

artikel rita.pdf

by . .

Submission date: 23-Mar-2025 02:41PM (UTC+0800)

Submission ID: 2622280718

File name: artikel_rita.pdf (247.87K)

Word count: 2986

Character count: 17854

Pengelompokan Bulan Berdasarkan Data Jenis Surat Tahun 2023 di UPTD/SAMSAT Kabupaten Bungo Menggunakan Analisis Metode K-Means Clustering

Nurmarita¹, Gusmi Kholijah², Wardi Syafmen³

^{1,2,3}Universitas Jambi, Jambi, Indonesia

ABSTRAK

Pengelolaan data menjadi kunci efisiensi pelayanan publik dalam era digital ini, termasuk di UPTD/SAMSAT Kabupaten Bungo yang menangani pajak kendaraan dan pengolahan data surat. Oleh karena itu, untuk memahami pola data surat pada tahun 2023, diterapkan metode K-Means clustering yaitu salah satu metode non hierarki yang bertujuan untuk mengelompokkan data berdasarkan kesamaan karakteristik. Penelitian ini mengelompokkan bulan berdasarkan jenis surat ke dalam 3 cluster. Cluster 1 terdiri dari tiga bulan yaitu September, November, dan Desember yang memiliki karakteristik aktivitas surat tertinggi. Cluster 2 terdiri dari 7 bulan yaitu Januari, Februari, April, Mei, Juni, Agustus, dan Oktober yang memiliki karakteristik aktivitas surat yang standar. Sedangkan pada cluster 3 terdiri dari dua bulan yaitu Maret dan Juli yang memiliki karakteristik jumlah surat masuk tinggi dibanding cluster lain. Hasil ini menunjukkan bahwa metode K-means clustering efektif dalam mengidentifikasi pola dalam setiap bulan serta memberikan wawasan bagi SAMSAT Kabupaten Bungo untuk meningkatkan pengelolaan surat dan pelayanan publik secara lebih optimal.

ABSTRACT

Data management is the key to efficient public services in this digital era, including in the UPTD/SAMSAT of Bungo Regency which handles vehicle taxes and letter data processing. Therefore, to understand the pattern of letter data in 2023, the K-Means clustering method is applied, which is a non-hierarchical method that aims to group data based on similar characteristics. This study groups months based on letter type into 3 clusters. Cluster 1 consists of three months, namely September, November, and December, which have the characteristics of the highest letter activity. Cluster 2 consists of 7 months, namely January, February, April, May, June, August, and October, which have standard letter activity characteristics. While cluster 3 consists of two months, namely March and July, which have the characteristics of a high number of incoming letters compared to other clusters. These results show that the K-means clustering method is effective in identifying patterns in each month and providing insight for SAMSAT of Bungo Regency to improve letter management and public services more optimally.

Kata Kunci: Cluster, Jenis Surat, K-Means, UPTD/SAMSAT, Kabupaten Bungo

Email Address: ¹nurmarita@gmail.com

DOI: <http://dx.doi.org/10.30829/jistech.v9i2.xxxx>

Received 10 February 2025; Received in revised form 10 April 2025; Accepted 15 June 2025



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

Pendahuluan

UPTD/SAMSAT Kabupaten Bungo merupakan salah satu unit pelaksana teknis (UPT) yang bertanggung jawab atas pengelolaan dan pelayanan administrasi kendaraan bermotor. Instansi ini didirikan sebagai bagian dari upaya pemerintah untuk meningkatkan pelayanan publik, terutama yang berkaitan dengan pengelolaan perpajakan dan surat-surat kendaraan. SAMSAT (Sistem Administrasi Satu Atap) Kabupaten Bungo adalah unit pelaksana teknis di bawah naungan Dinas Pendapatan Daerah (DISPENDA) Provinsi Jambi. Adapun melalui SAMSAT ini, sehingga dapat menyelenggarakan pelayanan dan pengkajian bahan petunjuk teknis di bidang pajak dan retribusi daerah (Sutanto, 2020)[1].

UPTD/SAMSAT menjalankan tugas berupa mengelola data administrasi kendaraan bermotor serta mengelola data surat yang masuk dan keluar. Adapun tujuan dari penelitian ini yaitu untuk memahami pola data surat sepanjang tahun 2023 dengan menggunakan metode K-Means clustering. Metode ini digunakan untuk mengelompokkan bulan berdasarkan jenis surat. K-Means clustering adalah salah satu metode analisis data yang digunakan untuk mengelompokkan objek berdasarkan kesamaan karakteristiknya. Metode ini bersifat fleksibel dan dapat diterapkan pada dataset dengan ukuran besar. Metode K-Means clustering bekerja dengan cara

mengelompokkan data ke dalam beberapa cluster berdasarkan jarak antar objek. Semakin dekat jarak antara dua objek, maka semakin besar kemungkinan berada dalam satu cluster. Penelitian ini objek yang dianalisis adalah bulan Januari hingga Desember tahun 2023. Pengelompokan ini bertujuan untuk mengidentifikasi pola dalam data surat, mengevaluasi bulan dengan volume surat tertinggi dan terendah, serta memahami perbedaan distribusi jenis surat di setiap bulan [2].

12 Metodologi Penelitian

Jenis penelitian yang digunakan pada penelitian ini yaitu dengan menggunakan data kuantitatif. Data kuantitatif merupakan data yang berbentuk angka yang dapat diukur secara matematis. Data yang digunakan dalam pengelompokan bulan berdasarkan jenis data surat tahun 2023 di UPTD/SAMSAT Kabupaten Bungo dengan menggunakan metode K-Means clustering yang datanya yaitu terdiri dari 4 jenis data surat. Data tersebut adalah data primer yang diperoleh langsung melalui observasi dan pencatatan jumlah surat yang dikelola oleh kantor SAMSAT dalam periode tertentu. Pada laporan penelitian ini, peneliti ingin mengelompokkan bulan berdasarkan data jenis surat di UPTD/SAMSAT Kabupaten Bungo pada tahun 2023. Terdapat 4 jenis surat yang terdiri dari surat masuk, surat keluar, surat perintah tugas (SPT), dan surat perintah perjalanan dinas (SPPD).

Tabel 1. Variabel penelitian

Notasi Variabel	Variabel	Definisi Variabel
X_1	Surat Masuk	Banyaknya surat yang diterima dalam periode waktu tertentu
X_2	Surat Keluar	Surat resmi yang dibuat dan dikirimkan oleh instansi kepada pihak eksternal
X_3	Surat Perintah Tugas (SPT)	Surat resmi yang dikeluarkan pimpinan instansi untuk memberikan perintah pelaksanaan tugas tertentu kepada pegawai atau staf
X_4	Surat Perintah Perjalanan Dinas (SPPD)	Surat resmi yang berisi perintah kepada pegawai untuk melaksanakan tugas dinas ke lokasi tertentu dengan tujuan tertentu atas nama instansi

Analisis cluster adalah salah satu analisis multivariat yang bertujuan untuk mengelompokkan objek atau subjek berdasarkan kesamaan karakteristik yang dimiliki. Analisis cluster memiliki tingkat homogenitas yang tinggi di dalam satu kelompok (*Within Cluster*) dan tingkat heterogenitas yang tinggi antar kelompok (*Between Cluster*). Jika pengelompokan berhasil, individu dalam satu kelompok akan membentuk gerombolan saat dipetakan sementara objek dari kelompok lain akan terpisah dengan jelas. Analisis ini mengasumsikan bahwa variabel-variabel bersifat bebas (tidak berkorelasi) dan minimal menggunakan skala interval [3].

Algoritma K-Means adalah salah satu metode clustering non hierarki yang digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik. Data dengan karakteristik akan ditempatkan dalam kelompok yang sama. Hal ini memastikan setiap kelompok yang terbentuk memiliki tingkat perbedaan yang minim atau sedikit. Algoritma K-Means merupakan metode clustering berdasarkan pada jarak yang memisahkan data ke dalam sejumlah cluster dan hanya terkait pada atribut numerik [4].

Secara umum, algoritma analisis cluster K-means dimulai dengan langkah berikut [5]:

1. Menentukan nilai k sebagai jumlah cluster yang akan dibentuk.
2. Menentukan jumlah cluster k biasanya didasarkan pada berbagai pertimbangan, termasuk aspek teoritis dan konseptual yang dapat memberikan panduan dalam menentukan jumlah cluster yang optimal.
3. Menentukan k titik pusat cluster (*centroid*) awal secara acak.
4. Pemilihan *centroid* awal dilakukan secara acak dari objek-objek yang ada, sebanyak k cluster. Selanjutnya, untuk menghitung *centroid* cluster ke- i digunakan rumus sebagai berikut:

$$v = \frac{\sum_{i=1}^n x_i}{n} ; i = 1, 2, 3, \dots, n \quad (1)$$

dengan

v : *centroid* pada cluster

x_i : objek ke- i

n : jumlah objek

5. Menentukan jarak antara setiap objek dengan *centroid* masing-masing cluster

Jarak antara objek dan *centroid* dihitung dengan menggunakan rumus jarak Euclidean yaitu sebagai berikut:

$$d(x_i, x_j) = \sqrt{(|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2)} \quad (2)$$

$i = 1, 2, 3, \dots, p ; j = 1, 2, 3, \dots, p$

dengan

$d(x_i, x_j)$: Jarak objek ke- i ke centroid j

x_{i1} : Objek ke-25 pada atribut data ke-1

x_{j1} : Centroid ke-4 pada atribut data ke-1

4. Alokasikan setiap objek ke centroid yang paling dekat

5. Lakukan iterasi 20 untuk menentukan posisi centroid yang baru dengan menggunakan rumus berdasarkan persamaan (1)

6. Ulangi langkah ke 3 dan 4 hingga tidak ada lagi objek yang berpindah cluster

Hasil dan Pembahasan

Pada penelitian ini, akan dilakukan pengelompokan sekumpulan objek berdasarkan variabel yang dianalisis yaitu pengelompokan bulan berdasarkan data jenis surat tahun 2023.

a. Standarisasi data

Standarisasi data diperlukan jika data dalam suatu penelitian memiliki satuan yang bervariasi atau berbeda. Namun, jika semua variabel dalam data memiliki satuan yang sama maka standarisasi tidak diperlukan. Adapun dalam penelitian ini, karena semua variabel memiliki satuan yang sama maka pengujian untuk standarisasi data tidak perlu dilakukan [6].

b. Asumsi sampel mewakili populasi

Asumsi sampel mewakili populasi dapat ditentukan melalui uji Kaiser Mayer Olkin (KMO). Uji KMO dilakukan untuk mengetahui apakah faktor-faktor dalam penelitian valid atau tidak [7]. Secara umum statistik uji KMO yaitu sebagai berikut:

$$r_{x_i x_j} = \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x}_i)(x_j - \bar{x}_j)}{\sqrt{\sum_{i=1}^n (x_i - \bar{x}_i)^2} \sqrt{\sum_{j=1}^n (x_j - \bar{x}_j)^2}}$$

dengan

$r_{x_i x_j}$: Koefisien korelasi antara variabel i dan j

x_i, x_j : Nilai objek pada variabel ke- i dan ke- j

$\bar{x}_i, \bar{x}_j$: Nilai rata-rata variabel i dan j

Berikut akan ditampilkan hasil perhitungan uji korelasi dan kecukupan sampel menggunakan software Rstudio yang disajikan pada gambar di bawah ini:

```
> bart_spher(faktor)

Bartlett's Test of Sphericity

Call: bart_spher(x = faktor)

x2 = 7.357
df = 6
p-value = .28906

> km05(faktor)

Kaiser-Meyer-Olkin Statistics

Call: km05(x = faktor)

Measures of Sampling Adequacy (MSA):
      x1      x2      x3      x4
0.5226035 0.6347531 0.7251013 0.5658144

KMO-Criterion: 0.6206507
```

Gambar 1. KMO and Bartlett's Test

Berdasarkan gambar di atas terlihat bahwa taraf signifikan sebesar 0,620. Nilai tersebut <0,5 yang artinya H_0 tolak sehingga sampel yang digunakan layak untuk dianalisis secara lanjut.

c. Asumsi Non Multikolinearitas

Uji asumsi non multikolinearitas digunakan untuk mengidentifikasi hubungan antar variabel yang digunakan. Pengujian ini dapat dilakukan dengan melihat nilai Variance Inflation Factor (VIF) [8]. Berikut merupakan hasil perhitungan uji multikolinearitas yang dicari menggunakan software Rstudio:

```
> #uji Multikolinearitas
> library(car)
> vif(r1t)

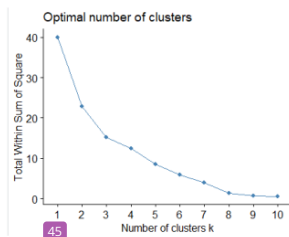
      x1      x2      x3      x4
1.262840 1.698616 1.438505 1.498048
```

Gambar 2. Nilai VIF setiap variabel

Berdasarkan gambar di atas, nilai VIF untuk setiap variabel <10 sehingga dapat disimpulkan bahwa tidak terdapat multikolinearitas pada semua variabel. Hal ini menunjukkan bahwa asumsi dalam analisis cluster K-Means telah terpenuhi.

d. Analisis cluster K-Means

Pada tahap awal yang dilakukan yaitu menentukan jumlah *cluster* yang optimal dan *centroid* awal. Untuk analisis *centroid* awal yang akan digunakan dalam perhitungan dengan metode *K-means* di tentukan jumlah *cluster* optimal yaitu $K = 3$ yang dapat dilihat dari grafik metode *Elbow* berikut ini:



Gambar 3. Grafik Metode *Elbow*.

Berdasarkan grafik metode *Elbow* yang diperoleh dari *output* Rstudio, menunjukkan hubungan antara jumlah *cluster* k dan nilai *Within Sum of Square* (TWSS) untuk setiap nilai K . Dalam konteks untuk analisis *cluster*, jumlah *cluster* optimal ditentukan ketika terjadi penurunan yang signifikan pada nilai eigen, diikuti oleh penurunan yang lebih landai atau stabil. Hal ini menandakan titik "siku" (*Elbow*) di mana nilai eigen tidak lagi turun secara signifikan [9]. Pada grafik di atas titik "Elbow" tampak di sekitar komponen kedua.

Berikut adalah hasil *k-means clustering* yang diperoleh menggunakan bantuan *software* Rstudio:

```
> #mencari hasil k-means clustering
> kclusteringskmeans(data=number1k,centers=3,nstart=25)
> kclusteringsk
k-means clustering with 3 clusters of sizes 3, 7, 2

Cluster means:
  Surat Masuk Surat keluar      SPT      SPPD
1  25.33333      36.00000 25.00000 23.33333
2  22.28571      20.42857 16.57143 13.00000
3   46.50000      15.00000 10.00000 13.50000

Clustering vector:
[1] 2 2 3 2 2 2 3 2 1 2 1 1

within cluster sum of squares by cluster:
[1] 489.3333 534.8571 143.0000
(between_SS / total_SS = 64.6 %)

Available components:
[1] "cluster" "centers" "totss" "withinss" "tot.withinss" "betweenss"
[7] "size"    "iter"    "ifault"
```

Gambar 4. Hasil Analisis *K-Means Clustering*

Pada proses hasil *K-means clustering* di atas dapat diinterpretasikan sebagai berikut:

- Jumlah *cluster* yang optimal menghasilkan 3 *cluster* yang dimana *cluster* 1 terdapat 3 anggota, *cluster* 2 terdapat 7 anggota, dan *cluster* 3 terdapat 2 anggota. Adapun untuk algoritma *K-means* didefinisikan dengan $nstart=25$ yang artinya algoritma dijalankan sebanyak 25 kali dengan inisialisasi acak untuk memastikan hasil yang terbaik.
- Rata-rata setiap *cluster* yang digunakan untuk menggambarkan karakteristik setiap *cluster* berdasarkan nilai-nilai variabelnya. Pada *cluster* 1 memiliki nilai rata-rata tertinggi pada surat masuk, surat keluar, dan juga nilai SPT dan SPPD yang cukup besar. Kemudian pada *cluster* 2 semua variabel memiliki nilai rata-rata yang lebih rendah dibanding *cluster* lainnya. Sedangkan pada *cluster* 3 memiliki nilai rata-rata sangat tinggi pada surat masuk, tetapi memiliki nilai rata-rata rendah pada surat keluar. Adapun perhitungan untuk menentukan nilai rata-rata dari 4 variabel yang terdapat dalam 3 *cluster* tersebut bisa digunakan rumus umum sebagai berikut:

$$\bar{X} = \mu + z \cdot \sigma$$
- Pada hasil vector *clustering* menunjukkan bahwa data pertama masuk ke *cluster* 2, data kedua masuk ke *cluster* 2 dan data ketiga hingga data ke sembilan dikelompokkan sesuai dengan urutan pada vektor yang ditampilkan. Vektor ini menunjukkan hasil klasterisasi, dimana setiap angka mempresentasikan *cluster* tempat sebuah data berada.
- Within Cluster Sum of Squares (WCSS) menunjukkan seberapa jauh data dalam *cluster* dari pusat *cluster* (*centroid*). Adapun nilai WCSS untuk *cluster* 1 yaitu sebesar 489,3333, sedangkan pada *cluster* 2 sebesar 534,8571 dan *cluster* 3 sebesar 143,000. Nilai yang lebih kecil menunjukkan bahwa data dalam *cluster* tersebut memiliki kedekatan yang lebih tinggi terhadap pusat *cluster*. Pada *cluster* 3 memiliki nilai WCSS terkecil, artinya penyebaran datanya paling kecil dibanding *cluster* lain. Sedangkan pada *cluster* 2 memiliki nilai WCSS terbesar, yang menunjukkan penyebaran lebih luas dalam *cluster* tersebut.
- Persentase Between-Cluster Sum of Squares (BCSS) terhadap Total Sum of Squares (TSS) adalah 64.6%. Persentase ini menunjukkan seberapa baik *cluster* terpisah satu sama lain. Semakin tinggi nilainya, semakin

baik pemisahan antar *cluster*. Nilai 64,6% menunjukkan pemisahan yang cukup baik, meskipun masih bisa ditingkatkan untuk mencapai *cluster* yang lebih terdefinisi.

Hasil k-means *clustering* menunjukkan bahwa data dikelompokkan menjadi tiga *cluster* dengan karakteristik dan ukuran yang berbeda. *Cluster* 3 memiliki data yang paling rapat di sekitar *centroid* (WCSS terkecil), sedangkan *cluster* 2 memiliki data yang lebih tersebar (WCSS terbesar). Pemisahan antar *cluster* cukup baik dengan BCSS sebesar 64,6%. Rata-rata nilai dari masing-masing *cluster* memberikan gambaran pola data yang dapat digunakan untuk analisis lebih lanjut atau pengambilan keputusan.

Selanjutnya untuk mengetahui bulan mana saja di tahun 2023 berdasarkan jumlah surat masuk, surat keluar SPT dan SPPD yang berada pada *cluster* 1, *cluster* 2 dan *cluster* 3, disajikan pada tabel berikut:

Tabel 2. Keanggotaan Cluster

No	Bulan	Surat Masuk	Surat Keluar	SPT	SPPD	Cluster
1	Januari	19	19	22	15	2
2	Februari	20	21	15	12	2
3	Maret	51	15	7	20	3
4	April	19	17	12	15	2
5	Mei	20	19	9	9	2
6	Juni	33	20	16	12	2
7	Juli	42	15	13	7	3
8	Agustus	6	28	24	10	2
9	September	25	47	17	21	1
10	Oktober	29	19	18	18	2
11	November	33	28	32	19	1
12	Desember	18	33	26	30	1

Berdasarkan tabel 2 diatas, dapat dilihat bahwa terdapat 3 anggota bulan yang berada pada *cluster* 1 yaitu bulan September, November dan Desember, *cluster* ini memiliki aktivitas surat yang paling banyak atau paling tinggi dibanding dari *cluster* yang lain. Kemudian pada *cluster* 2 yang terdiri dari 7 anggota bulan yaitu Januari, Februari, April, Mei, Juni, Agustus dan Oktober, pada *cluster* ini menunjukkan aktivitas surat dengan volume yang sedang atau normal tanpa adanya nilai ekstrem. *Cluster* ini menunjukkan aktivitas surat menyurat masih berada dalam jumlah yang standar, tidak terlalu tinggi dan tidak terlalu rendah. Sedangkan pada *cluster* 3 yang terdiri dari 2 anggota yaitu bulan Maret dan Juli yang menunjukkan data yang memiliki jumlah surat masuk yang tinggi dibandingkan dengan *cluster* lainnya. Misalnya pada bulan Maret dengan data surat masuk yang berjumlah 51 surat, jauh lebih tinggi dibandingkan dengan jumlah surat pada bulan-bulan lainnya yang sebagian besar memiliki nilai dibawah 30.

Adapun untuk visualisasi pada *cluster* yang telah terbentuk dapat dilihat pada plot *cluster* berikut ini:



Gambar 5. Plot Cluster

Berdasarkan gambar plot *cluster* diatas, dapat diinterpretasikan sebagai berikut:

1. Pada *cluster* 1 yang berwarna merah lingkaran terlihat bahwa titik-titik *cluster* ini terletak di bagian kiri bawah plot, dimana anggota yang termasuk ke dalam *cluster* ini yaitu bulan September, November dan Desember. Anggota *cluster* terlihat berdekatan, menunjukkan bahwa bulan-bulan dalam *cluster* ini memiliki kesamaan yang kuat.
2. Pada *cluster* 2 yang berwarna hijau segitiga terlihat bahwa titik tersebut berada di bagian tengah. Adapun anggota yang termasuk ke dalam *cluster* ini yaitu bulan Januari, Februari, April, Mei, Juni, Agustus dan Oktober. Variasi titik dalam *cluster* ini relatif lebih besar dibandingkan *cluster* lain, yang dapat menunjukkan keragaman karakteristik dalam *cluster*.

3. Pada *cluster* 3 yang berwarna biru persegi yang beranggotakan bulan Maret dan Juli menunjukkan bahwa titik-titik ini terletak jauh di sisi kanan bawah, terpisah dari *cluster* lain, hal ini menunjukkan bahwa *cluster* ini sangat berbeda dari yang lainnya.

Dapat disimpulkan bahwa *cluster* 1 memiliki kesamaan yang kuat, sedangkan pada *cluster* 2 menunjukkan keragaman yang lebih besar dibandingkan *cluster* lainnya. Kemudian pada *cluster* 3 sangat terisolasi dan berbeda secara signifikan dari dua *cluster* lainnya, hal ini menunjukkan kedua bulan tersebut memiliki sifat yang unik.

Kesimpulan

Hasil penelitian ini bermanfaat untuk meningkatkan efisiensi pengelolaan surat, seperti dengan mengimplementasikan sistem digital dan otomatisasi dalam pengarsipan serta pemrosesan surat. Penggunaan metode seperti analisis *cluster* juga dapat membantu dalam pengelompokan surat yang lebih terstruktur. Selain itu, pelatihan dan sosialisasi kepada pegawai mengenai pentingnya sistem yang terorganisir serta komunikasi antarinstansi yang lebih baik sangat diperlukan agar sistem administrasi surat dapat berjalan dengan optimal.

5 Ucapan Terima Kasih

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam penyelesaian penelitian ini. Berkat bantuan, bimbingan dan dukungan, baik secara langsung maupun tidak langsung, sehingga penelitian ini dapat diselesaikan dengan baik. Dukungan, kerja sama, dan kontribusi yang diberikan sangat berarti bagi keberhasilan penelitian ini, dan penulis sangat menghargai setiap bentuk perhatian yang telah diberikan.

Daftar Pustaka

- [1] B. Sutanto, "Manajemen Pelayanan Publik dan Administrasi Pemerintahan," Yogyakarta: Andi Publisher, 2020.
- [2] R. Hartono, "Implementasi Algoritma K-Means dalam Pengelompokan Data Surat Masuk di Dinas Komunikasi dan Informatika," *Sist. Inf.*, vol. 17, no. 3, hal. 89–97, 2021.
- [3] R. E. A. Joseph F. Hair Jr., William C. Black, Barry J. Babin, "Multivariate Data Analysis.pdf," 2010.
- [4] D. Putra, *Clustering Data dengan Algoritma K-Means*. Bandung: Informatika, 2023.
- [5] E. Rahmawati, *Pemanfaatan Algoritma K-Means dalam Pengelompokan Data*. Surabaya: Techno Data Press, 2024.
- [6] R. Fadilah, *Standarisasi Data dalam Analisis Statistik dan Machine Learning*. Jakarta: Data Science Press, 2024.
- [7] R. Hidayat, "Analisis Faktor dengan Uji KMO dan Bartlett dalam Penelitian Statistik," *Stat. Terap.*, 2020.
- [8] R. Susanto, *Analisis Statistik dengan Uji Multikolinearitas*. Jakarta: Statistika Press, 2024.
- [9] P. J. Kaufman, L., & Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley-Interscience, 2005.
- [10] S. M. Bungo, "Laporan Tahunan UPTD/Samsat Muara Bungo," Jambi, 2023.
- [11] F. A., "Perancangan Sistem Informasi Surat Perintah Perjalanan Dinas Berbasis Web Pada Kantor Imigrasi Kelas II TPI Belawan," Sumatera Utara: Universitas Muhammadiyah, 2023.
- [12] S. O. D. . Aprilia, R., Afsari, K., Rahma, R., Nasution, N., & Putri, "Analisis Cluster Dengan Metode K-Means Pada Jenis Data Surat Di BPPRD Sumatera Utara Amaliah," *Pengabd. Kpd. Masy.*, vol. 6, no. 2, hal. 369–373, 2022.
- [13] J. Sarwono, *Analisis Jalur untuk Riset Bisnis dengan SPSS*. Yogyakarta: Andi Offset, 2007.
- [14] J. Wu, *Advances in K-Means Clustering: A Data Mining Thinking*. Berlin: Springer-Verlag Berlin dan Heidelberg, 2012.
- [15] I. Purnama, "Akuntabilitas Penggunaan Surat Perintah Perjalanan Dinas (SPPD) Pegawai di Sekretariat DPRD Kota Makassar," Makassar: Universitas Muhammadiyah Makassar, 2017.

ORIGINALITY REPORT

24%

SIMILARITY INDEX

21%

INTERNET SOURCES

15%

PUBLICATIONS

9%

STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Universitas Lancang Kuning Student Paper	2%
2	repository.unja.ac.id Internet Source	1%
3	journal.unhas.ac.id Internet Source	1%
4	123dok.com Internet Source	1%
5	Mohammad Akbar Ilham, Sentot Achmadi, Karina Aulia Sari. "Implementation of Double Exponential Smoothing Method for Musical Instrument Sales Forecasting System", Jurnal Teknologi Terpadu, 2024 Publication	1%
6	ojs.uajy.ac.id Internet Source	1%
7	Aan Ardhiyanto, Dwi Hartanti, Joni Maulindar. "Implementasi Algoritma K-Means Untuk Rekomendasi Pengadaan Buku", Infotek: Jurnal Informatika dan Teknologi, 2025 Publication	1%
8	jurnal.uinsu.ac.id Internet Source	1%
9	es.scribd.com Internet Source	1%
10	Agnes Mariana Veronica Sandag, Charles E. Mongi, Mans L. Mananohas. "Pengelompokan Sekolah Menengah Atas (SMA) di Kabupaten	1%

Minahasa Tenggara Berdasarkan Stadar
Kompetensi Lulusan Tahun 2018
Menggunakan Analisis Gerombol",
d'CARTESIAN, 2021

Publication

11	mafiadoc.com Internet Source	1 %
12	jurnal.kopusindo.com Internet Source	1 %
13	kar.kent.ac.uk Internet Source	1 %
14	ejournal.unma.ac.id Internet Source	1 %
15	Submitted to Hoa Sen University Student Paper	1 %
16	Submitted to Institut Bisnis dan Teknologi Indonesia (INSTIKI) Student Paper	1 %
17	Zaifullah Zaifullah, Tony Yulianto. "Analisis Cluster Untuk Pengelompokkan Prestasi Mahasiswa Angkatan 2013 Fakultas MIPA Universitas Islam Madura", Zeta - Math Journal, 2022 Publication	1 %
18	Diana Nurfitriana, Apriade Voutama. "Penerapan K-Means dan Rank Order Centroid pada Proporsi Individu dengan Keterampilan Teknologi Informasi dan Komputer", Jurnal Teknologi Terpadu, 2023 Publication	<1 %
19	Submitted to Tarumanagara University Student Paper	<1 %
20	jurnal.unissula.ac.id Internet Source	<1 %

21	mdpi-res.com Internet Source	<1 %
22	Nugroho Irawan Febianto, Nicodias Palasara. "Analisa Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018", Jurnal Sisfokom (Sistem Informasi dan Komputer), 2019 Publication	<1 %
23	knastik.ukdw.ac.id Internet Source	<1 %
24	thesesups.ups-tlse.fr Internet Source	<1 %
25	www.neliti.com Internet Source	<1 %
26	Rifaldi Ali Hakim, Sri Harjanto, Ari Wibowo. "PENERAPAPAN METODE K-MEANS CLUSTERING DALAM PENGELOLAAN KESEJAHTERAAN SOSIAL DI DESA PENTUR", Jurnal Teknologi Informasi dan Komunikasi (TIKomSiN), 2024 Publication	<1 %
27	libeldoc.bsuir.by Internet Source	<1 %
28	media.neliti.com Internet Source	<1 %
29	jurnal.pancabudi.ac.id Internet Source	<1 %
30	repository.ub.ac.id Internet Source	<1 %
31	text-id.123dok.com Internet Source	<1 %
32	Aslam Fatkhudin, Ahmad Khambali, Fenilinas Adi Artanto, Mundriyah Mundriyah, Nabil	<1 %

Ahmad Putra Zade. "Implementasi Algoritma Clustering K-Means Dalam Pengelompokan Mahasiswa", Jurnal Minfo Polgan, 2023

Publication

33

Farid Raihan, Sutisna. "Analisis Pengaruh Media Sosial X Terhadap Sikap dan Perilaku Remaja Menggunakan K-Means Clustering", Jurnal Indonesia : Manajemen Informatika dan Komunikasi, 2024

Publication

<1 %

34

Indra Syahputra, Ilhamsyah Ilhamsyah, Syahru Rahmayuda, Ferdy Febrianto. "SISTEM KLASTERISASI DATA KESEHATAN PENDUDUK UNTUK MENENTUKAN RENTANG DERAJAT KESEHATAN DAERAH MENGGUNAKAN K-MEANS", Jurnal Khatulistiwa Informatika, 2022

Publication

<1 %

35

Sekar Ayu pratiwi, Subhan Hafiz Nanda Ginting. "Penerapan Algoritma k-means dalam Data Mining untuk Mengidentifikasi Strategi Promosi di Politeknik Ganesha Medan", Jurnal Minfo Polgan, 2024

Publication

<1 %

36

e-jurnal.stikesmitraadiguna.ac.id

Internet Source

<1 %

37

ejournal.iainpalopo.ac.id

Internet Source

<1 %

38

etheses.uin-malang.ac.id

Internet Source

<1 %

39

idoc.pub

Internet Source

<1 %

40

journal.ibrahimiy.ac.id

Internet Source

<1 %

41

jurnaltarbiyah.uinsu.ac.id

Internet Source

<1 %

42 kzyjc.alljournals.cn <1 %
Internet Source

43 Ahmad Fairus Zabidi. "Penerapan Algoritma K-Means untuk Pengelompokan Koleksi Perpustakaan dengan Data Mining", Media Jurnal Informatika, 2024 <1 %
Publication

44 Ignatius Bagas Ristanto, Yustina Retno Wahyu Utami, Teguh Susyanto. "Penerapan Metode Clustering dengan Fuzzy C-Means untuk Memetakan Daerah Rawan Kecelakaan Lalu Lintas di Surakarta", Jurnal Ilmiah SINUS, 2021 <1 %
Publication

45 Rapel Sigit. "Penerapan Algoritma K-Means Clustering dalam Menganalisis Pola Peminjaman Buku di Perpustakaan", The Indonesian Journal of Computer Science, 2024 <1 %
Publication

46 ejournal.unsrat.ac.id <1 %
Internet Source

47 kabar-terhangat.blogspot.com <1 %
Internet Source

48 repository.unej.ac.id <1 %
Internet Source

49 Resistania Anggita Putri, Nida Inayah Maghfirani, Galih Rendi Setyawan, Adam Achmad Rayhan, Nur Aini Rakhmawati. "Analisis Pengelompokan Peraturan Kementerian dengan Menggunakan K-Means Clustering", Jurnal Sisfokom (Sistem Informasi dan Komputer), 2020 <1 %
Publication

Exclude quotes Off
Exclude bibliography On

Exclude matches Off