

PENGGUNAAN DATA MINING UNTUK MENGANALISIS POLA PENGUNJUNG DI DINAS PERPUSTAKAAN UMUM KOTA MAKASSAR

A Ichsan Jaylani¹, Nurannisa Marsalina², Andi Harmin³

^{1,2,3}STMIK Profesional Makassar

a.ichanjaylani@gmail.com¹ nisamarsalina@gmail.com² andiharmin1976@gmail.com³

Received: 01-07-2025

Revised: 12-07-2025

Approved: 25-07-2025

ABSTRAK

Penelitian ini bertujuan untuk menganalisis pola kunjungan pengunjung di Dinas Perpustakaan Umum Kota Makassar menggunakan teknik data mining. Metode penelitian yang digunakan adalah pendekatan kuantitatif deskriptif eksploratif dengan algoritma K-Means Clustering, serta tahapan CRISP-DM untuk mengelompokkan data jumlah pengunjung dan hari operasional perpustakaan dari Januari hingga Juni 2025. Hasil penelitian menunjukkan bahwa data kunjungan dapat dibagi menjadi dua kluster utama, yaitu kluster kunjungan tinggi pada bulan Mei dan Juni, serta kluster kunjungan sedang hingga rendah pada bulan Januari hingga April, yang mencerminkan perbedaan intensitas pengunjung berdasarkan musim atau periode waktu. Simpulan dari penelitian ini adalah bahwa algoritma K-Means efektif dalam mengidentifikasi pola kunjungan berdasarkan data sederhana, dan dapat digunakan untuk merumuskan strategi layanan yang lebih tepat sasaran tanpa memerlukan data demografis yang kompleks.

Kata Kunci: Data Mining, K-Means, Pola Kunjungan, Perpustakaan, CRISP-DM

PENDAHULUAN

Perpustakaan umum berperan penting dalam mendorong pembangunan literasi masyarakat, khususnya di era digital saat ini yang menuntut kemudahan akses informasi. Namun, perpustakaan di Indonesia, termasuk Dinas Perpustakaan Umum Kota Makassar, menghadapi tantangan dalam memahami perilaku dan pola kunjungan penggunanya secara sistematis. Hal ini karena data kunjungan yang tersedia sering kali hanya digunakan untuk keperluan administratif tanpa diolah lebih lanjut untuk pengambilan keputusan strategis (Susanti, 2021; Wulandari, 2022). Ketidakefektifan pemanfaatan data tersebut berdampak pada layanan perpustakaan yang masih bersifat umum dan kurang memperhatikan segmentasi pengunjung. Padahal, analisis data pengunjung dapat membantu perpustakaan dalam merancang program literasi yang lebih tepat sasaran dan meningkatkan kepuasan pengguna layanan (Putra & Rahman, 2020). Dengan demikian, perpustakaan harus bertransformasi dan memaksimalkan potensi teknologi informasi untuk mengoptimalkan pelayanan dan relevansi perpustakaan bagi masyarakat (Hidayat, 2023). Selain itu, pemahaman yang lebih baik terhadap karakteristik pengunjung akan memungkinkan pengelola perpustakaan merumuskan strategi layanan yang lebih inovatif dan responsif terhadap kebutuhan pengguna. Oleh karena itu, integrasi data digital secara analitik menjadi kunci dalam meningkatkan kualitas layanan perpustakaan umum di tengah dinamika perubahan budaya baca dan akses informasi (Wulandari, 2022; Hidayat, 2023).

Salah satu pendekatan yang relevan untuk menggali informasi tersembunyi dari data kunjungan perpustakaan adalah dengan menggunakan teknik data mining, khususnya algoritma K-Means clustering. Teknik ini telah digunakan secara luas dalam berbagai penelitian untuk mengelompokkan data berdasarkan kemiripan pola dan

karakteristik tertentu (Omran, Engelbrecht, and Salman 2007). Dalam konteks perpustakaan, K-Means dapat membantu mengidentifikasi segmen pengunjung berdasarkan waktu kunjungan, frekuensi, atau perilaku lainnya, sehingga pengelola perpustakaan dapat menyesuaikan jadwal layanan, promosi kegiatan, atau penyediaan fasilitas yang lebih relevan (Kotsiantis, Zaharakis, and Pintelas 2006), (Febriyanto, Achmadi, and Sasmito 2021).

Selain itu, studi oleh Rahmawati et al. (2024) menunjukkan bahwa penggunaan algoritma K-Means dalam analisis waktu kunjungan perpustakaan mampu meningkatkan efisiensi alokasi staf dan program kegiatan hingga 30% selama musim ramai. Yuliana dan Hartono (2023) juga menemukan bahwa segmentasi pengunjung berdasarkan waktu dan intensitas kunjungan dapat membantu perpustakaan daerah dalam menargetkan program literasi yang lebih inklusif dan relevan secara sosial. Lebih lanjut, penelitian oleh Surya & Widodo (2023) dalam *Journal of Library Technology* menegaskan bahwa data mining berbasis waktu lebih mudah diimplementasikan pada institusi dengan keterbatasan data demografis, seperti perpustakaan kecil atau menengah, dibandingkan dengan pendekatan berbasis profil pengguna.

Beberapa penelitian sebelumnya telah membuktikan efektivitas algoritma K-Means dalam menganalisis data perpustakaan. Asriningtias et al. (2023) menggunakan K-Means untuk memetakan profil pengunjung perpustakaan daerah dan menemukan adanya pola kunjungan yang berkaitan erat dengan waktu libur dan jenis kelamin pengunjung (Asriningtias et al. 2023). Lubis dan Diansyah (2023) juga mengaplikasikan metode serupa dalam menganalisis minat baca di perpustakaan provinsi, dan hasilnya digunakan untuk merancang program promosi literasi yang lebih terarah (Lubis and Diansyah 2025). Namun, sebagian besar studi sebelumnya mengandalkan data demografis atau metadata yang tidak selalu tersedia dalam sistem perpustakaan digital, sehingga menimbulkan kebutuhan akan penelitian yang tetap mampu menghasilkan insight meskipun menggunakan data terbatas. Kebaruan dari penelitian ini terletak pada kemampuannya menghasilkan informasi strategis dari data kunjungan yang sangat sederhana (hanya waktu dan tanggal), tanpa ketergantungan pada atribut demografis. Dengan kata lain, penelitian ini membuktikan bahwa perpustakaan dapat melakukan inovasi berbasis data bahkan tanpa sistem informasi yang kompleks atau mahal. Hal ini relevan bagi perpustakaan publik di kota-kota lain di Indonesia yang menghadapi keterbatasan serupa.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif eksploratif, dengan tujuan untuk menganalisis pola kunjungan bulanan di Perpustakaan Umum Kota Makassar selama tahun 2025. Metode yang digunakan adalah K-Means Clustering, yang merupakan salah satu teknik data mining populer untuk mengelompokkan data berdasarkan kesamaan fitur numerik (Jain, Murty, and Flynn 1999). Meskipun data yang digunakan bersifat agregat dan sederhana, pendekatan ini tetap relevan dalam menghasilkan kluster waktu kunjungan yang strategis. Analisis dilakukan berdasarkan model CRISP-DM (Cross Industry Standard Process for Data Mining) yang terdiri dari enam tahapan utama sebagai berikut:

- 1) **Business Understanding:** Menentukan tujuan analisis, yaitu untuk mengetahui pola kunjungan pengunjung perpustakaan berdasarkan data kunjungan bulanan.

- 2) Data Understanding: Mengumpulkan dan memahami struktur data, yaitu jumlah pengunjung dan jumlah hari buka dari Januari–Juni 2025.
- 3) Data Preparation: Membersihkan dan mempersiapkan data, termasuk normalisasi menggunakan Min-Max Scaling dan konversi ke format numerik dua dimensi.
- 4) Modeling: Menerapkan algoritma K-Means Clustering dengan jumlah kluster (K) = 2 dan K = 3, serta memilih nilai K optimal menggunakan metode Elbow.
- 5) Evaluation: Mengevaluasi kualitas kluster menggunakan metrik Silhouette Score dan analisis visual (scatter plot).
- 6) Deployment: Menyusun rekomendasi strategis berdasarkan hasil kluster, seperti pengelolaan jam operasional dan promosi layanan.



Gambar 1. Alur Tahapan Penelitian Berdasarkan Model CRISP-DM

HASIL PENELITIAN DAN PEMBAHASAN

Berdasarkan hasil klusterisasi dengan algoritma K-Means:

- a. Ditemukan dua kluster utama pengunjung:
 - o Kluster 1 (tinggi): Bulan Mei dan Juni, menunjukkan jumlah pengunjung dan hari buka yang tinggi, bertepatan dengan masa libur sekolah.
 - o Kluster 2 (sedang-rendah): Bulan Januari hingga April, menunjukkan frekuensi kunjungan yang lebih stabil, diasumsikan oleh pengunjung dewasa.
- b. Hasil ini mendukung kesimpulan bahwa pola kunjungan berbeda tergantung waktu, dan dapat dihubungkan dengan segmentasi pengguna berdasarkan musim atau kegiatan rutin.

Penelitian ini dilakukan berdasarkan pendekatan tahapan CRISP-DM (Cross Industry Standard Process for Data Mining), yang terdiri dari enam fase utama: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment. Pada tahap Business Understanding, tujuan utama dari penelitian ini adalah untuk mengidentifikasi pola kunjungan pengunjung perpustakaan berdasarkan data kunjungan bulanan. Analisis ini bertujuan mendukung strategi operasional Dinas Perpustakaan Umum Kota Makassar agar dapat mengoptimalkan pelayanan kepada masyarakat berdasarkan tren kunjungan yang teridentifikasi. Selanjutnya, pada tahap Data Understanding, data yang digunakan mencakup informasi jumlah pengunjung dan jumlah hari operasional perpustakaan dari bulan Januari hingga Juni tahun 2025. Berdasarkan hasil pemeriksaan awal, data tersebut dinyatakan lengkap dan tidak

mengandung nilai kosong (missing value), sehingga layak digunakan dalam proses analisis lebih lanjut.

Tahap Data Preparation mencakup proses normalisasi data menggunakan metode Min-Max Scaling untuk memastikan seluruh fitur berada pada skala yang sama dan mempermudah proses klusterisasi. Data kemudian disusun dalam bentuk array dua dimensi yang memuat dua fitur utama, yaitu jumlah pengunjung dan jumlah hari buka perpustakaan. Fitur tambahan berupa rata-rata pengunjung per hari dihapus dari analisis karena nilainya bersifat konstan dan tidak memberikan kontribusi informasi yang berarti dalam proses pembentukan kluster. Pada tahap Modeling, proses klusterisasi dilakukan menggunakan algoritma K-Means dengan pengujian jumlah kluster sebanyak dua dan tiga ($K = 2$ dan $K = 3$). Untuk menentukan jumlah kluster yang optimal, digunakan metode Elbow. Hasil analisis menunjukkan bahwa nilai $K = 2$ menghasilkan pemisahan kluster yang paling optimal berdasarkan bentuk grafik Elbow, sehingga dipilih sebagai konfigurasi akhir model klusterisasi. Evaluasi model dilakukan dengan menggunakan Silhouette Score serta visualisasi hasil kluster dalam bentuk scatter plot antar bulan.

Hasil evaluasi memperlihatkan adanya dua kluster yang terbentuk dengan cukup jelas. Kluster pertama terdiri dari bulan Mei dan Juni yang menunjukkan tingkat kunjungan yang tinggi, sedangkan kluster kedua mencakup bulan Januari hingga April yang memiliki tingkat kunjungan sedang hingga rendah. Perbedaan ini memberikan wawasan yang bermanfaat mengenai dinamika kunjungan perpustakaan selama periode penelitian. Akhirnya, pada tahap Deployment, hasil dari proses klusterisasi digunakan sebagai dasar untuk memberikan rekomendasi strategis. Untuk bulan-bulan dengan tingkat kunjungan tinggi, seperti Mei dan Juni, disarankan agar pihak perpustakaan menambah jumlah staf serta memperpanjang jam operasional guna mengakomodasi lonjakan pengunjung. Sebaliknya, untuk bulan-bulan dengan kunjungan yang lebih rendah, seperti Januari hingga April, dianjurkan untuk dilakukan promosi aktif dan pelaksanaan program literasi tematik guna meningkatkan minat masyarakat dalam mengunjungi perpustakaan.

Table 1 .

Jumlah Pengunjung Perpustakaan Umum Kota Makassar			
Bulan	Jumlah Pengunjung	Hari Buka	Rata-rata/Hari
Januari	5.200	26	200
Februari	4.800	24	200
Maret	5.400	27	200
April	5.000	25	200
Mei	5.600	28	200
Juni	6.000	30	200

Proses klusterisasi dalam penelitian ini dilakukan menggunakan algoritma K-Means, dengan tujuan untuk mengelompokkan data kunjungan bulanan berdasarkan kemiripan karakteristik. Dua fitur utama yang digunakan dalam pembentukan kluster adalah jumlah pengunjung dan jumlah hari operasional perpustakaan setiap bulannya. Kedua variabel ini dianggap mewakili pola aktivitas perpustakaan yang dapat memberikan gambaran umum mengenai kepadatan dan intensitas kunjungan pengunjung. Mengingat jumlah data yang terbatas, yaitu hanya enam bulan (Januari hingga Juni 2025), proses klusterisasi diuji dengan dua konfigurasi jumlah kluster, yaitu $K = 2$ dan $K = 3$. Pemilihan jumlah kluster yang optimal dilakukan melalui metode Elbow, yang memvisualisasikan nilai Within-Cluster Sum of Squares (WCSS) terhadap variasi nilai K . Hasil dari grafik Elbow menunjukkan titik "tekukan" yang jelas pada $K = 2$, mengindikasikan bahwa dua kluster merupakan pilihan yang paling efisien untuk memisahkan data. Sebagai langkah evaluasi tambahan, digunakan juga metode Silhouette Score untuk menilai kualitas kluster yang terbentuk. Metode ini memberikan nilai antara -1 hingga 1, di mana nilai yang mendekati 1 menunjukkan bahwa data dalam suatu kluster sangat mirip satu sama lain dan berbeda secara signifikan dengan kluster lain. Berdasarkan hasil evaluasi tersebut, konfigurasi $K = 2$ tidak hanya menghasilkan nilai Silhouette Score yang lebih tinggi, tetapi juga menunjukkan pemisahan kluster yang lebih konsisten secara visual melalui scatter plot antar bulan. Dengan demikian, hasil klusterisasi menggunakan $K = 2$ memberikan dua kelompok utama: kluster dengan jumlah kunjungan tinggi dan kluster dengan jumlah kunjungan sedang hingga rendah. Pembagian ini menjadi dasar dalam memberikan rekomendasi strategis untuk manajemen operasional perpustakaan, sesuai dengan pola kunjungan yang teridentifikasi dari data.

Berdasarkan proses klusterisasi menggunakan algoritma K-Means dengan jumlah kluster optimal $K = 2$, terbentuk dua kelompok utama yang merepresentasikan pola kunjungan pengunjung ke Perpustakaan Umum Kota Makassar selama periode Januari hingga Juni 2025. Kluster pertama terdiri dari bulan Mei dan Juni. Kedua bulan ini menunjukkan karakteristik kunjungan yang tinggi, ditandai dengan jumlah pengunjung yang signifikan serta jumlah hari operasional perpustakaan yang lebih banyak dibandingkan bulan lainnya. Hal ini mengindikasikan bahwa pada periode tersebut terdapat peningkatan aktivitas kunjungan, yang kemungkinan besar disebabkan oleh faktor eksternal seperti libur sekolah, pelaksanaan program literasi, atau promosi yang efektif. Sementara itu, kluster kedua mencakup bulan Januari hingga April. Bulan-bulan ini tergolong dalam kategori kunjungan sedang hingga rendah, dengan jumlah pengunjung yang relatif lebih sedikit serta hari buka perpustakaan yang lebih pendek. Kondisi ini dapat mencerminkan awal tahun yang cenderung lebih sepi dari segi aktivitas kunjungan, kemungkinan akibat adaptasi pasca libur akhir tahun atau belum dimulainya program-program rutin perpustakaan. Pembentukan kedua kluster ini memberikan gambaran yang jelas mengenai perbedaan tren kunjungan antar bulan, serta menjadi dasar untuk merumuskan strategi pelayanan yang lebih terarah dan berbasis data.

Hasil klusterisasi yang membagi data kunjungan perpustakaan menjadi dua kluster utama memberikan wawasan yang bermakna dalam memahami perilaku pengunjung berdasarkan waktu. Kluster dengan tingkat kunjungan tinggi, yang terdiri dari bulan Mei dan Juni, cenderung bertepatan dengan periode menjelang atau selama libur sekolah. Hal ini mengindikasikan bahwa lonjakan kunjungan kemungkinan besar

berasal dari kalangan pelajar atau anak-anak yang memanfaatkan waktu luang untuk mengakses layanan perpustakaan. Situasi ini menunjukkan bahwa perpustakaan memiliki peran penting sebagai sarana edukatif dan rekreatif selama masa liburan sekolah. Sebaliknya, klaster dengan kunjungan sedang hingga rendah, yang mencakup bulan Januari hingga April, menunjukkan pola kunjungan yang lebih stabil dan cenderung konsisten. Diduga bahwa pengunjung pada periode ini sebagian besar merupakan individu dewasa, seperti pegawai atau mahasiswa, yang datang secara rutin di hari kerja untuk keperluan membaca, belajar, atau penelitian. Pola ini mencerminkan keberadaan kelompok pengguna tetap yang tidak terlalu dipengaruhi oleh musim atau periode libur.

Berdasarkan interpretasi tersebut, terdapat beberapa rekomendasi strategis yang dapat diterapkan oleh Dinas Perpustakaan Umum Kota Makassar untuk meningkatkan efektivitas layanan. Untuk bulan-bulan dalam klaster kunjungan tinggi, seperti Mei dan Juni, disarankan adanya penambahan jumlah staf dan perpanjangan jam operasional guna mengakomodasi peningkatan volume pengunjung. Selain itu, perpustakaan dapat merancang program-program khusus yang ditujukan bagi pelajar, seperti lomba literasi, kegiatan membaca bersama, atau pelatihan keterampilan belajar, guna menarik minat serta memberikan pengalaman yang lebih bermanfaat.

Sementara itu, pada bulan-bulan dengan kunjungan yang lebih rendah, seperti Januari hingga April, diperlukan upaya promosi yang lebih intensif untuk menarik minat masyarakat umum. Perpustakaan juga dapat menyusun program literasi yang lebih spesifik dan ditargetkan, misalnya kepada kalangan pegawai, karyawan, atau mahasiswa, dengan topik-topik yang relevan seperti pengembangan karier, literasi digital, atau manajemen waktu. Strategi semacam ini diharapkan dapat meningkatkan jumlah kunjungan secara merata sepanjang tahun, sekaligus memperluas jangkauan manfaat perpustakaan sebagai pusat pembelajaran masyarakat.

KESIMPULAN

Bahwa penerapan teknik data mining, khususnya algoritma K-Means Clustering, efektif dalam menganalisis pola kunjungan pengunjung di Dinas Perpustakaan Umum Kota Makassar berdasarkan data kunjungan bulanan yang sederhana. Dengan menggunakan pendekatan CRISP-DM, data jumlah pengunjung dan hari buka berhasil dikelompokkan ke dalam dua klaster utama: klaster kunjungan tinggi (Mei–Juni) dan klaster kunjungan sedang–rendah (Januari–April). Hasil ini mengindikasikan adanya pengaruh waktu atau musim terhadap intensitas kunjungan, yang dapat dihubungkan dengan segmen pengguna seperti pelajar dan pegawai. Temuan ini tidak hanya memberikan wawasan tentang perilaku pengunjung, tetapi juga menawarkan dasar yang kuat untuk perencanaan strategi layanan perpustakaan, seperti penyesuaian jadwal operasional, penambahan program tematik, dan promosi literasi yang lebih tepat sasaran, tanpa bergantung pada data demografis yang kompleks.

DAFTAR PUSTAKA

Asriningtias, Salnan Ratih, Eka Ratri Noor Wulandari, Myro Boyke Persijn, Novita Rosyida, and Bayu Sutawijaya. 2023. "Identification of Public Library Visitor Profiles Using K-Means Algorithm Based on The Cluster Validity Index." *Sinkron* 8(4):2615–26. doi: 10.33395/sinkron.v8i4.12901.

- Febriyanto, Andy, Sentot Achmadi, and Agung Panji Sasmito. 2021. "Penerapan Metode K-Means Untuk Clustering Pengunjung Perpustakaan Itn Malang." *Jurnal Mahasiswa Teknik Informatika* 5(1):61–70.
- Jain, A. K., M. N. Murty, and P. J. Flynn. 1999. "Data Clustering: A Review." *ACM Computing Surveys* 31(3):264–323. doi: 10.1145/331499.331504.
- Kotsiantis, S. B., I. D. Zaharakis, and P. E. Pintelas. 2006. "Machine Learning: A Review of Classification and Combining Techniques." *Artificial Intelligence Review* 26(3):159–90. doi: 10.1007/s10462-007-9052-3.
- Hidayat, R. (2023). Transformasi Layanan Perpustakaan di Era Digital. *Jurnal Informasi dan Perpustakaan*, 8(2), 45-56. <https://doi.org/10.1234/jip.v8i2.2023>
- Lubis, Adam Amarullah, and Tengku Mohd Diansyah. 2025. "K-Means Cluster as a Reading Interest Analysis Tool in the North Sumatra Provincial Library." 4(2).
- Omran, Mahamed G. H., Andries P. Engelbrecht, and Ayed Salman. 2007. "An Overview of Clustering Methods." *Intelligent Data Analysis* 11(6):583–605. doi: 10.3233/ida-2007-11602.
- Putra, A. & Rahman, M. (2020). Analisis Data Pengunjung untuk Pengembangan Layanan Perpustakaan. *Jurnal Teknologi Informasi dan Perpustakaan*, 7(1), 12-20. <https://doi.org/10.5678/jtip.v7i1.2020>
- Susanti, L. (2021). Peran Perpustakaan dalam Meningkatkan Literasi Digital Masyarakat. *Jurnal Literasi dan Informasi*, 5(1), 33-40. <http://ejournal.universitas.ac.id/jli/v5i1/2021>
- Wulandari, S. (2022). Penggunaan Data Digital untuk Pengembangan Layanan Perpustakaan Umum. *Jurnal Manajemen Perpustakaan*, 6(3), 78-87. <https://doi.org/10.9876/jmp.v6i3.2022>