



Analisis Ternak Menggunakan K-Means Clustering Dalam Business Intelligence

Sulaiman Saverosagi*, Safrizal

Fakultas Teknologi dan Desain, Program Studi Sistem Informasi, Universitas Pembangunan Jaya, Tangerang Selatan, Indonesia

Email: ^{1,*}sulaiman.saverosagi@student.upj.ac.id, ²safrizal.abdurahman@upj.ac.id

Email Penulis Korespondensi: sulaiman.saverosagi @student.upj.ac.id

Abstrak—Penelitian ini bertujuan untuk menganalisis pola populasi ternak dan mengelompokkan wilayah di Provinsi Jawa Tengah menggunakan data tahun 2020–2023 yang mencakup sapi, kambing, dan ayam dari sumber resmi pemerintah. Metode Knowledge Discovery in Database (KDD) dan algoritma K-Means Clustering digunakan, dengan penentuan jumlah cluster optimal melalui metode Elbow dan Silhouette Score. Hasil menunjukkan bahwa jumlah cluster optimal adalah dua ($K=2$) dengan nilai Silhouette Score sebesar 0,328 yang mengindikasikan struktur cluster yang relatif lemah. Meskipun demikian, hasil clustering tetap mampu menunjukkan segmentasi yang bermakna ketika dikombinasikan dengan analisis Business Intelligence. Cluster 0 terdiri dari wilayah dengan populasi lebih rendah namun memiliki pertumbuhan lebih tinggi, sedangkan Cluster 1 memiliki populasi lebih tinggi dengan pertumbuhan rendah atau negatif. Analisis menunjukkan bahwa hubungan antara populasi, pertumbuhan, dan produksi tidak bersifat linear, di mana produksi tinggi tidak selalu diikuti pertumbuhan yang baik. Temuan ini menegaskan pentingnya membedakan antara kapasitas produksi saat ini dan potensi pertumbuhan di masa depan dalam mendukung pengambilan keputusan berbasis data di sektor peternakan.

Kata Kunci: K-Means Clustering; Populasi Ternak; Data Mining; Business Intelligence

Abstract—This study aims to analyze livestock population patterns and classify regions in Central Java using data from 2020–2023 covering cattle, goats, and chickens from official sources. The Knowledge Discovery in Database (KDD) framework and K-Means Clustering were applied, with the optimal number of clusters determined using the Elbow Method and Silhouette Score. The results show that the optimal number of clusters is two ($K=2$), with a Silhouette Score of 0.328, indicating a relatively weak clustering structure with potential overlap. Despite this limitation, the results reveal meaningful segmentation when combined with Business Intelligence analysis. Cluster 0 represents regions with lower population but higher growth, while Cluster 1 represents regions with higher population but lower or negative growth. Further analysis indicates that the relationship between population, growth, and production is not linear, where high production does not necessarily correspond to strong growth. These findings highlight the importance of distinguishing between current production capacity and future growth potential, providing more informative insights for data-driven decision-making in livestock sector management.

Keywords: K-Means Clustering; Livestock Population; Data Mining; Business Intelligence

1. PENDAHULUAN

Sektor peternakan merupakan salah satu komponen penting dalam mendukung ketahanan pangan dan perekonomian daerah, khususnya di Provinsi Jawa Tengah. Populasi ternak seperti sapi, kambing, dan ayam berperan sebagai sumber utama penyedia protein hewani bagi masyarakat. Oleh karena itu, perkembangan populasi ternak perlu dipantau dan dianalisis secara berkelanjutan agar dapat mendukung perencanaan dan pengambilan keputusan yang lebih efektif di sektor peternakan [1]. Hal ini didukung oleh penelitian sebelumnya yang menyatakan bahwa tanpa pemahaman yang baik terhadap kondisi populasi peternakan, akan sulit untuk mengoptimalkan hasil produksi serta merumuskan strategi pengembangan yang tepat.

Seiring dengan meningkatnya ketersediaan data populasi ternak dari berbagai kabupaten/kota, muncul tantangan dalam mengolah dan memanfaatkan data tersebut secara optimal. Data yang tersedia umumnya masih disajikan dalam bentuk statistik deskriptif seperti laporan rutin yang dikeluarkan oleh Badan Pusat Statistik (BPS), yang hanya menggambarkan kondisi populasi ternak tanpa mampu mengidentifikasi pola tersembunyi atau karakteristik kemiripan antar wilayah.

Akibatnya, informasi yang dihasilkan cenderung terbatas pada penyajian angka tanpa memberikan pemahaman yang lebih mendalam terkait segmentasi wilayah berdasarkan potensi dan karakteristik peternakan. Oleh karena itu, terdapat kesenjangan penelitian (research gap) antara penyajian data secara deskriptif dengan kebutuhan analisis yang mampu mengelompokkan wilayah secara berbasis pola dan karakteristik data. Penelitian ini berupaya mengisi kesenjangan tersebut dengan menerapkan teknik clustering untuk menghasilkan segmentasi wilayah yang lebih informatif dibandingkan analisis statistik konvensional.

Hal ini sejalan dengan penelitian yang menyatakan bahwa untuk menunjang produktivitas sektor peternakan diperlukan data yang mampu menggambarkan persebaran ternak serta pengelompokan wilayah berdasarkan karakteristik yang dimiliki [2].

Salah satu pendekatan yang dapat digunakan adalah teknik data mining, khususnya metode clustering. Terdapat berbagai algoritma clustering yang dapat digunakan, seperti K-Means, DBSCAN, dan Hierarchical Clustering, yang masing-masing memiliki karakteristik berbeda. DBSCAN efektif dalam mendeteksi pola berbasis densitas dan menangani noise, sedangkan Hierarchical Clustering mampu membentuk struktur cluster bertingkat namun memiliki kompleksitas komputasi yang lebih tinggi [3].

Dalam penelitian ini, metode K-Means dipilih karena memiliki keunggulan dalam mengelompokkan data numerik dengan jumlah besar secara efisien serta mampu menghasilkan cluster yang jelas berdasarkan centroid. Selain itu,



karakteristik data populasi ternak yang bersifat numerik dan relatif terstruktur membuat K-Means lebih sesuai digunakan dibandingkan metode lainnya. Pendekatan clustering digunakan untuk mengelompokkan objek ke dalam beberapa kelompok yang relatif homogen berdasarkan karakteristik tertentu, sehingga objek dalam satu kelompok memiliki tingkat kemiripan yang lebih tinggi dibandingkan dengan kelompok lainnya.

Untuk memastikan kualitas hasil pengelompokan, digunakan metode evaluasi Silhouette Score dalam menentukan jumlah cluster yang optimal.

Selain itu, pemanfaatan Business Intelligence dalam penelitian ini berperan dalam menyajikan hasil analisis dalam bentuk visualisasi yang informatif dan mudah dipahami. Visualisasi tersebut diharapkan dapat membantu pemangku kepentingan dalam menginterpretasikan hasil clustering secara lebih efektif, sehingga dapat mendukung proses pengambilan keputusan berbasis data [4].

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis perkembangan populasi ternak di Jawa Tengah serta mengelompokkan wilayah berdasarkan karakteristiknya menggunakan metode K-Means Clustering. Hasil penelitian diharapkan dapat memberikan gambaran distribusi populasi ternak antar wilayah dan menjadi dasar dalam penyusunan kebijakan yang lebih efektif dan tepat sasaran di sektor peternakan.

2. METODOLOGI PENELITIAN

2.1 Sumber dan Jenis Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari sumber resmi pemerintah. Data populasi ternak diperoleh melalui portal data terbuka Pemerintah Provinsi Jawa Tengah, yaitu melalui situs data.jatengprov.go.id. Selain itu, data pendukung berupa produksi daging dan telur diperoleh dari Badan Pusat Statistik [2]. Variabel yang digunakan dalam penelitian ini berasal dari data mentah yang diperoleh dari sumber tersebut dan dijelaskan pada Tabel 1.

Tabel 1. Variable Jenis data

Variabel	Tipe Data	Deskripsi
Kab/Kota	Kategorikal	Nama kabupaten/kota di Provinsi Jawa Tengah
Sapi	Numerik	Jumlah populasi ternak sapi
Kambing	Numerik	Jumlah populasi ternak kambing
Ayam	Numerik	Jumlah populasi ternak ayam
Growth	Numerik	Tingkat pertumbuhan populasi ternak
Produksi Daging Sapi	Numerik	Produksi daging ternak sapi (kg)
Produksi Daging Kambing	Numerik	Produksi daging ternak kambing (kg)
Produksi Daging Kerbau	Numerik	Produksi daging ternak kerbau (kg)
Produksi Telur Ayam Kampung	Numerik	Produksi telur ayam kampung (kg)
Produksi Telur Ayam Petelur	Numerik	Produksi telur ayam petelur (kg)
Produksi Telur Ayam Buras	Numerik	Produksi telur ayam buras (kg)
Produksi Telur Itik	Numerik	Produksi telur itik (kg)

Data yang digunakan mencakup periode tahun 2020 hingga 2023 pada tingkat kabupaten/kota di Provinsi Jawa Tengah. Variabel utama yang digunakan dalam analisis meliputi jumlah populasi ternak sapi, kambing, dan ayam. Variabel tambahan berupa growth digunakan sebagai informasi pendukung dalam tahap analisis untuk memberikan gambaran kondisi populasi ternak antar wilayah.

2.2 Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini dilakukan dengan metode dokumentasi, yaitu dengan mengakses dan mengunduh data dari sumber resmi yang telah tersedia secara publik melalui portal data Pemerintah Provinsi Jawa Tengah serta publikasi dari Badan Pusat Statistik. Penggunaan data sekunder yang diperoleh dari sumber resmi pemerintah dan publikasi institusi terkait merupakan pendekatan yang umum digunakan dalam penelitian data mining karena mampu menyediakan data yang terstruktur dan relevan untuk proses analisis [5].

Data yang diperoleh kemudian dikompilasi dan digabungkan dari beberapa sumber sesuai kebutuhan penelitian. Selanjutnya, dilakukan penyesuaian format dan struktur data agar konsisten, sehingga dapat digunakan dalam proses pengolahan dan analisis pada tahap berikutnya.

2.3 Bahan dan Alat

Bahan yang digunakan dalam penelitian ini berupa dataset populasi ternak yang mencakup jumlah ternak sapi, kambing, dan ayam pada tingkat kabupaten/kota di Provinsi Jawa Tengah periode 2020–2023. Selain itu, digunakan pula data pendukung berupa produksi daging dan telur yang diperoleh dari publikasi resmi Badan Pusat Statistik untuk memperkaya proses analisis dan interpretasi hasil. Data sekunder dari instansi resmi merupakan sumber yang umum digunakan dalam penelitian kuantitatif untuk mendukung analisis dan pengambilan kesimpulan [2].



Alat yang digunakan dalam penelitian ini adalah perangkat lunak berbasis bahasa pemrograman Python. Proses pengolahan dan analisis data dilakukan dengan bantuan pustaka pandas untuk pengolahan data, scikit-learn untuk analisis clustering, serta matplotlib untuk visualisasi data.

2.4 Tahapan Penelitian

Penelitian ini menggunakan metode Knowledge Discovery in Database (KDD) sebagai kerangka kerja dalam proses analisis data. KDD merupakan serangkaian tahapan yang digunakan untuk mengekstraksi informasi atau pengetahuan yang bermanfaat dari kumpulan data berukuran besar melalui proses yang sistematis dan terstruktur [6].

Tahapan dalam KDD meliputi data selection, data preprocessing, data transformation, data mining, evaluation, serta knowledge. Setiap tahapan memiliki peran penting dalam memastikan bahwa data yang digunakan telah melalui proses pengolahan yang tepat sehingga hasil analisis yang diperoleh menjadi lebih akurat dan dapat diinterpretasikan dengan baik.

Dalam penelitian ini, setiap tahapan KDD diimplementasikan secara terintegrasi untuk memastikan bahwa proses analisis berjalan secara sistematis. Tahap data selection digunakan untuk menentukan variabel yang relevan, dilanjutkan dengan data preprocessing dan data transformation untuk meningkatkan kualitas dan konsistensi data. Selanjutnya, tahap data mining dilakukan menggunakan algoritma K-Means untuk mengidentifikasi pola pengelompokan wilayah, yang kemudian dievaluasi menggunakan metode Silhouette Score dan Elbow Method. Hasil dari proses tersebut diinterpretasikan pada tahap knowledge untuk menghasilkan insight yang dapat mendukung pengambilan keputusan berbasis data di sektor peternakan [7].

2.4.1 Data Selection

Tahap data selection merupakan proses pemilihan data yang relevan dari keseluruhan dataset yang tersedia untuk digunakan dalam penelitian. Pada penelitian ini, data yang dipilih adalah data populasi ternak pada tingkat kabupaten/kota di Provinsi Jawa Tengah periode 2020–2023. Variabel utama yang digunakan dalam proses clustering meliputi jumlah populasi ternak sapi, kambing, dan ayam.

Selain itu, data pendukung berupa produksi dan variabel growth tetap dikumpulkan, namun tidak digunakan sebagai input utama dalam proses clustering. Hal ini dikarenakan penelitian berfokus pada pemetaan struktur populasi sebagai representasi kapasitas dasar wilayah, sedangkan growth digunakan sebagai variabel interpretatif untuk menganalisis dinamika perkembangan antar cluster.

2.4.2 Data Preprocessing

Tahap data preprocessing dilakukan untuk memastikan bahwa data berada dalam kondisi bersih, konsisten, dan siap dianalisis. Proses ini bertujuan untuk meningkatkan kualitas data sehingga hasil clustering menjadi lebih akurat dan representatif [8]. Preprocessing meliputi pembersihan data, penanganan missing value, penghapusan data tidak relevan, serta integrasi dataset. Selain itu dilakukan seleksi atribut agar hanya variabel yang relevan digunakan dalam analisis.

Selain tahapan tersebut, dilakukan pula proses validasi data untuk memastikan tidak terdapat duplikasi data maupun inkonsistensi nilai antar variabel. Penanganan terhadap nilai ekstrem (outlier) juga diperhatikan secara selektif agar tidak mengganggu proses pembentukan cluster, namun tetap mempertahankan informasi penting yang mencerminkan kondisi riil data. Proses ini menjadi penting karena kualitas data yang baik akan berpengaruh langsung terhadap hasil clustering, mengingat algoritma K-Means sangat sensitif terhadap distribusi dan skala data.

2.4.3 Data Transformation

Tahap data transformation merupakan proses mengubah data ke dalam format yang sesuai dengan kebutuhan algoritma. Data diubah dari long format menjadi wide format, kemudian dilakukan normalisasi menggunakan metode MinMaxScaler untuk menyamakan skala antar variabel sehingga tidak terjadi dominasi nilai tertentu dalam proses perhitungan jarak.

Transformasi dari long format ke wide format dilakukan untuk memastikan bahwa setiap baris data merepresentasikan satu entitas wilayah secara utuh dengan seluruh variabel yang relevan. Hal ini penting dalam proses clustering karena algoritma K-Means bekerja berdasarkan perhitungan jarak antar data dalam ruang multidimensi, sehingga struktur data yang konsisten dan terorganisir akan mempermudah proses pembentukan cluster yang optimal [5].

Selain itu, proses normalisasi menjadi langkah krusial karena perbedaan skala antar variabel dapat menyebabkan bias dalam perhitungan jarak. Variabel dengan nilai yang lebih besar berpotensi mendominasi hasil clustering jika tidak dinormalisasi. Oleh karena itu, penggunaan MinMaxScaler memastikan bahwa setiap variabel memiliki kontribusi yang seimbang dalam proses pengelompokan, sehingga hasil cluster yang terbentuk lebih representatif terhadap karakteristik data secara keseluruhan.

2.4.4 Data Mining

Tahap data mining merupakan proses utama untuk menemukan pola dari data [5]. Metode yang digunakan adalah K-Means Clustering dengan variabel populasi ternak sebagai input utama. Algoritma bekerja dengan menentukan centroid, menghitung jarak, dan memperbarui centroid secara iteratif hingga konvergen [5].

Dalam implementasinya, algoritma K-Means dijalankan dengan beberapa variasi jumlah cluster (K) untuk mengidentifikasi struktur pengelompokan yang paling sesuai dengan karakteristik data. Penentuan nilai K dilakukan



menggunakan metode Elbow dan Silhouette Score sehingga jumlah cluster yang dipilih memiliki dasar yang lebih objektif.

Selain itu, proses clustering dilakukan dengan memperhatikan stabilitas hasil melalui beberapa percobaan untuk mengurangi pengaruh inisialisasi centroid yang bersifat acak. Dalam proses perhitungan jarak, digunakan metode Euclidean Distance yang sesuai untuk data numerik, sehingga mampu merepresentasikan perbedaan antar wilayah secara kuantitatif.

Meskipun K-Means memiliki keunggulan dalam efisiensi dan kemudahan implementasi, algoritma ini memiliki keterbatasan seperti sensitivitas terhadap outlier dan kecenderungan membentuk cluster berbentuk sferis. Oleh karena itu, hasil clustering diinterpretasikan dengan mempertimbangkan keterbatasan tersebut agar tidak menghasilkan kesimpulan yang bias.

2.4.5 Evaluation/Interpretation

Evaluasi dilakukan untuk menilai kualitas hasil clustering dengan tujuan meminimalkan variasi dalam cluster dan memaksimalkan perbedaan antar cluster [9]. Metode yang digunakan adalah Silhouette Score untuk mengukur kualitas pemisahan cluster serta Elbow Method untuk menentukan jumlah cluster optimal.

Nilai Silhouette Score diinterpretasikan untuk mengetahui sejauh mana setiap data berada dalam cluster yang tepat dibandingkan dengan cluster lainnya. Semakin mendekati nilai 1 menunjukkan bahwa data memiliki tingkat kesesuaian yang tinggi dalam cluster-nya, sedangkan nilai yang mendekati 0 mengindikasikan adanya kemungkinan overlap antar cluster. Oleh karena itu, interpretasi hasil tidak hanya berfokus pada nilai numerik semata, tetapi juga mempertimbangkan pola distribusi data yang terbentuk.

Selain itu, hasil evaluasi digunakan tidak hanya untuk menentukan jumlah cluster terbaik, tetapi juga untuk memahami keterbatasan model yang digunakan. Nilai Silhouette Score yang tidak terlalu tinggi menunjukkan bahwa struktur data memiliki kompleksitas yang tidak sepenuhnya dapat dipisahkan secara jelas oleh K-Means. Dengan demikian, hasil clustering dalam penelitian ini diinterpretasikan secara hati-hati dengan mempertimbangkan kemungkinan adanya tumpang tindih karakteristik antar wilayah.

2.4.6 Knowledge/Insight

Tahap knowledge merupakan tahap akhir dalam KDD di mana hasil analisis diterjemahkan menjadi informasi yang dapat dimanfaatkan. Dalam penelitian ini, pendekatan Business Intelligence digunakan sebagai lapisan analitis tambahan yang mengintegrasikan hasil clustering dengan visualisasi data untuk mendukung interpretasi yang lebih komprehensif, bukan sekadar penyajian grafik deskriptif. Hasil analisis digunakan untuk memahami distribusi populasi ternak serta menjadi dasar dalam pengambilan keputusan berbasis data di sektor peternakan.

Selain itu, insight yang dihasilkan memungkinkan identifikasi karakteristik masing-masing cluster, seperti perbedaan antara wilayah dengan kapasitas populasi tinggi dan wilayah dengan potensi pertumbuhan yang lebih besar. Informasi ini dapat digunakan untuk mendukung perumusan strategi yang lebih spesifik, seperti optimalisasi produksi pada wilayah dengan populasi tinggi serta pengembangan pada wilayah dengan pertumbuhan yang lebih potensial. Dengan demikian, hasil analisis tidak hanya bersifat deskriptif, tetapi juga memberikan nilai tambah dalam mendukung pengambilan keputusan yang lebih terarah dan berbasis data [2].

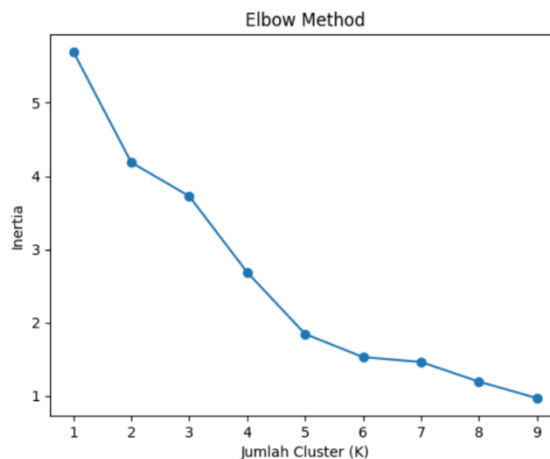
3. HASIL DAN PEMBAHASAN

3.1 Penentuan Jumlah Cluster

Penentuan jumlah cluster merupakan tahap penting dalam proses clustering untuk memastikan hasil pengelompokan yang optimal. Jumlah cluster yang tidak tepat dapat menyebabkan hasil pengelompokan menjadi kurang representatif terhadap struktur data yang sebenarnya [9]. Pada penelitian ini, penentuan jumlah cluster dilakukan menggunakan metode Elbow dan Silhouette Score, yang merupakan dua pendekatan umum dalam evaluasi hasil clustering untuk menentukan jumlah cluster yang optimal.

Berdasarkan grafik Elbow pada Gambar 1, terlihat bahwa penurunan nilai inerti mengalami perlambatan setelah $K = 2$, yang menunjukkan adanya titik siku (elbow point) sebagai indikasi jumlah cluster yang optimal. Hal ini mengindikasikan bahwa penambahan jumlah cluster setelah nilai tersebut tidak memberikan peningkatan yang signifikan terhadap kualitas model [10]. Hasil tersebut kemudian dibandingkan dengan metode Silhouette Score, di mana nilai tertinggi diperoleh pada $K = 2$ sebesar 0,328. Nilai ini menunjukkan bahwa struktur cluster yang terbentuk masih tergolong lemah dan terdapat kemungkinan overlap antar cluster, sehingga pemisahan data belum sepenuhnya optimal.

Meskipun demikian, pemilihan $K = 2$ tetap dipertahankan karena didukung oleh hasil Elbow Method yang menunjukkan titik optimal yang jelas. Oleh karena itu, jumlah cluster dalam penelitian ini ditentukan berdasarkan kombinasi kedua metode, dengan mempertimbangkan bahwa struktur data yang dianalisis memiliki kompleksitas yang tidak sepenuhnya dapat dipisahkan secara tegas oleh algoritma K-Means.



Gambar 1. Grafik Elbow Method untuk Penentuan Jumlah Cluster

3.2 Hasil Cluster

Setelah jumlah cluster optimal ditentukan, tahap selanjutnya adalah penerapan metode K-Means Clustering untuk mengelompokkan wilayah berdasarkan karakteristik populasi ternak. Hasil clustering menunjukkan bahwa data terbagi menjadi dua kelompok utama, yaitu cluster 0 dan cluster 1. Distribusi jumlah wilayah pada masing-masing cluster ditunjukkan pada Tabel 2.

Tabel 2. Anggota Cluster

Cluster	Anggota
0	24
1	7

Berdasarkan hasil pengelompokan, cluster 0 terdiri dari 24 wilayah, sedangkan cluster 1 terdiri dari 7 wilayah. Distribusi ini menunjukkan bahwa sebagian besar wilayah di Provinsi Jawa Tengah berada dalam satu kelompok yang dominan, sementara sebagian kecil wilayah lainnya membentuk kelompok dengan karakteristik yang berbeda. Ketimpangan jumlah anggota antar cluster ini merupakan fenomena yang umum dalam proses clustering, di mana distribusi data yang tidak seimbang dapat mencerminkan adanya dominasi pola tertentu dalam dataset.

Ketidakseimbangan jumlah anggota tersebut mengindikasikan bahwa sebagian besar wilayah memiliki karakteristik populasi ternak yang relatif homogen, sedangkan sebagian kecil wilayah memiliki karakteristik yang lebih spesifik atau berbeda secara signifikan. Dalam konteks analisis clustering, kondisi ini dapat menunjukkan adanya kelompok minoritas yang berperan sebagai pembeda utama dalam struktur data dan berpotensi menjadi fokus analisis lebih lanjut.

Dengan demikian, hasil clustering ini tidak hanya menunjukkan pembagian kelompok, tetapi juga memberikan indikasi awal mengenai adanya segmentasi wilayah berdasarkan karakteristik populasi ternak. Segmentasi ini menjadi dasar untuk analisis lanjutan dalam memahami perbedaan karakteristik antar cluster serta implikasinya terhadap pengelolaan sektor peternakan di masing-masing wilayah.

3.3 Karakteristik Tiap Cluster

Karakteristik masing-masing cluster dianalisis berdasarkan nilai rata-rata variabel populasi ternak, yaitu sapi, kambing, dan ayam, serta didukung oleh variabel tambahan berupa growth sebagai indikator perkembangan populasi. Berdasarkan hasil perhitungan, nilai rata-rata variabel pada masing-masing cluster ditunjukkan pada Tabel 3.

Tabel 3. Rata-rata Karakteristik Setiap Cluster

Cluster	Sapi	Kambing	Ayam	Growth
0	34.550	164.960	10.542.620	0,648
1	160.084	301.640	9.290.849	-0,073

Berdasarkan Tabel 3, cluster 0 merupakan kelompok dengan jumlah wilayah terbanyak dan memiliki rata-rata populasi ternak yang relatif lebih rendah dibandingkan cluster lainnya. Meskipun demikian, cluster ini menunjukkan nilai growth sebesar 0,648 yang mengindikasikan adanya tren pertumbuhan populasi yang positif. Kondisi ini menunjukkan bahwa wilayah dalam cluster 0 masih berada dalam tahap pengembangan, di mana peningkatan populasi ternak masih dapat terjadi seiring dengan ketersediaan sumber daya dan peluang ekspansi yang relatif terbuka.

Sebaliknya, cluster 1 terdiri dari jumlah wilayah yang lebih sedikit, namun memiliki rata-rata populasi ternak yang jauh lebih tinggi, khususnya pada ternak sapi dan kambing. Namun, nilai growth sebesar -0,073 menunjukkan adanya kecenderungan stagnasi atau penurunan dalam perkembangan populasi ternak. Kondisi ini mengindikasikan bahwa



wilayah dalam cluster 1 telah mencapai tingkat pemanfaatan sumber daya yang tinggi, sehingga ruang untuk pertumbuhan menjadi lebih terbatas.

Secara lebih mendalam, perbedaan ini dapat dipengaruhi oleh beberapa faktor struktural, seperti keterbatasan lahan peternakan, meningkatnya biaya produksi, serta tekanan terhadap sumber daya pakan yang umumnya terjadi pada wilayah dengan populasi ternak tinggi. Selain itu, wilayah dengan populasi besar cenderung menghadapi tantangan dalam menjaga efisiensi produksi, sehingga peningkatan jumlah ternak tidak selalu diikuti dengan pertumbuhan yang berkelanjutan [2].

Jika dibandingkan, cluster 1 memiliki keunggulan pada jumlah populasi ternak, sedangkan cluster 0 unggul pada aspek pertumbuhan. Hal ini menunjukkan bahwa terdapat perbedaan karakteristik yang cukup signifikan antara kedua cluster, di mana cluster 0 dapat dikategorikan sebagai wilayah dengan potensi pengembangan (growth-oriented), sedangkan cluster 1 merupakan wilayah dengan kapasitas produksi tinggi (production-oriented) namun dengan dinamika pertumbuhan yang relatif terbatas.

Dengan demikian, perbedaan karakteristik ini tidak hanya mencerminkan kondisi saat ini, tetapi juga memberikan gambaran mengenai tahapan perkembangan sektor peternakan pada masing-masing wilayah. Informasi ini dapat digunakan sebagai dasar dalam analisis lanjutan dan perumusan strategi pengembangan yang lebih tepat sasaran, terutama dalam menentukan prioritas kebijakan antara peningkatan kapasitas produksi dan pengembangan potensi pertumbuhan.

3.4 Analisis Perbandingan Cluster

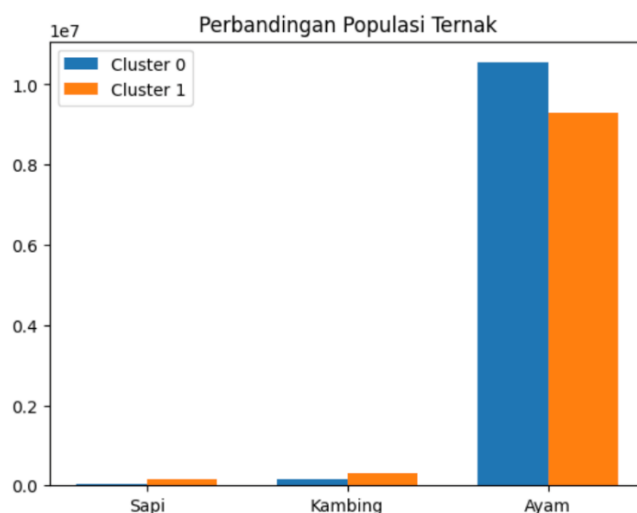
Analisis perbandingan cluster dilakukan untuk mengidentifikasi perbedaan karakteristik antar kelompok wilayah berdasarkan variabel populasi ternak. Berdasarkan hasil perbandingan nilai rata-rata, terlihat bahwa cluster 1 memiliki nilai yang lebih tinggi pada variabel populasi ternak sapi dan kambing dibandingkan dengan cluster 0. Hal ini menunjukkan bahwa wilayah dalam cluster 1 memiliki kapasitas populasi ternak yang lebih besar, khususnya pada kedua jenis ternak tersebut.

Sementara itu, pada variabel ayam, perbedaan antara kedua cluster tidak terlalu signifikan, meskipun cluster 0 cenderung memiliki nilai yang sedikit lebih tinggi. Selain itu, variabel growth menunjukkan perbedaan yang cukup mencolok, di mana cluster 0 memiliki tingkat pertumbuhan yang lebih tinggi dibandingkan cluster 1.

Jika dianalisis lebih lanjut, perbedaan ini menunjukkan bahwa hubungan antara jumlah populasi ternak dan tingkat pertumbuhan tidak bersifat linear. Wilayah dengan populasi ternak yang tinggi tidak selalu menunjukkan pertumbuhan yang tinggi, bahkan cenderung mengalami stagnasi atau penurunan. Kondisi ini dapat mengindikasikan adanya fenomena diminishing return, di mana peningkatan jumlah populasi tidak lagi diikuti oleh peningkatan pertumbuhan yang sebanding.

Selain itu, perbedaan pola ini juga dapat dipengaruhi oleh faktor efisiensi produksi dan kapasitas sumber daya. Wilayah dengan populasi tinggi kemungkinan menghadapi tekanan yang lebih besar terhadap ketersediaan pakan, lahan, serta biaya operasional, sehingga pertumbuhan menjadi lebih sulit dipertahankan. Sebaliknya, wilayah dengan populasi yang lebih rendah memiliki ruang pengembangan yang lebih besar, sehingga mampu menunjukkan pertumbuhan yang lebih tinggi.

Dengan demikian, perbandingan antar cluster tidak hanya menunjukkan perbedaan jumlah populasi ternak, tetapi juga mengungkap adanya perbedaan pola perkembangan dan efisiensi antar wilayah. Hal ini memberikan indikasi bahwa strategi pengembangan sektor peternakan tidak dapat disamaratakan, melainkan perlu disesuaikan dengan karakteristik masing-masing wilayah.



Gambar 2. Perbandingan Nilai Rata-rata Variabel pada Setiap Cluster

3.5 Analisis Business Intelligence

Selain analisis berdasarkan variabel utama dalam proses clustering, penelitian ini juga memanfaatkan data pendukung berupa produksi daging dan telur untuk memperkaya analisis melalui pendekatan Business Intelligence. Pendekatan ini

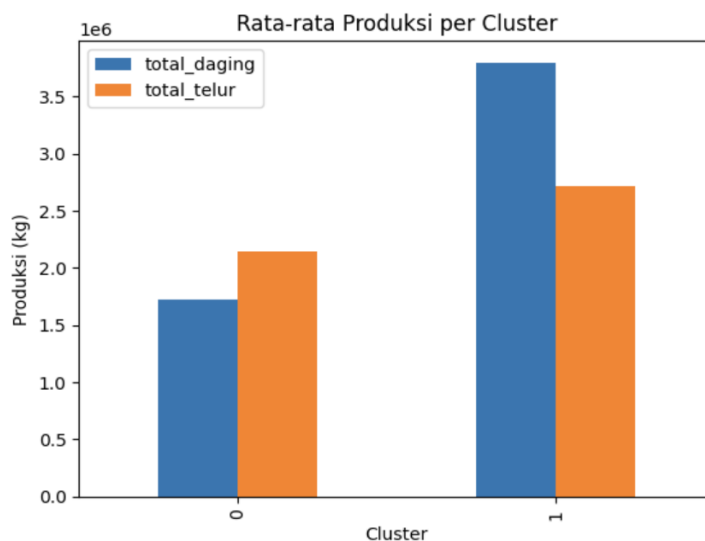


bertujuan untuk memberikan pemahaman yang lebih komprehensif mengenai hubungan antara populasi ternak, tingkat pertumbuhan, dan kapasitas produksi pada masing-masing cluster [6]. Rata-rata produksi daging dan telur pada masing-masing cluster ditunjukkan pada Tabel 4.

Table 4. Rata-rata Produksi Daging dan Telur pada Setiap Cluster

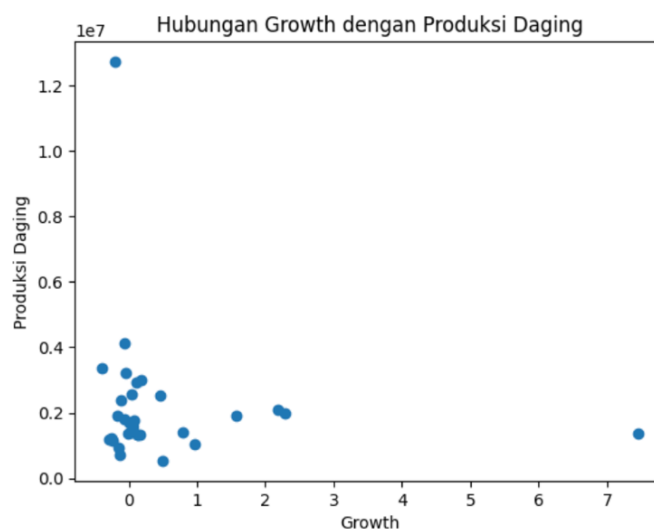
Cluster	Produksi Daging	Produksi Telur
0	1.724.596	2.146.987
1	3.796.785	2.716.253

Berdasarkan Tabel 4, cluster 1 memiliki tingkat produksi yang lebih tinggi dibandingkan cluster 0, baik pada produksi daging maupun telur. Kondisi ini sejalan dengan hasil sebelumnya yang menunjukkan bahwa cluster 1 memiliki populasi ternak yang lebih besar, sehingga mengindikasikan adanya hubungan positif antara jumlah populasi ternak dan kapasitas produksi. Visualisasi perbandingan produksi antar cluster ditunjukkan pada Gambar 3.



Gambar 3. Perbandingan Produksi Daging dan Telur Antar Cluster

Namun demikian, analisis tidak berhenti pada perbandingan nilai rata-rata. Untuk memahami hubungan antar variabel secara lebih mendalam, dilakukan visualisasi hubungan antara growth dan produksi daging menggunakan scatter plot, sebagaimana ditunjukkan pada Gambar 4. Pendekatan ini umum digunakan dalam analisis data untuk mengidentifikasi pola hubungan dan distribusi data secara visual.



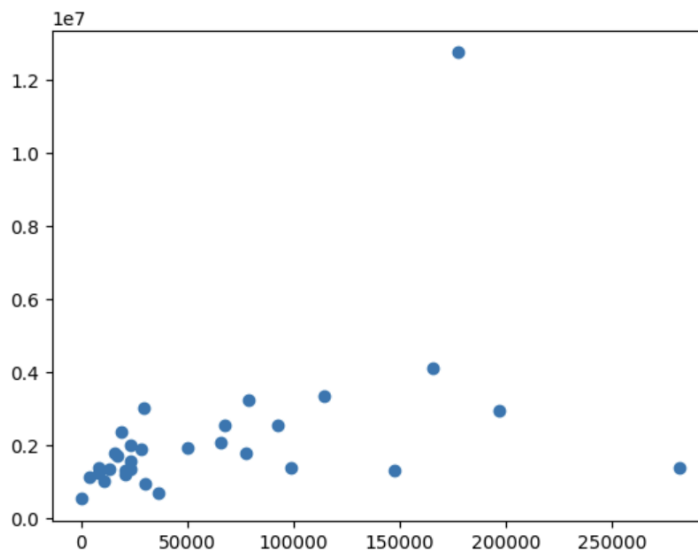
Gambar 4. Hubungan Growth dengan Produksi Daging

Berdasarkan Gambar 4, hubungan antara variabel growth dan produksi daging tidak menunjukkan pola linear yang kuat. Hal ini terlihat dari penyebaran titik data yang tidak membentuk tren tertentu. Beberapa wilayah dengan nilai growth tinggi tidak diikuti oleh produksi yang tinggi, sementara wilayah dengan produksi tinggi justru memiliki nilai pertumbuhan yang rendah atau negatif.



Kondisi ini menunjukkan bahwa peningkatan pertumbuhan populasi ternak tidak secara langsung berbanding lurus dengan peningkatan produksi. Dengan kata lain, hubungan antara pertumbuhan dan produksi bersifat kompleks dan dipengaruhi oleh faktor lain seperti efisiensi produksi, teknologi, serta ketersediaan sumber daya.

Selain itu, terdapat beberapa titik data yang menyimpang (outlier), yang menunjukkan adanya wilayah dengan karakteristik produksi yang sangat berbeda dibandingkan wilayah lainnya. Keberadaan outlier ini mengindikasikan bahwa struktur data tidak homogen dan memerlukan pendekatan analisis yang lebih mendalam dalam interpretasi hasil clustering. Untuk memperkuat analisis, dilakukan visualisasi hubungan antara populasi ternak dan produksi daging, sebagaimana ditunjukkan pada Gambar 5.



Gambar 5. Hubungan Populasi Sapi dengan Produksi Daging

Berbeda dengan hubungan antara growth dan produksi, hubungan antara populasi ternak dan produksi menunjukkan kecenderungan yang lebih konsisten. Wilayah dengan populasi ternak yang lebih tinggi cenderung memiliki tingkat produksi yang lebih tinggi, meskipun tidak sepenuhnya linear akibat adanya variasi efisiensi antar wilayah. Temuan ini menunjukkan bahwa kapasitas produksi lebih dipengaruhi oleh jumlah populasi ternak dibandingkan dengan laju pertumbuhannya. Dengan demikian, peningkatan produksi dalam jangka pendek lebih bergantung pada skala populasi, sedangkan growth lebih mencerminkan potensi perkembangan jangka panjang.

Secara keseluruhan, hasil analisis menunjukkan adanya perbedaan antara current performance dan future potential. Cluster 1 dapat dikategorikan sebagai wilayah dengan kapasitas produksi tinggi (current performance) namun memiliki risiko stagnasi pertumbuhan, sedangkan cluster 0 menunjukkan potensi pertumbuhan yang lebih baik (future potential) meskipun kapasitas produksinya masih relatif lebih rendah [6]. Dengan demikian, pendekatan Business Intelligence dalam penelitian ini tidak hanya berfungsi sebagai alat visualisasi, tetapi juga mampu mengungkap hubungan antar variabel serta menghasilkan insight yang lebih mendalam untuk mendukung pengambilan keputusan berbasis data di sektor peternakan.

3.6 Interpretasi & Implikasi

Berdasarkan hasil analisis clustering dan pendekatan Business Intelligence yang telah dilakukan, diperoleh beberapa implikasi penting yang dapat digunakan sebagai dasar dalam pengambilan kebijakan di sektor peternakan. Hasil penelitian menunjukkan bahwa wilayah di Provinsi Jawa Tengah dapat dikelompokkan menjadi dua kategori utama, yaitu wilayah dengan potensi pertumbuhan tinggi dan wilayah dengan kapasitas produksi yang besar. Segmentasi ini memberikan perspektif yang lebih mendalam dibandingkan analisis deskriptif konvensional karena mampu membedakan antara kondisi saat ini dan potensi perkembangan di masa mendatang.

Cluster 0 merepresentasikan wilayah dengan tingkat pertumbuhan populasi ternak yang relatif tinggi, meskipun jumlah populasinya masih lebih rendah. Kondisi ini menunjukkan bahwa wilayah dalam cluster ini memiliki potensi pengembangan yang cukup besar, terutama dalam konteks ekspansi populasi ternak. Sebaliknya, cluster 1 menunjukkan wilayah dengan jumlah populasi ternak yang besar, namun memiliki kecenderungan pertumbuhan yang stagnan atau menurun, yang mengindikasikan adanya keterbatasan dalam mempertahankan laju pertumbuhan pada tingkat populasi yang tinggi.

Implikasi dari temuan tersebut menunjukkan bahwa pendekatan kebijakan tidak dapat diseragamkan. Pada cluster 0, strategi yang lebih tepat adalah mendorong ekspansi melalui peningkatan akses terhadap sumber daya produksi seperti pakan, lahan, dan teknologi, sehingga potensi pertumbuhan dapat dimaksimalkan. Sementara itu, pada cluster 1, fokus kebijakan sebaiknya diarahkan pada optimalisasi sistem yang sudah ada, seperti peningkatan efisiensi produksi, pengelolaan sumber daya yang lebih efektif, serta inovasi teknologi untuk mengatasi potensi stagnasi [4].



Selain itu, hasil analisis menunjukkan bahwa tingginya produksi tidak selalu diikuti oleh pertumbuhan yang baik. Hal ini mengindikasikan bahwa kebijakan yang hanya berorientasi pada peningkatan output produksi tanpa mempertimbangkan dinamika pertumbuhan berpotensi tidak berkelanjutan dalam jangka panjang. Oleh karena itu, diperlukan pendekatan yang lebih seimbang antara peningkatan kapasitas produksi dan pengelolaan pertumbuhan populasi ternak.

Dengan demikian, penelitian ini menunjukkan bahwa segmentasi berbasis clustering dapat digunakan sebagai dasar dalam penyusunan kebijakan yang lebih terarah dan berbasis data. Pendekatan ini memungkinkan pemangku kepentingan untuk merancang strategi yang kontekstual sesuai dengan karakteristik masing-masing wilayah, sehingga dapat meningkatkan efektivitas pengelolaan sektor peternakan serta mendukung pembangunan yang berkelanjutan.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa pendekatan K-Means Clustering yang dikombinasikan dengan analisis Business Intelligence mampu mengidentifikasi pola segmentasi wilayah dalam sektor peternakan di Provinsi Jawa Tengah yang tidak terlihat melalui analisis deskriptif konvensional. Temuan utama penelitian ini tidak hanya terletak pada pembagian wilayah ke dalam dua cluster, tetapi pada adanya perbedaan mendasar antara kapasitas produksi saat ini (current performance) dan potensi pertumbuhan di masa mendatang (future potential). Hasil analisis mengindikasikan bahwa wilayah dengan populasi ternak yang tinggi cenderung memiliki kapasitas produksi yang besar, namun tidak selalu diikuti oleh pertumbuhan yang berkelanjutan. Sebaliknya, wilayah dengan populasi yang lebih rendah justru menunjukkan laju pertumbuhan yang lebih tinggi. Temuan ini menegaskan bahwa hubungan antara populasi, produksi, dan pertumbuhan bersifat tidak linear serta dipengaruhi oleh faktor efisiensi dan keterbatasan sumber daya. Kontribusi utama penelitian ini terletak pada penggunaan pendekatan clustering untuk membedakan karakteristik wilayah berdasarkan dinamika pertumbuhan dan kapasitas produksi secara simultan, bukan hanya berdasarkan besaran populasi. Pendekatan ini memberikan nilai tambah dibandingkan analisis konvensional yang umumnya bersifat agregatif dan kurang mampu menangkap pola segmentasi antar wilayah. Secara praktis, hasil penelitian ini mengindikasikan bahwa kebijakan pengembangan sektor peternakan tidak dapat diseragamkan. Wilayah dengan kapasitas produksi tinggi memerlukan strategi optimalisasi dan efisiensi untuk mengatasi potensi stagnasi, sedangkan wilayah dengan pertumbuhan tinggi memerlukan dukungan ekspansi untuk meningkatkan skala produksi. Dengan demikian, pendekatan berbasis data seperti ini dapat membantu dalam penyusunan kebijakan yang lebih terarah, kontekstual, dan berkelanjutan. Namun demikian, penelitian ini memiliki keterbatasan, terutama pada nilai validitas cluster yang relatif rendah, sebagaimana ditunjukkan oleh nilai Silhouette Score yang belum optimal. Hal ini menunjukkan bahwa pemisahan antar kelompok belum sepenuhnya kuat dan masih terdapat kemungkinan tumpang tindih karakteristik antar wilayah. Selain itu, penelitian ini belum memasukkan variabel eksternal seperti faktor ekonomi, infrastruktur, dan kondisi geografis yang berpotensi memengaruhi pola yang terbentuk. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan metode clustering yang lebih robust, seperti Hierarchical Clustering atau DBSCAN, serta menambahkan variabel pendukung yang lebih komprehensif agar hasil segmentasi menjadi lebih akurat dan interpretatif.

REFERENCES

- [1] R. Widyastuti, "Evaluasi Manajemen Kesehatan dan Kesejahteraan Hewan pada Peternakan Domba, Sapi Perah, Sapi Potong, dan Burung Puyuh di Wilayah Jatinangor, Jawa Barat, Indonesia," *Farmers J. community Serv.*, vol. 7, no. 1, pp. 114–118, 2026. <https://doi.org/10.24198/fjcs.v7i1.69115>
- [2] R. Ardhani, S. Marshelle, A. Fitrianto, and L. M. R. Dwi, "Perbandingan K-Medoids Dan Clara (Clustering Large Application)," *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 5, no. 3, pp. 1744–1763, 2024. <https://doi.org/10.46306/lb.v5i3.764>
- [3] P. Bisecting and K. D. A. N. Agglomerative, "Hierarchical Clustering Untuk Pengelompokan Daerah," *VARIANSI: Journal of Statistics and Its Application on Teaching and Research.*, vol. 8, no. 2, pp. 175–182, 2025. <https://doi.org/10.35580/variansiunm38>
- [4] L. Sumaryanti and D. Muchlis, "Mathematical Modelling of Engineering Problems Analysis of Regional Potential in Merauke Regency Based on Superior Livestock Population Using a Hybrid Algorithm," *Mathematical Modelling of Engineering Problems.*, vol. 11, no. 3, pp. 754–764, 2024. <https://doi.org/10.18280/mmep.110320>
- [5] Fajriana, "Analisis Algoritma K-Medoids pada Sistem Klusterisasi Produksi Perikanan Tangkap," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 7, no. 2, pp. 263–269, 2021. <https://doi.org/10.26418/jp.v7i2.47795>
- [6] B. Nikum and H. Yuliansyah, "Evaluation of K-Means Clustering Using Silhouette Score Method on Customer Segmentation," *jurnal.fikom.*, vol. 16, no. 3, pp. 330–342, 2025. <https://doi.org/10.33096/ilkom.v16i3.2325.330-342>
- [7] I. A. A. Iis Setiawan Mangku Negara¹, Purwono², "Analisa Cluster Data Transaksi Penjualan Minimarket Selama Pandemi Covid-19 dengan Algoritma K-means 1,3," *Journal of Information.*, vol. 3, no. 28, pp. 153–160, 2022. <https://doi.org/10.31328/jointecs.v6i3.2693>
- [8] D. Lasandi, R. Astuti, K. Anam, A. P. Putra, and B. Nurhakim, "Penerapan Algoritma K-Means Clustering Untuk Analisis Pola Populasi Ternak Di Kota," *Jurnal Informatika dan Teknik Elektro Terapan.*, vol. 14, no. 2. <https://doi.org/10.23960/jitet.v14i2.9248>
- [9] R. H. Bhahari, "Clustering Analysis of Socio-Economic Districts / Cities In East Java Province Using PCA And Hierarchical Clustering Methods," *Sinkron: jurnal dan penelitian teknik informatika.*, vol. 8, no. 4, pp. 2242–2251, 2024. <https://doi.org/10.33395/sinkron.v8i4.14078>
- [10] T. S. Irma Khoirunisa, Rini Astuti, "Penerapan Algoritma K-Medoids Dalam Menentukan Cluster Kabupaten Dan Kota



- Berdasarkan Populasi Peternakan Di Provinsi Jawa Barat,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3753–3758, 2023. <https://doi.org/10.36040/jati.v7i6.8202>
- [11] M. T. J. Stefanus Arman, Kris Suryowati, “Pengelompokan Provinsi Di Indonesia Berdasarkan Indikator Peternakan Dengan Menggunakan Algoritma K-Means Dan K-Medoids,” *Jurnal Statistika Industri dan Komputasi.*, vol. 09, no. 02, pp. 51–59, 2024. <https://doi.org/10.34151/statistika.v9i2.4858>
- [12] N. H. Cahyaningrum, N. Mas, and R. Yotenka, “Perbandingan Algoritma K-Means dan K-Medoids dalam Pengelompokan Peternakan Unggas di Daerah Istimewa Yogyakarta Tahun 2023,” *Emerging Statistics and Data Science Journal.*, vol. 3, no. 2, pp. 630–639, 2025. <https://doi.org/10.20885/esds.vol3.iss.2.art16>
- [13] R. Rafsanjani, S. M. Kristin, and A. Firman, “Peran Subsektor Peternakan dalam Perekonomian Jawa Tengah The Role of Livestock Subsector in Central Java Regional Economy,” *Jurnal Peikiran Masyarakat Ilmiah Berwawasan Agribisnis.*, vol. 11, pp. 789–798, 2025. <https://doi.org/10.25157/ma.v11i1.16477>
- [14] Y. A. Wijaya *et al.*, “K-Means Di Sekolah Menengah Kejuruan Wahidin Kota Cirebon,” *Emerging Statistics and Data Science Journal.*, vol. 6, no. 2, pp. 552–559, 2022. <https://doi.org/10.36040/jati.v6i2.5236>
- [15] D. K. Zega, “Analisis Kelayakan Pengembangan Usaha Ternak Sapi Potong (Beef Cattle): Suatu Kajian Literatur,” *Business and Entrepreneurship Journal (BEJ).*, vol. 7, no. 1, pp. 1–9, 2026. <https://doi.org/10.57084/bej.v7i1.2204>
- [16] R. F. Miwan Kurniawan Hidayat, “Implementasi k-means dan k-medoids dalam pengelompokan wilayah potensial produksi daging ayam implementation of k-means and k-medoids in grouping potential areas of chicken meat production,” *Journal of Agroindustrial Technology/Jurnal Teknologi Industri.*, vol. 32, no. 158, pp. 239–247, 2022. <https://doi.org/10.24961/j.tek.ind.pert.2022.32.3.239>
- [17] N. O. Syamsiah, I. Purwandani, M. Rosmiati, and S. Nurwahyuni, “Triangulation Approach Using K-Means , Hierarchical Clustering , and DBSCAN for Beef Production Analysis,” *IJNMT (International Journal of New Media Technology).*, vol. 12, no. 2, 2026. <https://doi.org/10.31937/ijnmt.v12i2.4481>
- [18] L. Trapanese, G. Bifulco, M. Santinello, N. Pasquino, G. Campanile, and A. Salzano, “Comparison of K-Means and Hierarchical Clustering Methods for Buffalo Milk Production Data,” *Animals* vol. 12, no. 2, pp. 1–14, 2025. <https://doi.org/10.3390/ani15223246>
- [19] F. Mukaromah and I. K. D. Nuryana, “Clustering of Goat Buyers in West Java with K-Means Algorithm,” *Journal of Emerging Information Systems and Business Intelligence.*, vol. 6, no. 4, pp. 485–498, 2025. <https://doi.org/10.26740/jeisbi.v6i4.72022>
- [20] H. Rhomadhona, W. Kusri, W. Aprianti, and J. Permadi, “Implementation of K - Means Clustering for Social Assistance Recipients with Silhouette Score Evaluation,” *Brilliance: Research of Artificial Intelligence.*, vol. 5, no. 1, pp. 136–143, 2025. <https://doi.org/10.47709/brilliance.vxix.xxxx>