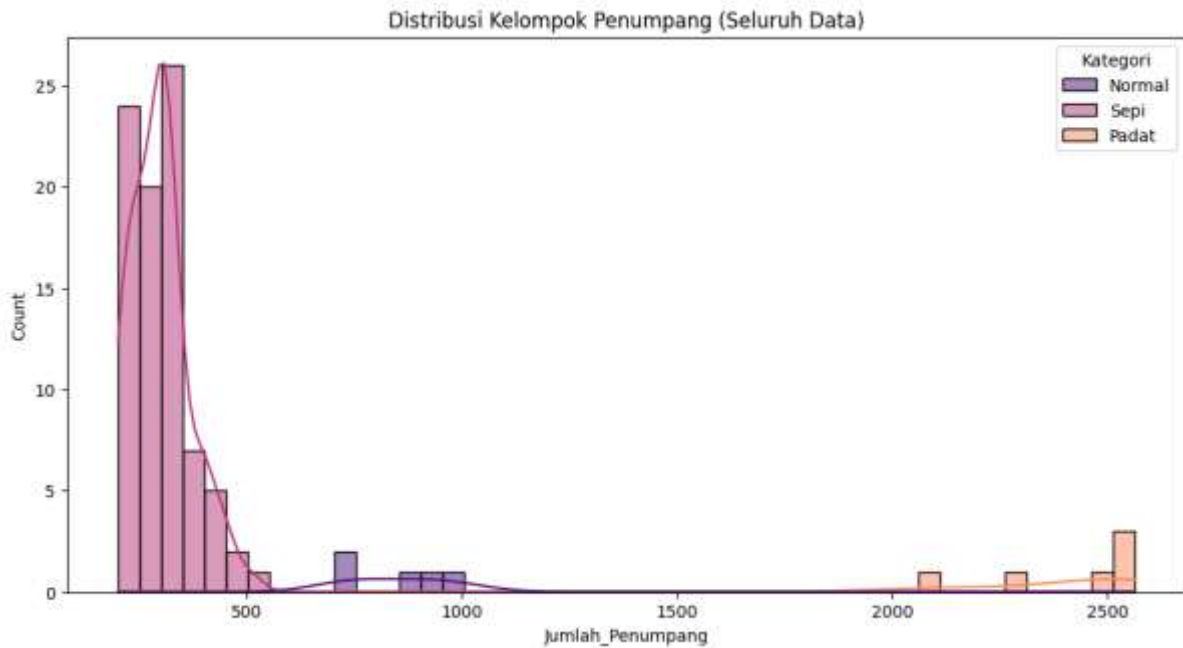
Berdasarkan visualisasi hasil *clustering* pada Gambar 2, terlihat bahwa pola kepadatan penumpang memiliki perbedaan yang cukup signifikan antara kategori sepi, normal, dan padat. Untuk memperjelas karakteristik masing-masing *cluster* yang terbentuk, terlebih dahulu disajikan ringkasan statistik hasil pengelompokan data penumpang yang meliputi jumlah data, nilai minimum, maksimum, dan rata-rata penumpang pada setiap kategori *cluster*. Ringkasan tersebut digunakan untuk memberikan gambaran umum mengenai karakteristik kepadatan penumpang pada setiap kelompok hasil *clustering* dan ditunjukkan pada Tabel 1.

Tabel 1. Statistik Hasil *Clustering* Jumlah Penumpang Berdasarkan Kategori Kepadatan

Kategori	Count	Min	Max	Mean
Normal (Medium)	196	576	1406	807
Padat ( <i>Peak Season</i> )	83	1430	2572	2035
Sepi ( <i>Low</i> )	756	180	564	334

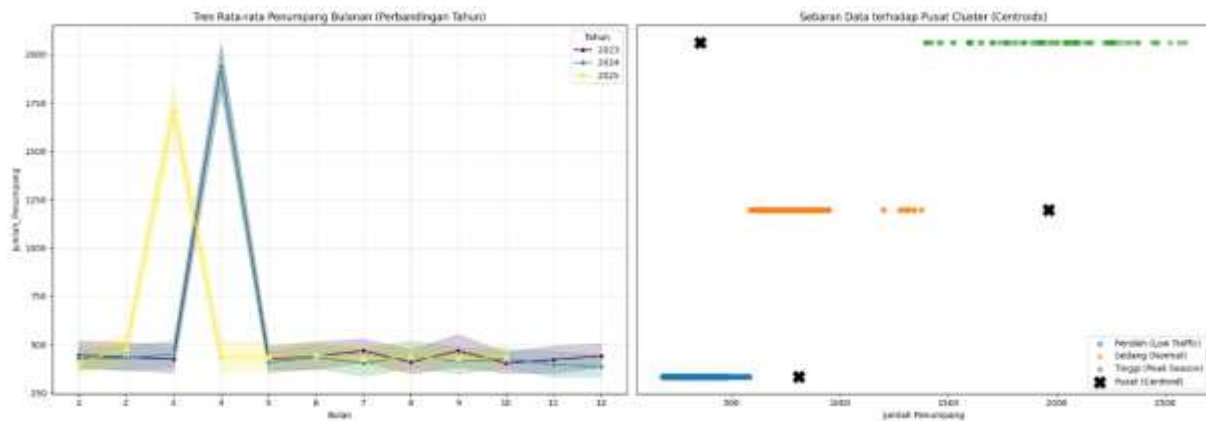
(Sumber : Olah Data, 2026)

Selanjutnya, untuk memperoleh pemahaman yang lebih menyeluruh mengenai karakteristik data penumpang pada setiap periode operasional, dilakukan analisis lanjutan melalui visualisasi distribusi data penumpang secara keseluruhan. Visualisasi ini bertujuan untuk menunjukkan penyebaran jumlah penumpang, tingkat dominasi data pada setiap rentang nilai, serta pola distribusi penumpang yang terbentuk selama periode pengamatan Januari 2023 hingga Oktober 2025. Adapun hasil distribusi data penumpang tersebut ditunjukkan pada Gambar 3.



Gambar 3. Hasil Visualisasi Distribusi Kelompok Penumpang

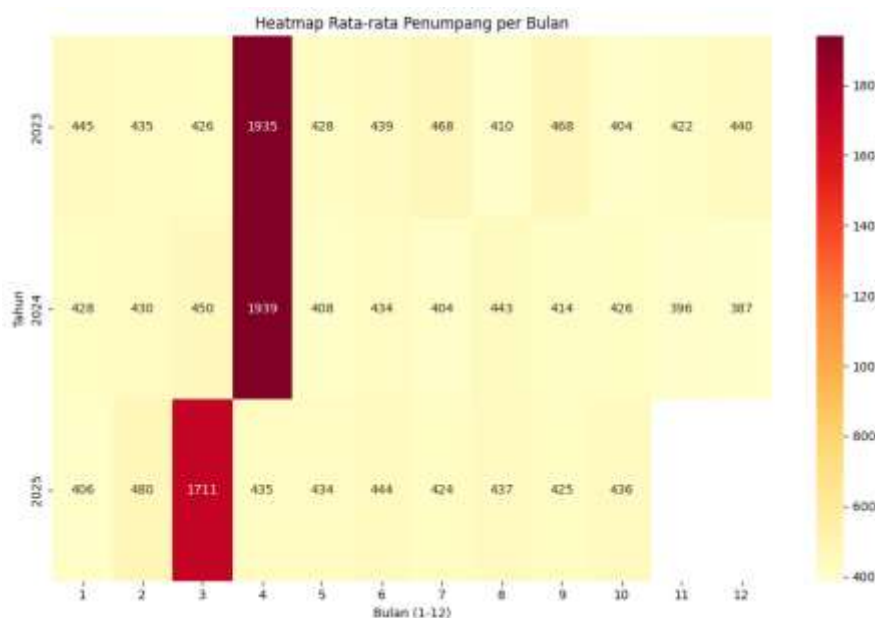
Berdasarkan hasil distribusi data penumpang, dilakukan analisis lanjutan untuk mengidentifikasi pola tren penumpang bulanan serta karakteristik penyebaran data terhadap pusat *cluster* (*centroid*) yang dihasilkan oleh algoritma *K-Means Clustering*. Analisis ini bertujuan untuk melihat perubahan rata-rata jumlah penumpang pada setiap tahun pengamatan serta mengetahui tingkat kedekatan data terhadap masing-masing *cluster* yang terbentuk. Selain itu, visualisasi *centroid* digunakan untuk memperjelas proses pengelompokan data ke dalam kategori rendah (*low traffic*), sedang (*normal*), dan tinggi (*peak season*) sehingga pola kepadatan penumpang dapat dianalisis secara lebih representatif pada lintasan penyeberangan Torobulu–Tampo sebagaimana yang ditunjukkan pada Gambar 4.



Gambar 4. Analisis Tren Penumpang

Berdasarkan hasil visualisasi tren rata-rata penumpang bulanan dan sebaran data terhadap pusat *cluster* sebagaimana pada Gambar 4, terlihat bahwa pola kepadatan penumpang pada lintasan penyeberangan Torobulu–Tampo memiliki karakteristik musiman yang cukup jelas. Lonjakan penumpang tertinggi terjadi pada periode Maret–April yang bertepatan dengan musim mudik dan libur idul fitri, dimana jumlah penumpang masuk ke dalam kategori tinggi (*peak season*) dengan pusat *cluster* berada pada kisaran lebih dari 2.000 penumpang. Sementara itu, sebagian besar periode operasional reguler berada pada kategori rendah dan sedang dengan jumlah penumpang relatif stabil pada kisaran 300–900 penumpang. Hasil *clustering* menunjukkan bahwa distribusi data pada *cluster* rendah memiliki penyebaran yang paling dominan, sedangkan *cluster* tinggi hanya muncul pada periode tertentu dengan intensitas penumpang yang sangat signifikan.

Berdasarkan analisis temporal, pola kepadatan penumpang menunjukkan karakteristik musiman yang konsisten pada periode tertentu, terutama menjelang idul fitri dan libur nasional. Selain itu, pola peningkatan penumpang juga terlihat pada akhir pekan dibandingkan hari operasional reguler. Hal ini menunjukkan bahwa dinamika mobilitas penumpang pada lintasan penyeberangan Torobulu–Tampo dipengaruhi oleh faktor waktu dan periode operasional tertentu yang berulang setiap tahun. Adapun analisis ini dapat ditunjukkan pada Gambar 5.



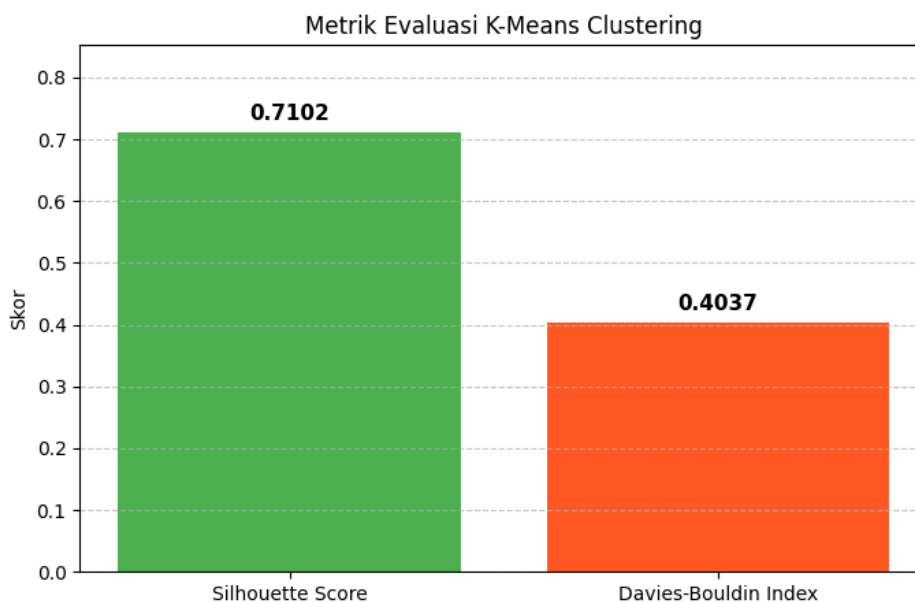
Gambar 5. Visualisasi Temporal Kepadatan Penumpang Harian pada Lintasan Torobulu–Tampo

Berdasarkan visualisasi *heatmap* pada Gambar 5, terlihat bahwa kepadatan penumpang cenderung meningkat pada periode tertentu, terutama menjelang hari besar nasional dan musim mudik. Intensitas warna yang lebih tinggi

menunjukkan konsentrasi jumlah penumpang yang lebih besar pada waktu tertentu. Selain itu, pola peningkatan penumpang pada akhir pekan juga terlihat lebih dominan dibandingkan hari operasional reguler. Visualisasi ini memperkuat hasil clustering bahwa pola mobilitas penumpang pada lintasan Torobulu–Tampo memiliki karakteristik temporal dan musiman yang cukup jelas. Peningkatan intensitas warna juga terlihat berulang pada periode akhir pekan dan musim mudik tahunan.

Temuan penelitian menunjukkan bahwa jadwal operasional kapal yang bersifat tetap berpotensi kurang optimal dalam menghadapi fluktuasi permintaan penumpang pada lintasan penyeberangan Torobulu–Tampo. Pada periode *peak season*, operator pelabuhan dapat mempertimbangkan penambahan frekuensi perjalanan, penyesuaian jam keberangkatan, maupun penyediaan armada tambahan untuk mengurangi kepadatan dan antrean penumpang di pelabuhan. Sebaliknya, pada periode *low traffic*, penyesuaian operasional kapal dapat dilakukan untuk meningkatkan efisiensi penggunaan armada dan biaya operasional. Dengan demikian, hasil *clustering* tidak hanya berfungsi sebagai segmentasi pola kepadatan penumpang, tetapi juga dapat dimanfaatkan sebagai pendukung analisis operasional transportasi penyeberangan berbasis pola historis penumpang guna mendukung pengambilan keputusan operasional yang lebih efektif, efisien, dan responsif terhadap dinamika mobilitas penumpang.

Setelah dilakukan analisis tren data penumpang secara keseluruhan, tahap selanjutnya adalah mengevaluasi kualitas hasil *clustering* yang dihasilkan oleh algoritma *K-Means Clustering*. Evaluasi ini dilakukan untuk mengetahui tingkat pemisahan antar *cluster* serta tingkat kemiripan data dalam setiap *cluster* yang terbentuk. Pada penelitian ini, pengukuran performa *clustering* dilakukan menggunakan metrik *Silhouette Score* dan *Davies-Bouldin Index*. Hasil evaluasi tersebut digunakan untuk menilai apakah proses *clustering* telah mampu menghasilkan pengelompokan data penumpang yang optimal dan representatif terhadap pola kepadatan penumpang pada lintasan penyeberangan Torobulu–Tampo sebagaimana ditunjukkan pada Gambar 6.



Gambar 6. Hasil Evaluasi *Cluster*

Berdasarkan hasil evaluasi *K-Means Clustering* pada Gambar 6, diperoleh nilai *Silhouette Score* sebesar 0,7102 dan *Davies-Bouldin Index* sebesar 0,4037. Nilai *Silhouette Score* yang berada di atas 0,7 menunjukkan bahwa hasil *clustering* memiliki tingkat *cohesiveness* dan *separation* yang baik, dimana data dalam satu *cluster* memiliki kemiripan yang tinggi serta pemisahan antar *cluster* berlangsung secara jelas [16]. Sementara itu, nilai *Davies-Bouldin Index* yang relatif rendah mengindikasikan bahwa tingkat *overlap* antar *cluster* tergolong kecil sehingga proses pengelompokan data dapat merepresentasikan pola distribusi penumpang secara optimal [17]. Hasil evaluasi tersebut menunjukkan bahwa data penumpang berhasil dikelompokkan ke dalam tiga kategori utama, yaitu rendah (*low traffic*), sedang (*normal*), dan tinggi (*peak season*). Dengan demikian, metode *K-Means Clustering* dinilai mampu mengidentifikasi pola kepadatan penumpang berdasarkan karakteristik jumlah penumpang harian pada lintasan penyeberangan Torobulu–Tampo secara representatif dan terukur.

Pemilihan metode *K-Means Clustering* pada penelitian ini didasarkan pada kemampuannya dalam melakukan segmentasi data numerik secara efisien dengan kompleksitas komputasi yang relatif rendah dibandingkan metode *clustering* lainnya seperti *hierarchical clustering* (Taufan & Redjeki, 2025). Selain itu, *K-Means Clustering* juga

memiliki kemampuan yang baik dalam mengidentifikasi pola distribusi data berdasarkan kedekatan karakteristik antar data penumpang.

Hasil evaluasi ini juga memperkuat bahwa pengelompokan data yang terbentuk dapat digunakan sebagai dasar dalam mendukung pengambilan keputusan operasional transportasi penyeberangan. *Cluster* dengan kategori *peak season* dapat menjadi indikator perlunya penambahan frekuensi keberangkatan kapal pada periode tertentu, sedangkan *cluster low traffic* dapat digunakan untuk melakukan penyesuaian jadwal operasional agar lebih efisien. Oleh karena itu, penerapan *K-Means Clustering* tidak hanya mampu menghasilkan segmentasi pola penumpang, tetapi juga memberikan kontribusi dalam pengembangan analisis operasional transportasi berbasis data untuk meningkatkan efektivitas layanan transportasi penyeberangan.

#### 4. Kesimpulan

Penelitian ini berhasil menerapkan metode *K-Means Clustering* untuk mengidentifikasi pola kepadatan penumpang pada lintasan penyeberangan Torobulu–Tampo menggunakan data historis penumpang harian periode 2023–2025. Hasil *clustering* menunjukkan bahwa pola penumpang dapat dikelompokkan ke dalam tiga kategori utama, yaitu rendah (*low traffic*), sedang (*normal*), dan tinggi (*peak season*). Kategori *peak season* didominasi oleh periode hari besar nasional dan musim mudik yang menunjukkan lonjakan jumlah penumpang secara signifikan dibandingkan periode operasional reguler. Selain itu, hasil evaluasi menggunakan *Silhouette Score* dan *Davies-Bouldin Index* menunjukkan bahwa model *clustering* yang dihasilkan memiliki kualitas pengelompokan yang baik dan mampu merepresentasikan karakteristik pola kepadatan penumpang secara optimal. Temuan penelitian menunjukkan bahwa pola mobilitas penumpang pada lintasan penyeberangan Torobulu–Tampo memiliki karakteristik musiman dengan peningkatan jumlah penumpang yang signifikan pada periode tertentu, khususnya saat musim mudik dan hari besar nasional. Hasil *clustering* menggunakan metode *K-Means Clustering* berhasil mengelompokkan pola kepadatan penumpang ke dalam kategori rendah, normal, dan padat sehingga mampu memberikan gambaran yang lebih jelas mengenai dinamika permintaan transportasi penyeberangan berdasarkan data historis. Kondisi tersebut menunjukkan bahwa jadwal operasional kapal yang bersifat statis berpotensi kurang responsif terhadap fluktuasi jumlah penumpang, sehingga diperlukan pendekatan operasional berbasis data untuk mendukung pengambilan keputusan yang lebih adaptif dan efisien. Informasi hasil *clustering* dapat dimanfaatkan sebagai dasar optimalisasi jadwal keberangkatan, penyesuaian frekuensi perjalanan, serta pengelolaan armada kapal sesuai tingkat kepadatan penumpang pada setiap periode operasional. Dengan demikian, penerapan *K-Means Clustering* pada penelitian ini memiliki potensi sebagai pendukung sistem pengambilan keputusan operasional transportasi penyeberangan berbasis data historis guna meningkatkan efektivitas layanan transportasi secara lebih adaptif dan terencana. Penelitian ini masih terbatas pada penggunaan variabel jumlah penumpang harian sehingga belum mempertimbangkan faktor eksternal lain seperti kondisi cuaca, jumlah kendaraan, kapasitas armada, maupun waktu keberangkatan kapal yang dapat memengaruhi dinamika operasional transportasi penyeberangan. Oleh karena itu, hasil *clustering* pada penelitian ini masih difokuskan pada identifikasi pola kepadatan berdasarkan data historis jumlah penumpang. Penelitian selanjutnya dapat dikembangkan dengan menambahkan variabel lain seperti data cuaca, jumlah kendaraan, waktu keberangkatan kapal, serta data hari libur nasional agar hasil *clustering* menjadi lebih akurat dan komprehensif. Selain itu, penelitian berikutnya juga dapat mengombinasikan metode *clustering* dengan metode prediksi atau optimasi jadwal untuk menghasilkan sistem penjadwalan kapal yang lebih cerdas dan dinamis.

#### Referensi

- [1] C. Hutanjalay and A. Hadianti, “Peran Transportasi Laut Dalam Mendukung Mobilitas Masyarakat di Kabupaten Kepulauan Aru,” *Urban Reg. Stud. J.*, vol. 7, no. 2, pp. 244–250, 2025.
- [2] M. F. Almaliki and M. Satyadharna, “Urgensi Ketersediaan dan Integrasi Data Dalam Mendukung Pengambilan Keputusan Operasional Pelabuhan Berbasis Teknologi Informasi,” *Rangas J. Tek. dan Ilmu Terap.*, vol. 1, no. 1, pp. 78–86, 2026.
- [3] R. Riskayanti, Y. Setiawan, K. Khomsiah, A. D. Yogmar, and D. Tribuana, “Implementasi Web-Based Transportation Management System di Industri Transportasi: Sebuah Tinjauan Sistematis Literatur,” *J. Teknol. dan Bisnis Cerdas*, vol. 1, no. 3, pp. 194–209, 2025.
- [4] A. O. Sinaga and R. Passarella, “Analisis Hubungan Antara Jumlah Bus dan Jumlah Penumpang Menggunakan Unsupervised Learning: Analysis of Passenger and Bus Numbers Using Unsupervised Learning,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 3, pp. 1029–1036, 2025.
- [5] A. Maimunah, D. E. M. Sari, F. Julianto, and A. B. Susetyo, “Peramalan Produksi Barang Angkutan Kereta Api Nasional Periode 2019-2026 Menggunakan Model ARIMA,” *J. Media Akad.*, vol. 3, no. 12, pp. 1–17, 2025.
- [6] M. Davi and E. Winarko, “Rancang Bangun Aplikasi Peramalan Jumlah Penumpang Menggunakan Long Short-Term Memory (LSTM),” *Infotekmesin*, vol. 14, no. 2, pp. 303–310, 2023.
- [7] A. Khairunnisa and S. Alfarizi, “Strategi Kebijakan Transportasi Umum di Jawa Barat Berdasarkan Hasil Clustering,” *JURIST J. Ilmu Huk. dan Ilmu Polit.*, vol. 3, no. 1, pp. 25–34, 2026.
- [8] A. F. Muhibuddin, D. S. A. Nur, A. T. Fada, U. H. Rizqiyah, and F. Firmawati, “Prediksi Kebutuhan Kapasitas Dermaga Berdasarkan Tren Perubahan Pola Penyeberangan,” *Urban Reg. Stud. J.*, vol. 7, no. 2, pp. 81–85, 2025.

- [9] B. J. A. Huwaisalt, D. A. Rakhman, and D. Febriani, "Prediksi Jumlah Penumpang Pesawat Domestik di Bandara Soekarno-Hatta dengan Metode SARIMA untuk Mendukung Pengelolaan Kebijakan Transportasi Publik," *Digit. POLICY INSIGHTS Adv. Data Min. Digit. Gov.*, vol. 1, pp. 132–144, 2025.
- [10] M. F. Almaliki, M. Satyadharma, M. Rajulan, H. Hado, G. A. Sulisty, and S. Mursalim, "Analisis Time Series Volume Penumpang Untuk Mendukung Pengambilan Keputusan Manajerial Pada Sektor Transportasi Udara Dalam Negeri," *J. Manaj. dan Penelit. Akunt.*, vol. 19, no. 1, pp. 134–143, 2026.
- [11] D. A. Setyawan, A. Purwatiningsih, and F. S. Budi, "Analisis Pola Lalu Lintas Kapal Selat Bali Berbasis AIS, K-Means, Traffic Flow," *J. FASILKOM*, vol. 15, no. 3, pp. 621–629, 2025.
- [12] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *J. Ilm. Inform. dan Ilmu Komput.*, vol. 3, no. 1, pp. 46–56, 2024.
- [13] C. Aurelya and Y. Widjaja, "Analisis Pengelompokan Wilayah Berdasarkan Frekuensi Kejadian Banjir Menggunakan K-Means Clustering," *Bull. Comput. Sci. Res.*, vol. 6, no. 1, pp. 259–267, 2025.
- [14] D. Anggraini and M. S. Hasibuan, "Initial Centroid Pada Algoritma K-Means dan K-Medoids," *J. Multimed. Dan Teknol. Inf.*, vol. 7, no. 03, pp. 487–500, 2025.
- [15] N. A. Maori and E. Evanita, "Metode elbow dalam optimasi jumlah cluster pada k-means clustering," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 2, pp. 277–288, 2023.
- [16] C. T. N. Handayani and M. N. Sitokdana, "Comparison Of K-Means++ and Agglomerative Hierarchical Methods in Clustering Healthcare Workers," *INOVTEK Polbeng-Seri Inform.*, vol. 10, no. 2, pp. 717–728, 2025.
- [17] M. D. Ananda, M. Mardiah, A. F. N. Masruriyah, and K. N. Malik, "Studi Komparatif Algoritma K-Means dan K-Medoids untuk Segmentasi Informasi Kesehatan," *Comput. Sci.*, vol. 5, no. 2, pp. 103–112, 2025.