



Klasterisasi Kesiapan Digital Daerah: Studi Kasus Indeks SPBE di Jawa Barat, Indonesia

Agustia Hananto^{1*}, Tukino², Elfina Novalia²

¹²³Fakultas Ilmu Komputer, Universitas Buana Perjuangan Karawang, Indonesia

Email : agustia.hananto@ubpkarawang.ac.id¹, tukino@ubpkarawang.ac.id², elfina.novalia@ubpkarawang.ac.id³

Alamat: Jalan Ronggo Waluyo Sirnabaya, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361

Korespondensi penulis: agustia.hananto@ubpkarawang.ac.id¹

Abstract. *Public sector digital transformation requires a deep understanding of the digital readiness of each administrative region. The Electronic Government System (EGIS) Index is used by the Government of Indonesia as a measuring tool to assess the digital maturity of government agencies. This study aims to cluster districts/cities in West Java Province based on their 2023 EGIS scores to identify hidden patterns of digital readiness. Three unsupervised learning algorithms—K-Means, DBSCAN, and Agglomerative Clustering—are used to explore data-driven regional segmentation. The analyzed dataset includes 27 administrative regions and a number of numerical features related to the EGIS dimensions. The results show that each method is able to form clusters that reflect variations in digital readiness, with DBSCAN producing the most detailed segmentation and being able to detect outliers. Agglomerative Clustering shows good hierarchical separation, while K-Means provides a fairly representative general division. This study provides an analytical basis for contextual and targeted cluster-based policy making in developing regional digital transformation..*

Keywords: *Agglomerative Clustering, clustering, DBSCAN, K-Means, SPBE, digital transformation.*

Abstrak. Transformasi sektor digital menuntut pemahaman mendalam atas kesiapan digital setiap wilayah administratif. Indeks Sistem Pemerintahan Berbasis Elektronik (SPBE) digunakan oleh Pemerintah Indonesia sebagai alat ukur untuk menilai kematangan digital instansi pemerintah. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Jawa Barat berdasarkan skor SPBE tahun 2023 guna mengidentifikasi pola kesiapan digital yang tersembunyi. Tiga algoritma pembelajaran tanpa pengawasan—K-Means, DBSCAN, dan Agglomerative Clustering—digunakan untuk mengeksplorasi segmentasi wilayah secara berbasis data. Dataset yang mencakup 27 wilayah administratif dan sejumlah fitur numerik terkait dimensi SPBE. Hasil menunjukkan bahwa masing-masing metode mampu membentuk kluster yang mencerminkan variasi kesiapan digital, dengan DBSCAN menghasilkan segmentasi paling rinci dan mampu mendeteksi outlier. Agglomerative Clustering menampilkan hierarkis yang baik, sedangkan K-Means memberikan pembagian umum yang cukup representatif. Studi ini memberikan landasan analitis sebagai pengambilan kebijakan berbasis kluster yang kontekstual dan tepat sasaran dalam pengembangan transformasi digital daerah.

Kata kunci: *Agglomerative Clustering, clustering, DBSCAN, K-Means, SPBE, transformasi digital.*

1. LATAR BELAKANG

Transformasi digital dalam sektor publik telah menjadi agenda utama di berbagai negara, termasuk Indonesia, untuk meningkatkan efisiensi layanan, transparansi pemerintahan, dan partisipasi publik (Al-Sharafi et al., 2020; Zuiderwijk et al., 2021). Digitalisasi ini tidak hanya mengubah cara layanan publik disediakan, tetapi juga mendorong reformasi struktural di berbagai tingkat pemerintahan. Dalam konteks ini, evaluasi kesiapan pemerintah digital menjadi sangat penting untuk memastikan transformasi digital berjalan efektif dan inklusif.

Sebagai upaya untuk mengukur kemajuan digitalisasi sektor publik, Pemerintah Indonesia mengembangkan Indeks Sistem Pemerintahan Berbasis Elektronik (SPBE). Indeks ini berfungsi sebagai alat ukur standar untuk menilai sejauh mana instansi pemerintah

menerapkan prinsip-prinsip pemerintahan berbasis elektronik dalam menjalankan tugas dan fungsi (Kementerian PANRB, 2021). SPBE tidak hanya menjadi indikator kinerja internal, tetapi juga menjadi acuan penting dalam kebijakan penguatan kapasitas digital nasional.

Namun demikian, penerapan SPBE di tingkat daerah menunjukkan adanya ketimpangan yang cukup signifikan antar wilayah. Beberapa kabupaten/kota menunjukkan kinerja digital yang relatif tinggi, sementara yang lain masih tertinggal (Nugroho & Subekti, 2022). Ketimpangan ini mengindikasikan perlunya analisis yang lebih mendalam untuk memahami pola kesiapan digital di tingkat lokal, guna merumuskan kebijakan intervensi yang lebih tepat sasaran.

Sejauh ini, penelitian mengenai SPBE lebih banyak fokus pada evaluasi agregat nasional atau studi kasus individu, sementara penelitian yang melakukan segmentasi wilayah berdasarkan pola kesiapan digital masih sangat terbatas. Padahal, pendekatan segmentasi dapat memberikan wawasan strategi tentang karakteristik kelompok daerah yang memiliki tingkat kesiapan serupa, sehingga intervensi dapat disesuaikan dengan kebutuhan spesifik setiap kelompok.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Jawa Barat berdasarkan kesamaan skor indeks SPBE mereka. Selain itu, penelitian ini juga berusaha mengidentifikasi karakteristik utama dari setiap klaster yang terbentuk, sehingga dapat memberikan pemahaman yang lebih komprehensif tentang variasi kesiapan digital antar daerah.

Kontribusi utama dari penelitian ini adalah penerapan metode pembelajaran tanpa pengawasan, yaitu K-Means, DBSCAN, dan Agglomerative Clustering, untuk mengungkap pola tersembunyi dalam data SPBE tingkat daerah. Dengan pendekatan ini, penelitian ini diharapkan dapat memberikan rekomendasi berbasis klaster bagi pengambil kebijakan untuk mendukung perencanaan pembangunan digital yang lebih inklusif dan efektif.

2. KAJIAN TEORITIS

Penilaian terhadap kesiapan pemerintah digital telah menjadi fokus utama dalam berbagai kajian e-Government. Di tingkat global, beberapa indeks seperti United Nations E-Government Development Index (EGDI) dan Digital Government Index (DGI) dari OECD telah digunakan untuk mengukur kemajuan penerapan pemerintahan digital (OECD, 2020; United Nations, 2022). Di Indonesia, upaya serupa dilakukan melalui pengembangan Indeks Sistem Pemerintahan Berbasis Elektronik (SPBE), yang menilai aspek-aspek strategi, tata kelola, layanan, dan kebijakan internal instansi pemerintah (Kementerian PANRB, 2021).

Studi-studi sebelumnya menyoroti pentingnya SPBE sebagai alat ukur strategi dalam transformasi digital sektor publik (Wibowo & Utomo, 2021).

Namun sebagian besar penelitian mengenai SPBE masih fokus pada analisis evaluatif atau perbandingan nilai indeks antar instansi tanpa mempertimbangkan pendekatan berbasis kelompok atau segmentasi. Beberapa penelitian mulai memperkirakan pendekatan geografis atau spasial dalam menilai kesiapan digital, seperti yang dilakukan oleh Hermawan et al. (2022), yang menganalisis disparitas indeks SPBE antar provinsi di Indonesia. Studi ini mengindikasikan perlunya analisis yang lebih granular untuk memahami konteks lokal dalam kebijakan pengembangan digital.

Dalam konteks ini, pendekatan segmentasi atau klasterisasi wilayah berdasarkan indikator kesiapan digital menjadi relevan untuk mengidentifikasi pola tersembunyi dalam data dan merancang intervensi yang lebih tepat sasaran. Beberapa penelitian di tingkat internasional telah mengadopsi metode clustering untuk pemetaan kesiapan digital antar negara atau wilayah, seperti yang dilakukan oleh Wahab dkk. (2020), yang menggunakan K-Means untuk mengelompokkan negara berdasarkan infrastruktur digital dan layanan e-pemerintah. Pendekatan serupa dapat diterapkan di tingkat sub-nasional seperti kabupaten/kota, namun penerapannya masih jarang ditemukan, khususnya dalam konteks SPBE di Indonesia.

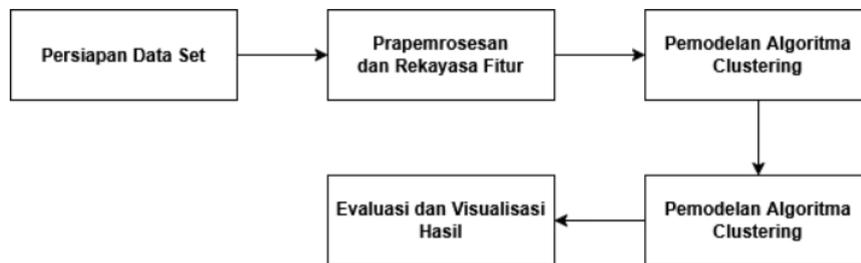
Lebih lanjut, penerapan algoritma unsupervised learning seperti K-Means, DBSCAN, dan Agglomerative Clustering telah terbukti efektif dalam menganalisis struktur data yang kompleks tanpa label, termasuk untuk tujuan pengambilan keputusan kebijakan (Bhuyan et al., 2021). Dalam studi kebijakan publik, metode ini digunakan untuk mengelompokkan wilayah, memahami pola perilaku sosial, dan bahkan mengidentifikasi daerah-daerah prioritas untuk digitalisasi program (Kankanhalli et al., 2021). Meskipun demikian, penerapan metode ini dalam konteks transformasi digital pemerintahan lokal masih sangat terbatas, sehingga membuka ruang kontribusi ilmiah yang signifikan.

Oleh karena itu, penelitian ini menempatkan diri pada celah-celah tersebut dengan menerapkan pendekatan unsupervised learning untuk melakukan segmentasi wilayah kabupaten/kota berdasarkan indeks SPBE. Upaya ini diharapkan dapat memberikan landasan analitis bagi penyusunan strategi pembangunan digital yang lebih berbasis data dan kontekstual..

3. METODE PENELITIAN

Secara garis besar Langkah-langkah penelitian yang dilakukan dapat dilihat pada gambar

1. diagram blok penelitian dibawah ini :



Gambar 1. Diagram Blok Penelitian

a. Persiapan Datatet

Penelitian ini menggunakan dataset publik yang memuat nilai indeks Sistem Pemerintahan Berbasis Elektronik (SPBE) tahun 2023 untuk seluruh kabupaten dan kota di Provinsi Jawa Barat, Indonesia. Dataset ini mencakup sejumlah komponen penilaian seperti kebijakan internal, tata kelola, manajemen informasi teknologi, dan layanan publik berbasis digital. Setiap entitas wilayah (kabupaten/kota) diwakili oleh satu baris data dengan beberapa fitur numerik yang mencerminkan skor dari tiap komponen SPBE.

Dataset diperoleh dari sumber resmi pemerintah dan telah melalui proses kuras awal. Secara total, data mencakup 27 wilayah administratif (18 kabupaten dan 9 kota) sebagai unit analisis. Untuk menjaga akurasi analisis, hanya fitur numerik yang digunakan sebagai input bagi klasterisasi algoritma.

b. Prapemrosesan dan Rekayasa Fitur

Sebelum dilakukan proses klasterisasi, data terlebih dahulu diproses melalui beberapa tahapan preprocessing sebagai berikut:

1. Normalisasi Data: Semua fitur dinormalisasi menggunakan metode Min-Max Scaling ke rentang $[0,1]$ untuk memastikan bahwa setiap fitur memiliki kontribusi yang seimbang dalam proses klasterisasi.
2. Menangani Nilai yang Hilang: Meskipun data relatif bersih, pengecekan nilai hilang dilakukan. Jika terdapat nilai kosong, diisi dengan nilai rata-rata dari fitur terkait.
3. Reduksi Dimensi (opsional): Untuk keperluan visualisasi dan validasi kluster, Principal Component Analysis (PCA) digunakan untuk memproyeksikan data ke dua dimensi utama. Namun, PCA tidak digunakan sebagai input utama dalam proses klasterisasi.

c. **Pemodelan Algoritma Clustering**

Penelitian ini mengimplementasikan tiga metode unsupervised learning untuk melakukan segmentasi wilayah berdasarkan skor SPBE, yaitu:

1. **K-Means Clustering:** Algoritma ini membagi data ke dalam k kluster berdasarkan minimisasi jarak Euclidean terhadap centroid masing-masing kluster. Nilai k ditentukan dengan menggunakan metode Elbow dan Silhouette Score.
2. **DBSCAN (Pengelompokan Spasial Berbasis Kepadatan Aplikasi dengan Kebisingan):** Digunakan untuk mengidentifikasi kluster berdasarkan kepadatan lokal. DBSCAN dapat mengenali kluster dengan bentuk arbitrer dan mendeteksi outlier yang tidak termasuk dalam kluster manapun.
3. **Agglomerative Clustering:** Suatu metode hirarki clustering bottom-up yang menggabungkan titik-titik data secara bertahap berdasarkan kesamaan jarak. Dendrogram digunakan untuk menentukan jumlah kluster optimal secara visual.

Pemilihan algoritma ketiga ini bertujuan untuk mengeksplorasi kelebihan pendekatan berbasis centroid, berbasis kepadatan, dan hierarki dalam konteks segmentasi kesiapan wilayah digital.

d. **Evaluasi dan Visualisasi Hasil.**

Mengingat klusterisasi adalah pendekatan tanpa pengawasan, evaluasi dilakukan menggunakan metrik validasi internal:

1. **Silhouette Coefficient:** Digunakan untuk mengukur seberapa baik suatu objek cocok dengan klusternya sendiri dibandingkan dengan kluster lain.
2. **Davies-Bouldin Index:** Mengukur jarak antar kluster dan konsistensi internal kluster.
3. **Calinski-Harabasz Score:** Rasio antara variasi antar kluster terhadap variasi dalam kluster, semakin tinggi semakin baik.

Selain itu, visualisasi dua dimensi dari hasil klusterisasi juga digunakan untuk menilai keterpisahan dan kepadatan tiap kluster, terutama pada K-Means dan Agglomerative Clustering. Untuk mendukung interpretasi kebijakan analisis, karakteristik setiap kluster dilakukan dengan menghitung rata-rata skor dari setiap fitur di masing-masing kluster.

4. HASIL DAN PEMBAHASAN

Deskripsi Data Set

Dataset SPBE tahun 2023 yang dianalisis dalam penelitian ini mencakup 27 entitas wilayah administratif di Provinsi Jawa Barat, terdiri dari 18 kabupaten dan 9 kota. Dataset diunggah ke lingkungan Google Colaboratory dalam format Comma Separated Values (.csv) dan dibaca menggunakan fungsi `read_csv()` dari pustaka `pandas`.

Berdasarkan eksplorasi awal, dataset memiliki 9 kolom utama, yakni:

1. Nama Wilayah (Kabupaten/Kota)
2. Nilai Indeks SPBE (total)
3. Domain Tata Kelola SPBE
4. Aspek Kelembagaan
5. Aspek Strategi dan Perencanaan
6. Aspek Teknologi Informasi dan Komunikasi
7. Aspek Layanan SPBE
8. Kategori Indeks SPBE (kualitatif, seperti "Baik", "Sangat Baik")
9. Kategori Domain Tata Kelola (kualitatif)

Sebagian besar kolom bersifat numerik dan merepresentasikan skor kuantitatif masing-masing aspek, sehingga sangat sesuai untuk dianalisis menggunakan metode `unsupervised learning`. Kolom non-numerik (seperti nama wilayah dan kategori) disimpan untuk keperluan interpretasi hasil klasterisasi, namun tidak digunakan secara langsung dalam proses model pelatihan. Dataset tersebut dapat dilihat pada gambar 2. dibawah ini.

id	kode_provinsi	nama_provinsi	kode_kabupaten_kota	nama_kabupaten_kota	predikat	indeks_spbe	satuan	tahun
0	1	32	JAWA BARAT	3201	KABUPATEN BOGOR	BAIK	2.72	POIN 2021
1	2	32	JAWA BARAT	3202	KABUPATEN SUKABUMI	KURANG	1.56	POIN 2021
2	3	32	JAWA BARAT	3203	KABUPATEN CIANJUR	CUKUP	2.46	POIN 2021
3	4	32	JAWA BARAT	3204	KABUPATEN BANDUNG	BAIK	2.98	POIN 2021
4	5	32	JAWA BARAT	3205	KABUPATEN GARUT	BAIK	3.25	POIN 2021

Gambar 2. Dataset

Hasil pengecekan juga menunjukkan bahwa tidak terdapat data kosong (`missing values`) dalam kolom numerik yang digunakan, sehingga tidak diperlukan proses imputasi. Distribusi nilai indeks SPBE menunjukkan variasi antar wilayah yang cukup signifikan, yang mengindikasikan adanya potensi ketimpangan kesiapan digital antardaerah—menegaskan urgensi dari segmentasi wilayah berbasis data ini.

Prapemrosesan dan Rekayasa Fitur

Sebelum dilakukan proses klasterisasi, data numerik pada dataset SPBE dinormalisasi terlebih dahulu guna menyetarakan skala antar fitur. Hal ini penting karena algoritma unsupervised learning seperti K-Means, DBSCAN, dan Agglomerative Clustering sangat sensitif terhadap perbedaan skala nilai. Fitur dengan skala lebih besar dapat mendominasi proses perhitungan jarak (distance calculation) dan memengaruhi hasil pembentukan klaster.

Pada dataset ini, kolom-kolom numerik yang digunakan meliputi: id, kode_provinsi, kode_kabupaten_kota, indeks_spbe, dan tahun. Meskipun beberapa fitur bersifat identifikasi administratif, seluruh kolom numerik tetap dinormalisasi terlebih dahulu menggunakan metode Min-Max Scaling. Metode ini merescale nilai-nilai dalam setiap fitur ke dalam rentang [0, 1], dengan rumus:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Hasil normalisasi dapat dilihat pada gambar 3. Dibawah ini :

```

Data Asli (5 baris pertama):
  id  kode_provinsi  nama_provinsi  kode_kabupaten_kota  nama_kabupaten_kota \
0  1             32  JAWA BARAT             3201  KABUPATEN BOGOR
1  2             32  JAWA BARAT             3202  KABUPATEN SUKABUMI
2  3             32  JAWA BARAT             3203  KABUPATEN CIANJUR
3  4             32  JAWA BARAT             3204  KABUPATEN BANDUNG
4  5             32  JAWA BARAT             3205  KABUPATEN GARUT

  predikat  indeks_spbe  satuan  tahun
0  BAIK      2.72  POIN  2021
1  KURANG   1.56  POIN  2021
2  CUKUP    2.46  POIN  2021
3  BAIK     2.98  POIN  2021
4  BAIK     3.25  POIN  2021

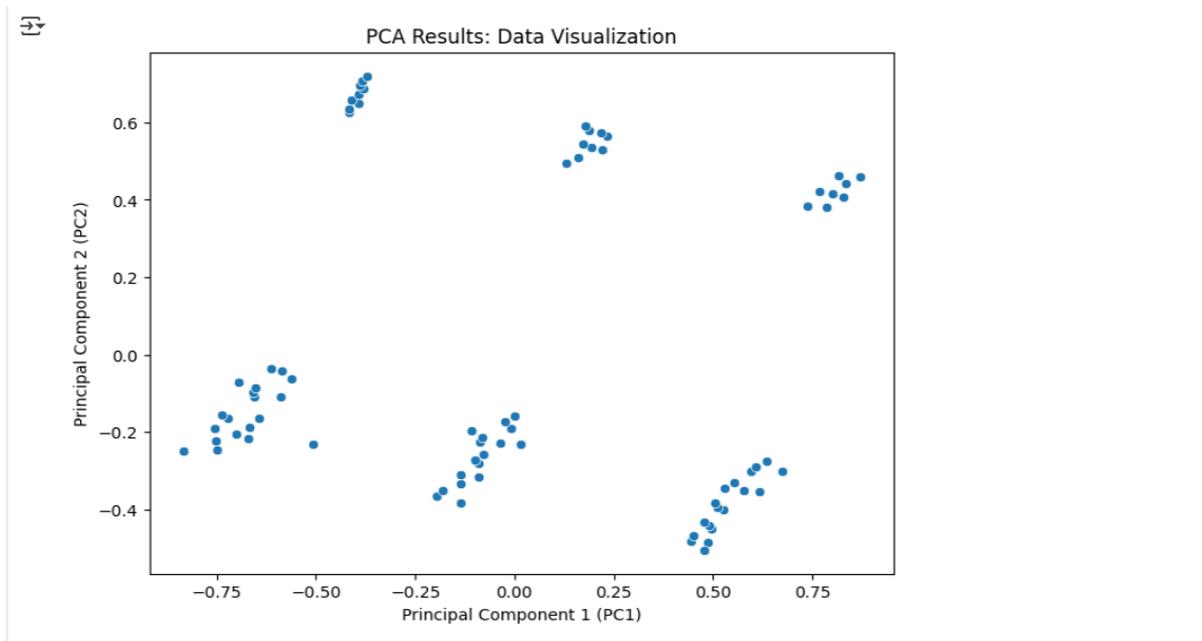
Data Numerik (5 baris pertama):
  id  kode_provinsi  kode_kabupaten_kota  indeks_spbe  tahun
0  1             32             3201          2.72  2021
1  2             32             3202          1.56  2021
2  3             32             3203          2.46  2021
3  4             32             3204          2.98  2021
4  5             32             3205          3.25  2021

Data Normalisasi (5 baris pertama):
  id  kode_provinsi  kode_kabupaten_kota  indeks_spbe  tahun
0  0.000000      0.0             0.000000      0.411348  0.0
1  0.012658      0.0             0.012821      0.000000  0.0
2  0.025316      0.0             0.025641      0.319149  0.0
3  0.037975      0.0             0.038462      0.503546  0.0
4  0.050633      0.0             0.051282      0.599291  0.0

```

Gambar 3. Hasil Normalisasi

Melalui *Principal Component Analysis* PCA pada gambar 4. dibawah ini menunjukkan bahwa data cenderung membentuk tiga kelompok yang relatif terpisah. Kelompok pertama mengklik di sekitar nilai PC1 yang rendah dan PC2 yang tinggi. Kelompok kedua berada di tengah-tengah plot dengan nilai PC1 dan PC2 yang moderat. Sementara kelompok ketiga terletak di area dengan nilai PC1 yang tinggi dan PC2 yang rendah. Terdapat beberapa titik yang tersebar jauh dari kelompok utama, kemungkinan adanya outlier dalam data.

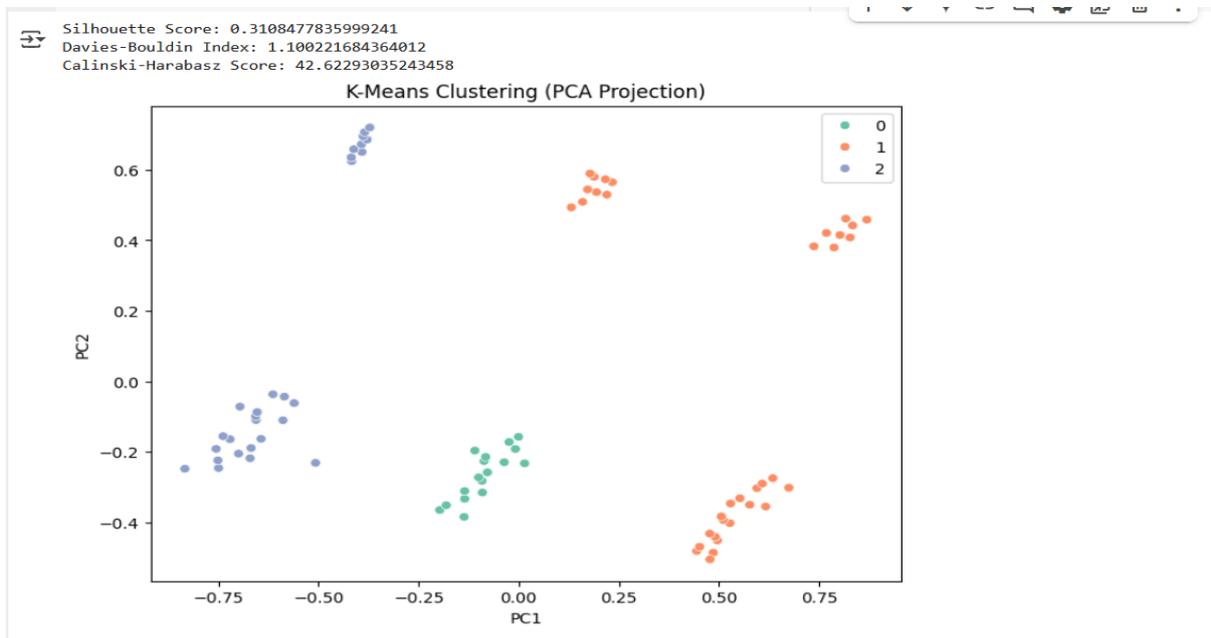


Gambar 4. PCA (*Principal Component Analysis*)

Pemodelan Algoritma Clustering

a. Hasil Clustering Menggunakan Algoritma K-Means

Gambar 5. dibawah ini menunjukkan hasil clustering terhadap data indeks SPBE wilayah Jawa Barat menggunakan algoritma K-Means.



Gambar 5. Hasil Clustering dan evaluasi K-Means

Dari gambar 5. tersebut divisualisasikan dalam dua dimensi melalui proyeksi Principal Component Analysis (PCA). Pada eksperimen ini, jumlah kluster optimal ditentukan sebanyak tiga ($k=3$), berdasarkan analisis awal menggunakan metode silhouette.

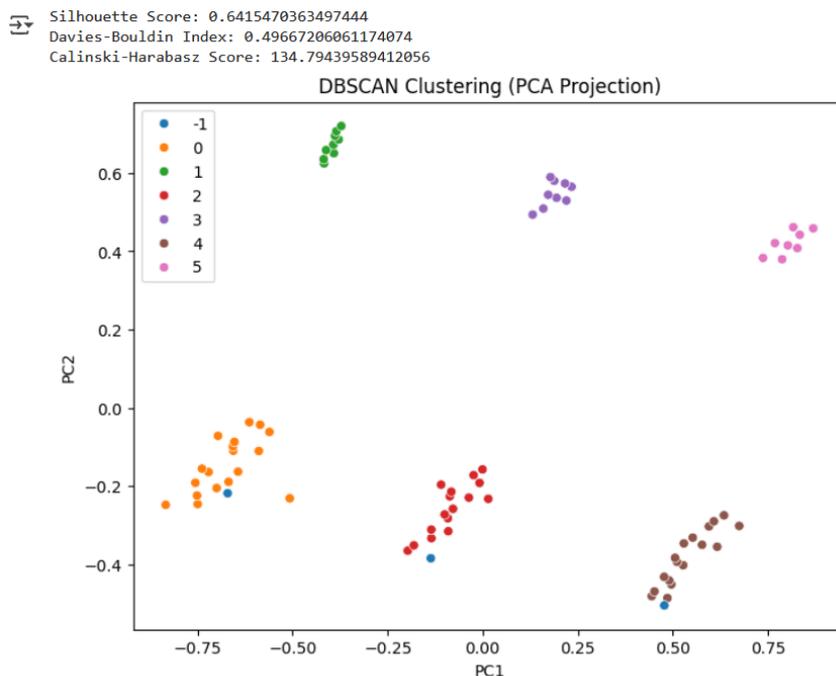
Hasil evaluasi kualitas clustering menghasilkan metrik sebagai berikut:

1. **Silhouette Score:** 0,3108, yang mengindikasikan adanya pemisahan kluster yang cukup baik meskipun belum sempurna.
2. **Davies-Bouldin Index:** 1,1002, menunjukkan tingkat kompaksi dan pemisahan antar kluster yang masih dapat ditingkatkan.
3. **Calinski-Harabasz Score:** 42,6229, mengindikasikan rasio dispersi antar kluster terhadap kompaksi dalam kluster yang relatif baik.

Secara visual, plot tersebut memperlihatkan bahwa algoritma K-Means mampu memisahkan data ke dalam tiga kelompok yang berbeda dengan kohesi internal yang cukup kuat dan pemisahan antar kelompok yang jelas, meskipun terdapat sedikit tumpang tindih. Hasil ini dapat mencerminkan perbedaan tingkat kesiapan digital antar daerah di Provinsi Jawa Barat.

b. Hasil Clustering Menggunakan Algoritma DBSCAN

Gambar 6 dibawah ini memperlihatkan hasil clustering terhadap data indeks SPBE wilayah Jawa Barat menggunakan algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN),



Gambar 6. Hasil Clustering dan evaluasi DBSCAN

Gambar 6. tersebut mengvisualisasikan dalam dua dimensi melalui proyeksi Principal Component Analysis (PCA). Berbeda dengan K-Means, DBSCAN tidak memerlukan penentuan jumlah kluster di awal, melainkan membentuk kluster berdasarkan kepadatan data.

Dalam visualisasi, DBSCAN berhasil mengelompokkan data ke dalam enam kluster utama (label 0 hingga 5), dengan sejumlah kecil data yang diklasifikasikan sebagai noise (ditandai dengan label -1). Setiap kluster menunjukkan kohesi internal yang kuat dan pemisahan antar kluster yang jelas, memperlihatkan kemampuan DBSCAN dalam mendeteksi struktur alami dari data tanpa harus menetapkan jumlah kluster secara eksplisit.

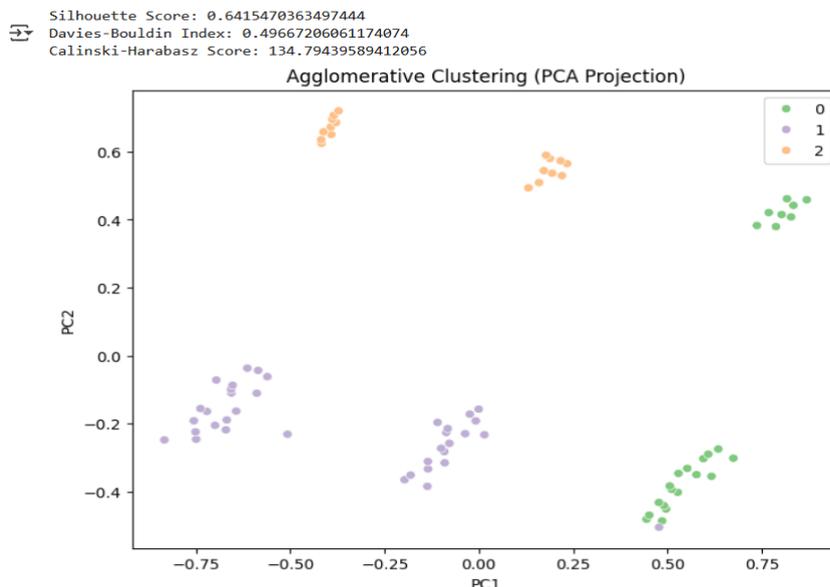
Hasil evaluasi kualitas clustering menghasilkan metrik sebagai berikut:

1. **Silhouette Score:** 0,6415, menunjukkan kualitas pemisahan kluster yang sangat baik.
2. **Davies-Bouldin Index:** 0,4967, mengindikasikan kompaksi intra-kluster yang kuat dan pemisahan antar-kluster yang signifikan.
3. **Calinski-Harabasz Score:** 134,7944, memperkuat bahwa struktur kluster yang terbentuk memiliki dispersi antar-kluster yang tinggi dibandingkan dengan variansi dalam kluster.

Temuan ini memperlihatkan bahwa DBSCAN sangat efektif untuk menggambarkan variasi tingkat kesiapan digital antar wilayah di Provinsi Jawa Barat, khususnya dalam mendeteksi outlier atau daerah-daerah yang memiliki karakteristik SPBE yang unik.

c. Hasil Clustering Menggunakan Algoritma Agglomerative Clustering

Gambar 7 dibawah ini memperlihatkan hasil penerapan algoritma Agglomerative Clustering terhadap data indeks SPBE wilayah Jawa Barat, dengan visualisasi hasil klasterisasi ditampilkan dalam dua dimensi melalui proyeksi Principal Component Analysis (PCA).



Gambar 6. Hasil Clustering dan evaluasi algoritma Agglomerative Clustering

Berdasarkan hasil clustering, diperoleh tiga kluster utama (label 0, 1, dan 2). Setiap kluster memiliki distribusi yang relatif jelas dengan pemisahan yang cukup baik antar kelompok. Tidak

terdapat data yang dikategorikan sebagai noise, mengingat pendekatan hierarkis agglomerative selalu mengelompokkan seluruh data ke dalam kluster tertentu.

Hasil evaluasi clustering dengan metrik internal quality menunjukkan nilai:

1. **Silhouette Score:** 0,6415, mengindikasikan bahwa objek-objek dalam kluster lebih dekat ke klasternya sendiri dibandingkan dengan kluster lain.
2. **Davies-Bouldin Index:** 0,4967, menandakan adanya separasi antar-kluster yang baik dan kompaksi internal yang kuat.
3. **Calinski-Harabasz Score:** 134,7944, memperkuat interpretasi bahwa struktur kluster yang terbentuk cukup terpisah dan terdefinisi dengan baik.

Agglomerative Clustering berhasil membentuk struktur kluster yang cukup solid, memperlihatkan grup-grup wilayah dengan karakteristik indeks SPBE yang mirip. Hasil ini memperlihatkan bahwa metode berbasis hierarki efektif digunakan untuk memahami segmentasi kesiapan digital antar daerah, terutama ketika struktur data bersifat hierarkis atau bertingkat.

Secara keseluruhan hasil klusterisasi dapat di lihat di gambar 7. Dibawah ini

```

Karakteristik Kluster K-Means:
      id  kode_provinsi  kode_kabupaten_kota  indeks_spbe  \
Cluster_KMeans
0      37.0      32.058824      3210.00      2.712941
1      63.0      32.028571      3241.20      3.287429
2      14.5      32.035714      3230.25      2.712143

      tahun
Cluster_KMeans
0      2022.000000
1      2022.742857
2      2021.000000

Karakteristik Kluster DBSCAN:
      id  kode_provinsi  kode_kabupaten_kota  indeks_spbe  \
Cluster_DBSCAN
-1      35.000000      33.0      3208.000000      2.850000
0      10.500000      32.0      3209.000000      2.607778
1      23.000000      32.0      3275.000000      2.908889
2      37.187500      32.0      3210.187500      2.717500
3      50.000000      32.0      3275.000000      2.977778
4      63.647059      32.0      3209.647059      3.350588
5      76.500000      32.0      3274.500000      3.526250

      tahun
Cluster_DBSCAN
-1      2022.0
0      2021.0
1      2021.0
2      2022.0
3      2022.0
4      2023.0
5      2023.0

Karakteristik Kluster Agglomerative Clustering:
      id  kode_provinsi  kode_kabupaten_kota  indeks_spbe  \
Cluster_Agglo
0      67.760000      32.000000      3230.400000      3.406800
1      24.027027      32.081081      3209.432432      2.674865
2      36.500000      32.000000      3275.000000      2.943333

      tahun
Cluster_Agglo
0      2023.000000
1      2021.513514
2      2021.500000

```

Gambar 8. Hasil perbandingan algoritma K-Means, DBSCAN dan Agglomerative Clustering.

Hasil klasterisasi menggunakan algoritma K-Means menghasilkan tiga klaster utama dengan karakteristik yang berbeda. Cluster 0 menunjukkan wilayah dengan rata-rata indeks SPBE sebesar 2,7129 dan tahun data 2022, yang merepresentasikan daerah dengan kesiapan digital sedang berdasarkan data terbaru. Cluster 1 memiliki rata-rata indeks SPBE yang lebih tinggi, yakni 3,2874, dengan rata-rata tahun data mendekati 2023, menggambarkan wilayah yang lebih siap secara digital dan berasal dari data yang lebih mutakhir. Sementara itu, Cluster 2, dengan rata-rata indeks SPBE sebesar 2,7121, menunjukkan karakteristik kesiapan digital serupa dengan Cluster 0, namun berasal dari tahun data yang lebih lama, yaitu 2021. Secara umum, K-Means mampu mengelompokkan wilayah berdasarkan kombinasi tingkat kesiapan digital dan dimensi temporal.

Pada klasterisasi menggunakan DBSCAN, terbentuk enam klaster utama beserta satu klaster noise (-1). Klaster noise mencakup wilayah dengan rata-rata indeks SPBE 2,8500 dari tahun 2022, yang tidak cocok dengan karakteristik klaster lain. Cluster 0 dan Cluster 1 mewakili wilayah dengan indeks SPBE relatif rendah hingga sedang, berasal dari tahun 2021. Cluster 2 dan Cluster 3 mencakup wilayah dengan indeks SPBE yang sedikit lebih tinggi, dari tahun 2022. Sementara Cluster 4 dan Cluster 5 menunjukkan wilayah dengan indeks SPBE yang lebih baik, masing-masing sebesar 3,3506 dan 3,5263, dan berasal dari tahun 2023, menandakan kesiapan digital yang sudah sangat baik dan mutakhir. DBSCAN terbukti efektif membedakan wilayah tidak hanya berdasarkan tingkat kesiapan digital, tetapi juga berdasarkan dimensi waktu, sekaligus mampu mengidentifikasi data outlier.

Sementara itu, hasil klasterisasi menggunakan Agglomerative Clustering membentuk tiga klaster dengan pembagian yang cukup tegas. Cluster 0 memiliki rata-rata indeks SPBE sebesar 3,4068 dan tahun data 2023, menunjukkan wilayah dengan kesiapan digital sangat baik dan data paling baru. Sebaliknya, Cluster 1 dan Cluster 2 terdiri dari wilayah dengan indeks SPBE yang lebih rendah, masing-masing sebesar 2,6749 dan 2,9433, keduanya berasal dari data tahun 2021. Meskipun tahun data hampir seragam, perbedaan indeks SPBE cukup jelas, mengindikasikan kemampuan metode agglomerative dalam memisahkan wilayah berdasarkan level kesiapan digital dengan struktur hierarkis.

Secara keseluruhan, ketiga metode klasterisasi menunjukkan pola pengelompokan berdasarkan kombinasi antara indeks kesiapan digital (indeks SPBE) dan dimensi waktu (tahun pengambilan data). DBSCAN menghasilkan segmentasi paling rinci dan mampu mendeteksi anomali, sedangkan Agglomerative Clustering menampilkan pemisahan berbasis hierarki kesiapan digital, dan K-Means memberikan pembagian umum berdasarkan tingkat indeks dan

tahun. Pilihan metode klusterisasi ke depan dapat disesuaikan dengan kebutuhan penelitian, apakah fokus pada ketelitian segmentasi, deteksi outlier, atau penyederhanaan klasifikasi.

5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa pendekatan klusterisasi menggunakan metode unsupervised learning dapat mengungkap pola tersembunyi dalam data SPBE daerah. Ketiga algoritma yang digunakan—K-Means, DBSCAN, dan Agglomerative Clustering—berhasil membentuk kluster wilayah berdasarkan tingkat kesiapan digital dan dimensi waktu pengambilan data. Hasil menunjukkan bahwa:

1. K-Means efektif dalam membentuk kluster berdasarkan kombinasi nilai indeks SPBE dan tahun data, meskipun cenderung menghasilkan segmentasi umum.
2. DBSCAN mampu mengidentifikasi struktur kluster alami, termasuk mendeteksi outlier, serta menghasilkan segmentasi yang paling tajam dan kaya informasi.
3. Agglomerative Clustering membentuk kluster dengan struktur hierarkis yang baik, berguna untuk memahami tingkat kesiapan digital secara bertingkat.

Temuan ini menegaskan pentingnya pendekatan segmentasi dalam perumusan kebijakan digital yang inklusif dan berbasis bukti. Ke depan, penelitian ini dapat dikembangkan dengan melibatkan variabel tambahan seperti indeks kematangan SDM digital, infrastruktur TIK, serta pendekatan explainable AI (seperti SHAP/LIME) untuk mendukung transparansi interpretasi model.

DAFTAR REFERENSI

- Al-Sharafi, M. A., Arshah, R. A., Abu-Shanab, E., & Alajmi, Q. (2020). Digital transformation in the public sector: A systematic literature review. *IEEE Access*, 8, 21279-21293. <https://doi.org/10.1109/ACCESS.2020.2964214>
- Kementerian PANRB. (2021). Peraturan Menteri PANRB No. 59 Tahun 2020 tentang Pemantauan dan Evaluasi Sistem Pemerintahan Berbasis Elektronik.
- Nugroho, R., & Subekti, R. (2022). Evaluasi Indeks SPBE di Pemerintah Daerah: Analisis Tantangan dan Strategi Peningkatan. *Jurnal Transformasi Digital Pemerintahan*, 2(1), 15-27.
- Bhuyan, M. K., Bhowmick, P., & Roy, S. S. (2021). A survey of clustering techniques in smart city applications. *Artificial Intelligence Review*, 54(3), 2257–2284. <https://doi.org/10.1007/s10462-020-09839-1>
- Hermawan, R., Nugroho, H., & Putri, I. A. (2022). Disparitas Kesiapan Pemerintah Daerah dalam Implementasi SPBE di Indonesia. *Jurnal Administrasi Digital*, 1(2), 45–58.

- Kankanhalli, A., Charalabidis, Y., & Mellouli, S. (2021). Digital transformation in governments: A review of public sector research. *Government Information Quarterly*, 38(1), 101398. <https://doi.org/10.1016/j.giq.2020.101398>
- Kementerian PANRB. (2021). Indeks SPBE Nasional 2021.
- OECD. (2020). *Digital Government Index: 2019 Results*. OECD Publishing. <https://doi.org/10.1787/4de9f5bb-en>
- United Nations. (2022). *UN E-Government Survey 2022: The Future of Digital Government*.
- Wahab, S., Roslan, A., & Hussin, N. (2020). Clustering countries by e-government readiness using K-means algorithm. *Journal of Theoretical and Applied Information Technology*, 98(11), 2133–2142.
- Wibowo, H., & Utomo, R. (2021). Analisis Peran SPBE dalam Peningkatan Kualitas Layanan Publik Pemerintah Daerah. *Jurnal Transformasi Digital Pemerintahan*, 2(2), 65–78.
- Zuiderwijk, A., Chen, Y. C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, 38(3), 101577. <https://doi.org/10.1016/j.giq.2021.101577>