

Segmentasi Responden Survei Persepsi Kualitas Pelayanan Publik Menggunakan Algoritma K-Means dan DBSCAN

Agus Suprianto

Program Studi : Sistem Informasi, Universitas Darwan Ali

Email : supriantoagus338@gmail.com

ABSTRACT— Public service quality is commonly used to evaluate the effectiveness of government performance and bureaucratic reform. The Investment and One-Stop Integrated Service Office (DPMPSTSP) of Seruyan Regency conducts the Public Service Quality Perception Survey (SPKP) to assess community satisfaction. Although the overall index indicates a very good performance, aggregated scores do not fully capture differences in respondents' perceptions. Consequently, a segmentation-based analysis is required to reveal hidden patterns within the survey data. This research utilizes clustering techniques to group respondents according to similarity of service assessments. Two algorithms, K-Means and Density-Based Spatial Clustering of Applications with Noise (DBSCAN), were implemented and compared using internal validation measures, including Silhouette, Davies-Bouldin, and Calinski-Harabasz scores. The experiments show that K-Means generated four primary clusters describing general respondent groups, whereas DBSCAN detected density-based groups and several outliers. The segmentation results provide more detailed insights than average-based evaluation and can support more focused and evidence-based strategies for improving public service delivery.

Keywords— *public service quality, clustering, K-Means, DBSCAN, data mining, respondent segmentation.*

ABSTRAK— Kualitas pelayanan publik sering dijadikan tolok ukur kinerja instansi pemerintah serta keberhasilan reformasi birokrasi. DPMPSTSP Kabupaten Seruyan melaksanakan Survei Persepsi Kualitas Pelayanan (SPKP) untuk menilai tingkat kepuasan masyarakat. Meskipun indeks menunjukkan kategori sangat baik, nilai agregat belum sepenuhnya menggambarkan keragaman persepsi responden. Oleh sebab itu, diperlukan pendekatan segmentasi untuk mengidentifikasi pola penilaian yang lebih rinci. Penelitian ini memanfaatkan teknik clustering untuk mengelompokkan responden berdasarkan kemiripan penilaian pelayanan menggunakan algoritma K-Means dan DBSCAN. Evaluasi kualitas kluster dilakukan melalui metrik Silhouette, Davies-Bouldin, dan Calinski-Harabasz. Hasil pengujian menunjukkan bahwa K-Means membentuk empat kluster utama yang mewakili kelompok responden secara umum, sedangkan DBSCAN mampu mengenali kelompok berbasis kepadatan serta mendeteksi data pencilan. Temuan segmentasi memberikan informasi yang lebih spesifik dibandingkan analisis rata-rata sehingga dapat mendukung penyusunan strategi peningkatan pelayanan publik yang lebih tepat sasaran.

Kata kunci— *kualitas pelayanan publik, clustering, K-Means, DBSCAN, data mining, segmentasi responden.*

I. PENDAHULUAN

Pelayanan publik merupakan indikator strategis dalam menilai keberhasilan reformasi birokrasi, khususnya dalam mewujudkan tata kelola pemerintahan yang efektif, transparan, dan berorientasi pada kepuasan masyarakat. Seiring dengan perkembangan digitalisasi layanan pemerintahan, masyarakat tidak hanya menuntut kemudahan akses, tetapi juga mengharapkan kualitas pelayanan yang konsisten, responsif, dan adil. Ketidakmampuan instansi publik dalam memenuhi ekspektasi tersebut berpotensi menurunkan tingkat kepercayaan masyarakat terhadap pemerintah sebagai penyelenggara layanan. Saputra dkk.[1] menegaskan bahwa masih terdapat kesenjangan antara harapan masyarakat dan kualitas pelayanan yang diterima, terutama pada sektor pelayanan administrasi dan layanan publik berbasis proses.

Sebagai instrumen evaluasi kinerja pelayanan publik, pemerintah secara berkala melaksanakan Survei Persepsi Kualitas Pelayanan (SPKP). Survei ini dirancang

untuk mengukur persepsi masyarakat terhadap berbagai dimensi pelayanan publik, meliputi kejelasan informasi, kecepatan pelayanan, kompetensi petugas, hingga mekanisme penanganan pengaduan. Data SPKP berfungsi sebagai dasar empiris bagi instansi pemerintah dalam merumuskan kebijakan peningkatan pelayanan, sekaligus sebagai alat akuntabilitas publik terhadap kinerja penyelenggara layanan.

Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu (DPMPSTSP) Kabupaten Seruyan secara konsisten melaksanakan SPKP sejak tahun 2023 sebagai bagian dari upaya peningkatan mutu pelayanan publik. Hasil survei menunjukkan bahwa indeks SPKP berada pada kategori sangat baik, dengan nilai 90,22 pada tahun 2023, 89,55 pada tahun 2024, dan meningkat menjadi 93,45 pada periode Triwulan I–III tahun 2025[2],[3],[4]. Meskipun demikian, capaian nilai agregat tersebut belum sepenuhnya mencerminkan keseragaman persepsi masyarakat, karena beberapa indikator pelayanan, seperti kemudahan konsultasi, kejelasan informasi, serta

penanganan pengaduan, masih memperoleh skor relatif lebih rendah dibandingkan indikator lainnya.

Kondisi tersebut mengindikasikan adanya variasi persepsi masyarakat terhadap kualitas pelayanan publik yang tidak dapat dijelaskan hanya melalui nilai rata-rata survei. Hidayat dkk.[5] menjelaskan bahwa faktor latar belakang pengguna layanan, seperti tingkat pendidikan, jenis pekerjaan, dan pengalaman dalam mengakses layanan, berpengaruh signifikan terhadap persepsi kualitas pelayanan. Dengan kata lain, nilai indeks yang tinggi secara agregat dapat menutupi keberadaan kelompok masyarakat tertentu yang masih menghadapi hambatan atau ketidakpuasan dalam proses pelayanan.

Lebih lanjut, Taufiqurokhman dkk.[6] menyoroti bahwa dalam konteks pelayanan publik berbasis digital, tingkat kepercayaan dan kepuasan masyarakat tidak selalu sejalan dengan skor survei yang diperoleh. Hal ini terutama terjadi pada kelompok pengguna yang memiliki keterbatasan dalam adaptasi teknologi. Temuan tersebut menegaskan bahwa karakteristik demografis dan psikografis masyarakat berperan penting dalam membentuk persepsi terhadap kualitas pelayanan publik, sehingga analisis yang hanya bertumpu pada nilai agregat berpotensi menghasilkan rekomendasi kebijakan yang kurang tepat sasaran.

Namun demikian, pengolahan data SPKP di tingkat instansi pemerintah pada umumnya masih terbatas pada analisis deskriptif berupa perhitungan nilai rata-rata setiap indikator pelayanan. Pendekatan ini belum mampu mengidentifikasi pola persepsi yang beragam di antara kelompok responden. Utami dkk.[7] menekankan bahwa meskipun kualitas pelayanan dinilai tinggi secara keseluruhan, perbedaan persepsi antar kelompok masyarakat tetap dapat terjadi, sehingga diperlukan pendekatan analisis yang lebih mendalam dan bersifat segmentatif.

Untuk menjawab keterbatasan tersebut, diperlukan metode analisis berbasis data yang mampu mengelompokkan responden berdasarkan kesamaan pola persepsi terhadap kualitas pelayanan publik. Teknik data mining, khususnya metode clustering, menawarkan pendekatan yang sistematis dalam mengidentifikasi segmentasi responden. Algoritma K-Means banyak digunakan karena kemampuannya membentuk kelompok data berdasarkan kedekatan terhadap pusat cluster, sedangkan algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) unggul dalam mendeteksi cluster dengan kepadatan tidak merata serta mengidentifikasi data terluar (outlier). Ramadhan dkk.[8] menunjukkan bahwa perbandingan K-Means dan DBSCAN mampu mengungkap variasi karakteristik pengguna yang tidak terdeteksi apabila hanya menggunakan satu algoritma saja.

Secara metodologis, pendekatan clustering dalam analisis pelayanan publik semakin banyak digunakan untuk mengidentifikasi segmentasi masyarakat secara lebih presisi. Studi terbaru menunjukkan bahwa integrasi metode partitional dan density-based clustering mampu meningkatkan akurasi segmentasi data survei publik[9]. Selain itu, penerapan teknik segmentasi berbasis data

mining dinilai lebih adaptif terhadap heterogenitas karakteristik responden dibandingkan pendekatan agregatif konvensional [10].

Pendekatan ini efektif dalam membentuk struktur segmentasi yang relatif stabil dan mudah diinterpretasikan, terutama ketika distribusi data bersifat homogen atau mendekati simetris. Sebaliknya, DBSCAN mengidentifikasi kelompok berdasarkan kepadatan lokal dan tidak mengasumsikan bentuk kluster tertentu, sehingga mampu membentuk kluster dengan bentuk tidak beraturan serta mengklasifikasikan data dengan kepadatan rendah sebagai noise. Perbedaan karakteristik tersebut telah dijelaskan dalam studi perbandingan algoritma clustering yang menunjukkan bahwa K-Means cenderung unggul dalam efisiensi komputasi dan kestabilan struktur global, sedangkan DBSCAN lebih adaptif terhadap pola distribusi data yang kompleks dan tidak seragam [11].

Perbedaan paradigma tersebut mengimplikasikan bahwa hasil segmentasi yang dihasilkan oleh kedua algoritma dapat memiliki karakteristik struktural yang berbeda, baik dari sisi jumlah kluster, ukuran kluster, maupun distribusi anggota kluster. Oleh karena itu, evaluasi kualitas kluster menggunakan metrik internal menjadi langkah krusial untuk memastikan tingkat kekompakan (cohesion) dan pemisahan (separation) antar kelompok yang terbentuk secara objektif [12]. Evaluasi ini tidak hanya berfungsi untuk membandingkan performa algoritma, tetapi juga untuk memastikan bahwa struktur segmentasi yang dihasilkan benar-benar merepresentasikan pola persepsi masyarakat yang mendasarinya.

Berdasarkan uraian tersebut, dapat disimpulkan bahwa meskipun indeks SPKP DPMPTSP Kabupaten Seruyan menunjukkan kategori sangat baik, masih terdapat kebutuhan mendesak untuk melakukan analisis yang lebih mendalam terhadap variasi persepsi masyarakat. Tanpa segmentasi responden, strategi peningkatan pelayanan publik berisiko bersifat umum dan kurang responsif terhadap kebutuhan kelompok tertentu. Oleh karena itu, penelitian ini memfokuskan pada segmentasi responden SPKP DPMPTSP Kabupaten Seruyan periode 2023–2025 (Triwulan I–III) menggunakan algoritma K-Means dan DBSCAN, guna menghasilkan rekomendasi peningkatan pelayanan publik yang lebih terarah, inklusif, dan berbasis bukti empiris.

II. METODOLOGI PENELITIAN

Penelitian ini menerapkan pendekatan kuantitatif dengan memanfaatkan teknik *unsupervised learning* melalui metode *clustering* untuk mengelompokkan responden berdasarkan kemiripan persepsi terhadap kualitas pelayanan publik. Pendekatan tersebut dipilih karena memungkinkan identifikasi pola data tanpa memerlukan label kelas, sehingga sesuai untuk karakteristik data survei yang heterogen.

A. Tahapan Penelitian

Pelaksanaan penelitian disusun secara bertahap dan terstruktur, dimulai dari identifikasi permasalahan hingga penyusunan kesimpulan. Proses diawali dengan analisis hasil Survei Persepsi Kualitas Pelayanan (SPKP) guna

mengamati variasi penilaian responden terhadap indikator pelayanan. Tahapan berikutnya mencakup pengumpulan data, kajian literatur, pra-pemrosesan, penerapan algoritma K-Means dan DBSCAN, evaluasi performa klaster, interpretasi hasil segmentasi, serta perumusan rekomendasi perbaikan layanan. Seluruh rangkaian proses digambarkan dalam bentuk *flowchart* untuk memperjelas alur penelitian.

B. Lokasi Penelitian

Kegiatan penelitian dilaksanakan pada Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu (DPMPTSP) Kabupaten Seruyan, Provinsi Kalimantan Tengah. Instansi ini dipilih karena secara rutin menyelenggarakan SPKP sebagai bagian dari evaluasi mutu pelayanan publik, khususnya pada layanan perizinan dan nonperizinan. Sumber data diperoleh dari laporan resmi SPKP yang dikelola oleh bidang pengelolaan data dan sistem informasi penanaman modal.

C. Populasi dan Sampel

Populasi penelitian mencakup seluruh responden yang berpartisipasi dalam SPKP periode 2023–2025. Responden merupakan masyarakat yang secara langsung memanfaatkan layanan DPMPTSP sehingga memiliki pengalaman nyata terhadap proses pelayanan.

Berdasarkan laporan survei, jumlah responden terdiri atas 703 orang pada tahun 2023, 402 orang pada tahun 2024, dan 256 orang pada periode 2025, sehingga total keseluruhan data yang dianalisis sebanyak 1.363 responden.

Teknik sampling yang digunakan adalah *total sampling*, di mana seluruh populasi dimanfaatkan sebagai sampel penelitian. Pendekatan ini dipilih karena setiap responden relevan dengan tujuan analisis sehingga tidak diperlukan seleksi tambahan.

D. Metode Pengumpulan Data

Data penelitian berupa data sekunder yang diperoleh melalui dokumentasi hasil SPKP DPMPTSP Kabupaten Seruyan. Instrumen survei mengukur persepsi masyarakat terhadap delapan aspek kualitas pelayanan publik, meliputi kejelasan informasi, ketepatan waktu layanan, kompetensi petugas, sikap petugas, keadilan pelayanan, kenyamanan fasilitas, transparansi biaya, serta penanganan pengaduan. Penilaian dilakukan menggunakan skala Likert dengan rentang nilai 1 sampai 4.

E. Variabel Penelitian

Variabel yang dianalisis berasal dari delapan indikator SPKP yang merepresentasikan dimensi utama pelayanan publik. Seluruh indikator berbentuk numerik dan digunakan sebagai atribut dalam proses *clustering*. Kombinasi nilai dari setiap indikator mencerminkan pola persepsi responden, yang selanjutnya menjadi dasar pembentukan segmentasi kelompok.

F. Teknik Analisis Data

Pengolahan data dilakukan menggunakan bahasa pemrograman Python melalui platform Google Colab dengan dukungan pustaka *pandas*, *numpy*, *matplotlib*, dan *scikit-learn*. Tahapan analisis meliputi:

1. Eksplorasi Data

Peninjauan awal dilakukan untuk memahami struktur dataset, distribusi nilai, serta memastikan kelengkapan data sebelum proses analisis.

2. Pra-Pemrosesan Data

Data dibersihkan dari duplikasi dan nilai kosong, kemudian dilakukan penskalaan menggunakan metode Min–Max agar seluruh atribut berada pada rentang yang seragam sehingga perhitungan jarak tidak dipengaruhi oleh perbedaan skala.

3. Clustering dengan K-Means

Algoritma K-Means membentuk kelompok dengan meminimalkan jarak setiap data terhadap titik pusat klaster. Jumlah klaster ditentukan melalui evaluasi metode Elbow dan Silhouette Score untuk memperoleh konfigurasi yang paling optimal.

4. Clustering dengan DBSCAN

Pengelompokan berbasis kepadatan diterapkan menggunakan parameter epsilon (ϵ) dan minimum points (MinPts). Pendekatan ini juga memungkinkan identifikasi data yang tidak tergabung dalam klaster sebagai noise atau outlier.

5. Evaluasi Hasil Clustering

Kinerja klaster dinilai menggunakan tiga metrik internal, yaitu Silhouette Coefficient, Davies–Bouldin Index, dan Calinski–Harabasz Index untuk mengukur kekompakan dan pemisahan antar kelompok.

6. Interpretasi Hasil

Setiap klaster dianalisis berdasarkan nilai rata-rata indikator pelayanan untuk mengidentifikasi karakteristik kelompok responden, yang kemudian digunakan sebagai dasar rekomendasi peningkatan kualitas layanan.

Melalui tahapan tersebut, penelitian ini membandingkan performa algoritma K-Means dan DBSCAN dalam menghasilkan segmentasi responden yang paling representatif terhadap persepsi kualitas pelayanan publik.

III. DESAIN, HASIL DAN PEMBAHASAN

Tahap analisis dilakukan setelah proses eksplorasi dan pra-pemrosesan data Survei Persepsi Kualitas Pelayanan (SPKP) periode 2023–2025 dengan total 1.363 responden dan delapan indikator penilaian kualitas pelayanan publik. Seluruh indikator dinormalisasi menggunakan metode Min–Max agar setiap variabel memiliki bobot yang setara dalam perhitungan jarak. Selanjutnya dilakukan segmentasi menggunakan algoritma K-Means dan DBSCAN untuk mengidentifikasi pola persepsi masyarakat terhadap kualitas pelayanan di DPMPTSP Kabupaten Seruyan. Alur lengkap tahapan pengolahan data dan desain analisis penelitian disajikan pada Gambar 1.



Gambar 1. Alur desain analisis dan tahapan pengolahan data penelitian

A. Deskripsi Data dan Pra-Pemrosesan

Pemeriksaan awal menunjukkan bahwa seluruh variabel bersifat numerik dan tidak ditemukan *missing values*. Setiap variabel diskalakan ke interval 0 hingga 1 dengan teknik Min–Max sehingga perhitungan jarak antar data memiliki bobot yang seimbang. Tahap ini dilakukan untuk memastikan setiap indikator memiliki kontribusi yang seimbang dalam proses clustering. Seluruh proses disajikan pada gambar 2,3 dan 4 dibawah ini :

```
df_spkp.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1363 entries, 0 to 1362
Data columns (total 8 columns):
#   Column  Non-Null Count  Dtype
---  ---      -
0    U1      1363 non-null    int64
1    U2      1363 non-null    int64
2    U3      1363 non-null    int64
3    U4      1363 non-null    int64
4    U5      1363 non-null    int64
5    U6      1363 non-null    int64
6    U7      1363 non-null    int64
7    U8      1363 non-null    int64
dtypes: int64(8)
memory usage: 85.3 KB
```

Gambar 2. Pengecekan Struktur Data

```
# Pengecekan missing value
df_spkp.isnull().sum()
U1 0
U2 0
U3 0
U4 0
U5 0
U6 0
U7 0
U8 0
dtype: int64
```

Gambar 3. Pengecekan Missing Values

```
# Inisialisasi Min-Max Scaler
scaler = MinMaxScaler()

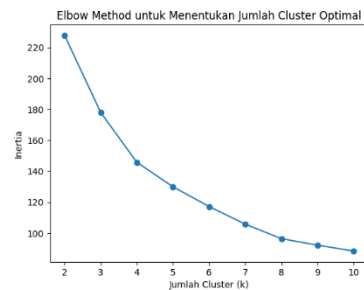
# Normalisasi data
df_norm = pd.DataFrame(
    scaler.fit_transform(df_spkp),
    columns=df_spkp.columns
)
df_norm.head()
U1 U2 U3 U4 U5 U6 U7 U8
0 0.666667 0.666667 0.666667 0.666667 1.0 1.0 0.0 1.0
1 0.666667 0.666667 0.666667 1.000000 1.0 1.0 0.0 0.5
2 1.000000 0.666667 0.666667 0.666667 1.0 1.0 1.0 0.5
3 1.000000 0.666667 1.000000 1.000000 1.0 1.0 1.0 0.5
4 1.000000 0.666667 0.666667 1.000000 1.0 1.0 0.0 0.5
```

Gambar 4. Hasil Normalisasi Data

B. Hasil Clustering Menggunakan K-Means

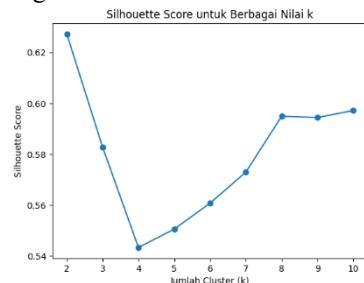
Algoritma K-Means diterapkan untuk mengelompokkan responden berdasarkan kedekatan jarak terhadap pusat kluster. Tahap awal yang dilakukan adalah menentukan jumlah kluster optimal. Penentuan ini menggunakan metode *Elbow* dengan mengamati penurunan nilai *Within Cluster Sum of Squares* (WCSS) pada setiap penambahan jumlah kluster.

Berdasarkan kurva *Elbow*, terlihat adanya penurunan nilai WCSS yang cukup tajam hingga titik tertentu, kemudian penurunannya mulai melandai. Titik tekuk tersebut muncul pada $k = 4$, sehingga jumlah empat kluster dipilih sebagai konfigurasi yang paling optimal.



Gambar 5. Kurva Elbow Method (WCSS)

Untuk memperkuat hasil tersebut, dilakukan perhitungan *Silhouette Score* pada beberapa jumlah kluster. Nilai *Silhouette* menunjukkan tingkat kekompakan dan pemisahan antar kluster. Hasil perhitungan menunjukkan bahwa nilai terbaik diperoleh pada jumlah kluster yang sama, sehingga konfigurasi $k = 4$ dinilai paling sesuai.



Gambar 6. Grafik Silhouette Score terhadap Jumlah Kluster

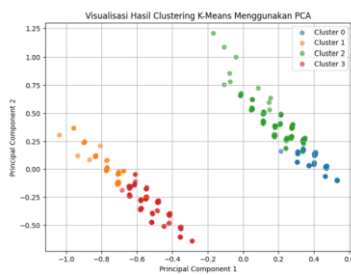
Empat kluster yang terbentuk menunjukkan adanya perbedaan pola penilaian responden terhadap indikator pelayanan publik. Setiap kluster memiliki rata-rata skor yang berbeda pada masing-masing indikator, sehingga mencerminkan variasi tingkat kepuasan masyarakat. Terdapat kelompok responden dengan nilai indikator

relatif tinggi pada hampir seluruh aspek pelayanan, kelompok dengan nilai sedang, serta kelompok dengan beberapa indikator yang cenderung lebih rendah. Perbedaan karakteristik ini menunjukkan bahwa meskipun nilai indeks kepuasan secara umum berada pada kategori baik, persepsi masyarakat sebenarnya tidak homogen dan terbagi ke dalam beberapa segmen. Distribusi jumlah responden pada masing-masing kluster disajikan pada tabel berikut.

No. Cluster	Jumlah Responden	Persentase %
0	652	47.84
1	322	23.62
2	193	14.16
3	196	14.38

Tabel 1. Distribusi Responden Hasil Evaluasi Clustering K-Means

Visualisasi menggunakan Principal Component Analysis (PCA) dua dimensi memperlihatkan empat kelompok data yang terbentuk secara jelas.



Gambar 7. Visualisasi PCA 2D Hasil K-Means (k = 4)

Empat kluster tersebut menggambarkan segmentasi responden berdasarkan tingkat kepuasan pelayanan, mulai dari kelompok dengan persepsi sangat baik hingga kelompok dengan persepsi lebih rendah pada beberapa indikator.

Selanjutnya evaluasi internal menghasilkan nilai Silhouette Coefficient, Davies–Bouldin Index, dan Calinski–Harabasz Index yang menunjukkan bahwa kluster yang terbentuk cukup kompak dan terpisah dengan baik.

Metrik Evaluasi	Nilai
Silhouette Coefficient	0.543
Davies–Bouldin Index	1.173
Calinski–Harabasz Index	1546.86

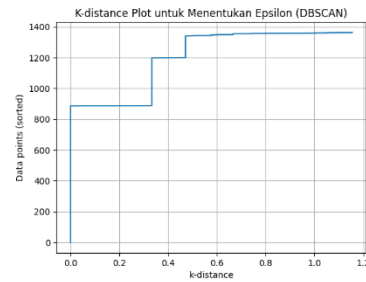
Tabel 2. Hasil Evaluasi Clustering K-Means

C. Hasil Clustering Menggunakan DBSCAN

Selain K-Means, segmentasi juga dilakukan menggunakan algoritma DBSCAN dengan cara membangun kluster dari area dengan kepadatan tinggi serta menandai data yang tidak memenuhi kriteria kepadatan sebagai noise. Pendekatan ini tidak memerlukan penentuan jumlah kluster di awal, melainkan menggunakan parameter epsilon (ϵ) dan *minimum points* (MinPts) sebagai dasar pembentukan kelompok.

Penentuan nilai ϵ dilakukan melalui analisis *k-distance plot* dengan mengurutkan jarak tetangga terdekat setiap titik data. Grafik menunjukkan adanya perubahan kemiringan yang menandakan batas radius kepadatan

optimal. Nilai tersebut kemudian digunakan sebagai parameter dalam proses clustering.



Gambar 8. Kurva k-Distance Plot Penentuan Nilai ϵ

Setelah parameter ditentukan, algoritma DBSCAN dijalankan pada data yang telah dinormalisasi. Hasilnya menunjukkan terbentuknya beberapa kluster utama serta sejumlah data yang dikategorikan sebagai *noise*. Keberadaan *noise* menunjukkan adanya responden dengan pola penilaian yang berbeda secara signifikan dibandingkan mayoritas responden.

Pembagian kluster DBSCAN menunjukkan adanya kelompok responden dengan karakteristik penilaian yang relatif serupa di dalam masing-masing kelompok. Selain itu, sejumlah responden tidak tergabung ke dalam kluster mana pun dan dikategorikan sebagai *noise*. Kelompok *noise* ini merepresentasikan responden dengan pola persepsi yang berbeda secara ekstrem atau tidak konsisten dibandingkan kelompok lain. Informasi ini tidak diperoleh pada metode K-Means karena seluruh data selalu dipaksa masuk ke dalam kluster tertentu. Distribusi jumlah anggota pada setiap kluster serta jumlah *noise* ditampilkan pada tabel berikut.

Cluster	Jumlah Responden	Persentase (%)	Keterangan
-1	476	34,92	Noise
0	16	1,17	Cluster
1	495	36,32	Cluster
2	24	1,76	Cluster
3	45	3,30	Cluster
4	17	1,25	Cluster
5	18	1,32	Cluster
6	209	15,33	Cluster
7	23	1,69	Cluster
8	17	1,25	Cluster
9	23	1,69	Cluster

Tabel 3. Distribusi Responden Hasil Evaluasi Clustering DBSCAN

Evaluasi ulang dilakukan dengan mengikutsertakan seluruh data, termasuk label noise (-1), agar hasil mencerminkan struktur segmentasi secara utuh. Hasil perhitungan disajikan pada tabel dibawah ini :

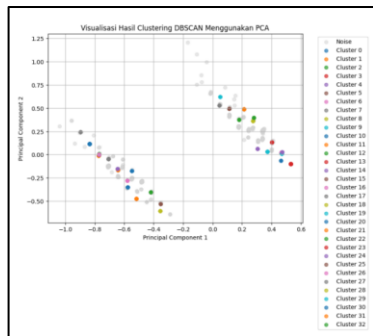
Metrik Evaluasi	Nilai
Silhouette Coefficient	0.467
Davies–Bouldin Index	1.041
Calinski–Harabasz Index	213.489

Tabel 4. Hasil Evaluasi Clustering DBSCAN

Nilai Silhouette yang berada pada kisaran 0,4–0,5 menunjukkan bahwa struktur kluster cukup terpisah namun tidak sekompak K-Means. Sementara itu, nilai

Davies–Bouldin yang lebih rendah dibandingkan K-Means mengindikasikan pemisahan antar kluster yang relatif baik. Namun, nilai Calinski–Harabasz yang lebih kecil menunjukkan bahwa variasi antar kluster secara global tidak sekuat struktur yang dibentuk oleh K-Means. Perbedaan ini mencerminkan karakteristik algoritma DBSCAN yang membentuk kluster berdasarkan kepadatan lokal serta mengidentifikasi sebagian data sebagai noise, sehingga menghasilkan segmentasi yang lebih selektif dibandingkan pendekatan centroid-based.

Visualisasi PCA dua dimensi memperlihatkan bahwa kluster terbentuk mengikuti pola kepadatan data. Beberapa kelompok tampak menyebar secara alami, sementara titik-titik *noise* berada terpisah dari kluster utama.



Gambar 9. Visualisasi PCA DBSCAN

D. Sintesis Hasil Clustering

Berdasarkan tahapan penerapan algoritma clustering yang telah dilakukan, penentuan jumlah kluster optimal merupakan bagian penting sebelum hasil segmentasi dianalisis lebih lanjut. Pada penelitian ini, penentuan kluster optimal dilakukan secara terpisah untuk algoritma K-Means dan DBSCAN sesuai dengan karakteristik masing-masing metode.

Pada algoritma K-Means, penentuan jumlah kluster optimal dilakukan menggunakan metode *Elbow* dan *Silhouette Score*. Hasil analisis *Elbow Method* menunjukkan adanya titik tekuk pada nilai $k = 4$, di mana penurunan nilai *Within Cluster Sum of Squares* (WCSS) mulai melandai. Kondisi ini mengindikasikan bahwa penambahan jumlah kluster setelah nilai tersebut tidak memberikan peningkatan signifikan terhadap kekompakan kluster. Hasil tersebut diperkuat oleh perhitungan *Silhouette Score* yang menunjukkan nilai terbaik pada jumlah kluster yang sama. Dengan demikian, jumlah kluster optimal untuk algoritma K-Means dalam penelitian ini ditetapkan sebanyak empat kluster.

Sementara itu, pada algoritma DBSCAN, penentuan kluster optimal tidak didasarkan pada jumlah kluster yang ditentukan di awal, melainkan pada pemilihan parameter epsilon (ϵ) dan *minimum points* (MinPts). Penentuan nilai ϵ dilakukan melalui analisis *k-distance plot* untuk mengidentifikasi perubahan kemiringan kurva yang menunjukkan batas kepadatan optimal. Berdasarkan hasil analisis tersebut, nilai ϵ dan MinPts ditetapkan sesuai konfigurasi yang menghasilkan pembentukan kluster paling stabil pada data SPKP yang digunakan.

Hasil penerapan DBSCAN dengan parameter tersebut menghasilkan beberapa kluster utama serta

sejumlah data yang dikategorikan sebagai *noise*. Keberadaan *noise* menunjukkan adanya responden yang tidak memenuhi kriteria kepadatan minimum sehingga tidak dapat dimasukkan ke dalam kluster tertentu. Dengan demikian, jumlah kluster pada DBSCAN bersifat adaptif dan ditentukan secara otomatis berdasarkan struktur kepadatan data, bukan melalui penetapan nilai kluster di awal seperti pada K-Means.

Berdasarkan hasil tersebut, dapat disintesis bahwa penentuan kluster optimal pada penelitian ini dilakukan melalui pendekatan yang berbeda sesuai karakteristik algoritma. K-Means menghasilkan jumlah kluster optimal sebanyak empat kluster berdasarkan evaluasi *Elbow* dan *Silhouette Score*, sedangkan DBSCAN menghasilkan kluster berdasarkan parameter kepadatan yang ditentukan melalui *k-distance plot*. Sintesis ini menjadi dasar untuk tahap selanjutnya, yaitu perbandingan hasil segmentasi dan evaluasi kinerja kedua algoritma pada subbab berikutnya.

E. Perbandingan Hasil Clustering K-Means dan DBSCAN

Sebagai ringkasan dan penegasan terhadap hasil evaluasi clustering yang telah dipaparkan pada subbagian sebelumnya, dilakukan perbandingan kinerja algoritma K-Means dan DBSCAN menggunakan tiga metrik evaluasi internal, yaitu *Silhouette Coefficient* (SC), *Davies–Bouldin Index* (DBI), dan *Calinski–Harabasz Index* (CHI). Ketiga metrik tersebut digunakan untuk menilai tingkat kekompakan kluster, pemisahan antar kluster, serta kualitas struktur pengelompokan secara keseluruhan.

Metode	SC	DBI	CHI
K-Means	0.543	1.173	1547
DBSCAN	0.467	1.041	213.489

Tabel 5. Tabel Perbandingan Hasil Evaluasi

Berdasarkan hasil perbandingan pada Tabel 5, terlihat bahwa algoritma K-Means dan DBSCAN menghasilkan karakteristik evaluasi clustering yang berbeda sesuai dengan paradigma pembentukan kluster masing-masing. K-Means memperoleh nilai *Silhouette Coefficient* yang lebih tinggi (0,543), yang menunjukkan tingkat kekompakan dan konsistensi struktur kluster yang relatif lebih baik secara global. Selain itu, seluruh responden berhasil dikelompokkan ke dalam empat kluster utama, sehingga segmentasi yang dihasilkan bersifat menyeluruh dan representatif terhadap populasi penelitian.

Di sisi lain, DBSCAN menghasilkan nilai *Silhouette* sebesar 0,467 yang masih berada pada kategori cukup baik, serta nilai *Davies–Bouldin Index* yang sedikit lebih rendah (1,041), mengindikasikan pemisahan antar kluster yang relatif memadai. Namun demikian, nilai *Calinski–Harabasz Index* yang lebih kecil menunjukkan bahwa variasi antar kluster secara keseluruhan tidak sekuat struktur yang dibentuk oleh K-Means. Hal ini wajar mengingat DBSCAN membentuk kluster berdasarkan kepadatan lokal serta mengidentifikasi sebagian data sebagai *noise*, sehingga struktur segmentasi menjadi lebih selektif dan tidak memaksakan seluruh data masuk ke dalam kluster tertentu.

Perbedaan hasil evaluasi tersebut mencerminkan perbedaan pendekatan antara metode partitional dan density-based clustering. K-Means menekankan homogenitas global dan keterwakilan populasi secara menyeluruh, sedangkan DBSCAN lebih sensitif terhadap variasi lokal dan pola persepsi yang ekstrem. Dalam konteks segmentasi persepsi kualitas pelayanan publik yang bertujuan mendukung perumusan kebijakan berbasis populasi, metode yang mampu merepresentasikan keseluruhan responden secara stabil menjadi lebih relevan.

Dengan demikian, K-Means dinilai lebih sesuai sebagai metode segmentasi utama karena menghasilkan pembagian kluster yang stabil, konsisten, dan mudah diinterpretasikan pada tingkat populasi. Sementara itu, DBSCAN berperan sebagai metode komplementer yang efektif untuk mengidentifikasi kelompok responden dengan pola penilaian yang berbeda secara signifikan atau tidak homogen.

IV. KESIMPULAN

Penelitian ini berhasil melakukan segmentasi responden Survei Persepsi Kualitas Pelayanan (SPKP) DPMPTSP Kabupaten Seruyan periode 2023–2025 menggunakan algoritma K-Means dan DBSCAN untuk mengidentifikasi pola persepsi masyarakat terhadap kualitas pelayanan publik. Hasil penerapan K-Means menunjukkan bahwa responden terbagi ke dalam empat kluster utama yang merepresentasikan perbedaan tingkat persepsi pelayanan, sehingga mengindikasikan bahwa karakteristik kepuasan masyarakat bersifat heterogen dan tidak dapat direpresentasikan hanya melalui nilai rata-rata indeks.

Berdasarkan evaluasi menggunakan Silhouette Coefficient, Davies–Bouldin Index, dan Calinski–Harabasz Index, kedua algoritma menunjukkan karakteristik kinerja yang berbeda. K-Means menghasilkan struktur kluster yang lebih kompak secara global serta mencakup seluruh responden sehingga lebih mudah diinterpretasikan sebagai segmentasi populasi secara menyeluruh. Sementara itu, DBSCAN menunjukkan pemisahan kluster yang cukup baik namun dengan variasi antar kluster yang lebih rendah secara global. Karakteristik ini dipengaruhi oleh pendekatan berbasis kepadatan yang memungkinkan identifikasi noise sebagai bagian dari struktur segmentasi. Oleh karena itu, K-Means dinilai lebih sesuai sebagai metode utama dalam segmentasi persepsi kualitas pelayanan publik, sedangkan DBSCAN berperan sebagai metode pendukung untuk mengidentifikasi pola persepsi yang lebih spesifik atau tidak homogen.

Hasil segmentasi yang diperoleh memberikan gambaran yang lebih rinci mengenai variasi persepsi masyarakat, sehingga dapat digunakan sebagai dasar dalam menentukan prioritas perbaikan pelayanan. Kelompok responden dengan tingkat persepsi lebih rendah menjadi fokus utama peningkatan kualitas layanan, sedangkan kelompok dengan persepsi tinggi dapat dipertahankan sebagai praktik pelayanan yang telah berjalan baik. Dengan demikian, pendekatan clustering

mampu mendukung perumusan kebijakan pelayanan publik yang lebih terarah dan berbasis data.

V. REFERENSI

- [1] A. Saputra and R. Yusuf, “Perbandingan Algoritma DBSCAN dan K-MEANS dalam Segmentasi Pelanggan Pengguna Transportasi Publik Transjakarta Menggunakan Metode RFM,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 4, pp. 1346–1361, 2024, doi: 10.57152/malcom.v4i4.1516.
- [2] DPMPTSP Seruyan, “Survei persepsi kualitas pelayanan,” 2023.
- [3] DPMPTSP Seruyan, “Survei persepsi kualitas pelayanan,” 2024.
- [4] DPMPTSP Seruyan, “Survei persepsi kualitas pelayanan,” 2025.
- [5] E. N. Hidayat, “MEASURING SERVICE QUALITY AT PUBLIC SERVICE MAL IN,” vol. V, no. 25, [Online]. Available: <https://doi.org/10.57266/epistemik.v5i1.275>
- [6] T. Taufiqurokhman, E. Satispi, A. Andriansyah, M. Murod, and E. Sulastri, “The impact of e-service quality on public trust and public satisfaction in e-government public services,” *Int. J. Data Netw. Sci.*, vol. 8, no. 2, pp. 765–772, 2024, doi: 10.5267/j.ijdns.2024.1.002.
- [7] F. R. Utami and L. Mursyidah, “Improving the Quality of Public Services from an Indonesian Village Perspective: Meningkatkan Kualitas Pelayanan Publik dilihat dari Wawasan Desa Indonesia,” *Indones. J. Public Policy Rev.*, vol. 25, no. 3 SE-Public Policy, p. 10.21070/ijppr.v25i3.1389, May 2024, doi: 10.21070/ijppr.v25i3.1389.
- [8] A. Ramadhan, F. Achmad, I. Zulkarnain, and M. Aritsugi, “Evaluation of K-Means, DBSCAN, and Hierarchical Clustering for Strategic Segmentation of Tourism SMEs in Rembang, Indonesia,” *J. Tek. Inform.*, vol. 6, no. 3, pp. 1605–1630, 2025, doi: 10.52436/1.jutif.2025.6.3.4602.
- [9] D. Chicco, A. Campagner, A. Spagnolo, D. Ciucci, and G. Jurman, “The Silhouette coefficient and the Davies-Bouldin index are more informative than Dunn index, Calinski-Harabasz index, Shannon entropy, and Gap statistic for unsupervised clustering internal evaluation of two convex clusters,” *PeerJ Comput. Sci.*, vol. 11, p. e3309, Nov. 2025, doi: 10.7717/peerj-cs.3309.
- [10] N. N. A. Bryar A. Hassan, Noor Bahjat Tayfor, Alla A. Hassan, Aram M. Ahmed, Tarik A. Rashid, “From A-to-Z Review of Clustering Validation Indices,” 2024, [Online]. Available: <https://doi.org/10.48550/arXiv.2407.20246>
- [11] N. P. Sutramiani, I. M. T. Arthana, and P. Fane, “The Performance Comparison of DBSCAN and K-Means Clustering for MSMEs Grouping based on Asset Value and Turnover,” vol. 10, no. 1, pp.

13–24, 2024.
[12] P. M. Hasugian, D. Mathelinea, S. Simamora, P. Barita, and N. Simangunsong, “Comparative Evaluation of Data Clustering Accuracy through

Integration of Dimensionality Reduction and Distance Metric,” vol. 24, no. 3, pp. 577–588, 2025, doi: 10.30812/matrik.v24i3.5057.