HEXATECH
Jurnal Ilmiah Teknik
Vol 4 No 2 Agustus 2025
ISSN: 2828-8696 (Print) ISSN: 2828-8548 (Electronic)
Open Access: https://jurnal.arkainstitute.co.id/index.php/hexatech/index

Penerapan algoritma *K-Means Clustering* dalam identifikasi pola penggunaan barang di gudang KPU Kabupaten Banyuwangi

Candra Edmond Safra[1], Ifani Hariyanti[2]
[1,2]Universitas Adhirajasa Reswara Sanjaya
email: [1]candrasafra0@gmail.com, [2]ifani@ars.ac.id

Info Artikel :
Diterima :
10 Juli 2025
Disetujui :
15 Agustus 2025
Dipublikasikan :
30 Agustus 2025

ABSTRAK

Pengelolaan logistik dalam penyelenggaraan pemilihan umum memiliki peran vital dalam menjamin kelancaran dan efisiensi proses demokrasi. Namun, pengelolaan barang logistik di gudang Komisi Pemilihan Umum (KPU) Kabupaten Banyuwangi masih dilakukan tanpa analisis data historis yang terstruktur, sehingga berpotensi menimbulkan ketidakefisienan dalam perencanaan dan distribusi barang. Penelitian ini bertujuan untuk mengidentifikasi pola penggunaan barang logistik Pilkada menggunakan algoritma K-Means Clustering pada data historis penggunaan barang di gudang KPU Kabupaten Banyuwangi. Data penelitian terdiri dari 206 record dengan 37 atribut yang mencakup berbagai jenis logistik seperti kotak suara, bilik suara, dan formulir. Analisis dilakukan menggunakan perangkat lunak RapidMiner Studio 9.10 dengan tahapan data preprocessing meliputi pembersihan, normalisasi, serta pembagian data training dan testing. Hasil analisis menunjukkan bahwa data dapat dikelompokkan ke dalam tiga cluster utama yang merepresentasikan tingkat kebutuhan tinggi, sedang, dan rendah. Dapat disimpulkan bahwa evaluasi model menggunakan Davies-Bouldin Index menunjukkan bahwa nilai K menghasilkan pemisahan cluster terbaik dengan struktur data yang kompak. Temuan ini diharapkan dapat menjadi dasar dalam pengambilan keputusan strategis untuk meningkatkan efisiensi pengelolaan gudang logistik Pilkada di KPU Kabupaten Banyuwangi.

Kata kunci*: K-Means Clustering*, KPU Kabupaten Banyuwangi, Manajemen Logistik, Pemilu, Pola Penggunaan Barang

ABSTRACT

Logistics management in the implementation of general elections plays a vital role in ensuring the smooth and efficient democratic process. However, the management of logistics goods in the warehouse of the General Election Commission (KPU) of Banyuwangi Regency is still carried out without structured historical data analysis, thus potentially causing inefficiencies in the planning and distribution of goods. This study aims to identify the pattern of use of logistics goods for the regional elections using the K-Means Clustering algorithm on historical data of goods usage in the warehouse of the KPU of Banyuwangi Regency. The research data consists of 206 records with 37 attributes covering various types of logistics such as ballot boxes, voting booths, and forms. The analysis was carried out using RapidMiner Studio 9.10 software with data preprocessing stages including cleaning, normalization, and division of training and testing data. The results of the analysis show that the data can be grouped into three main clusters representing high, medium, and low levels of need. It can be concluded that the model evaluation using the Davies-Bouldin Index shows that the K value produces the best cluster separation with a compact data structure. These findings are expected to be the basis for strategic decision making to improve the efficiency of the management of logistics warehouses for the regional elections at the KPU of Banyuwangi Regency.

Keywords : K-Means Clustering, Banyuwangi Regency KPU, Logistics Management, Election, Goods Usage Patterns

PENDAHULUAN

Perkembangan teknologi, dinamika sosial, dan kondisi global saat ini mendorong setiap sektor untuk beradaptasi secara cepat dan strategis. Berbagai tantangan kompleks, seperti ketidakstabilan geopolitik, perubahan iklim, dan disrupsi akibat pandemi COVID-19, telah memperlihatkan betapa rentannya sistem yang tidak adaptif. Di sisi lain, tantangan ini membuka peluang signifikan untuk transformasi, terutama melalui pemanfaatan teknologi guna meningkatkan efektivitas dan efisiensi proses pengambilan keputusan (Lestari, 2019). Selain itu, Isu keberlanjutan juga semakin mendorong penerapan green logistics, di mana integrasi prinsip lingkungan dalam manajemen logistik mampu mengurangi emisi karbon hingga 20% tanpa mengorbankan efisiensi operasional (Pricilia et al., 2024).

Dalam konteks ini, manajemen logistik menjadi salah satu titik krusial yang tidak dapat diabaikan. Logistik merupakan proses perencanaan, pelaksanaan, dan pengawasan terhadap pengadaan, penyimpanan, dan distribusi barang. Dalam instansi pemerintahan, pengelolaan logistik yang optimal tidak hanya mendukung kelancaran operasional, tetapi juga berfungsi sebagai indikator penting dalam menentukan tata kelola yang tepat. Di era digital dan administrasi modern, efisiensi pengelolaan barang menjadi faktor kunci untuk memastikan penggunaan sumber daya yang tepat sasaran. Salah satu tantangan utamanya adalah mengidentifikasi pola penggunaan barang secara akurat untuk meminimalkan pemborosan, mengoptimalkan stok, dan meningkatkan efektivitas distribusi. Tantangan ini menjadi semakin kompleks ketika dihadapkan pada dinamika perubahan kebijakan, keterbatasan infrastruktur, dan tuntutan transparansi publik (Purbasari et al., 2023).

Satu satu teknik yang dapat digunakan untuk mengatasi tantangan dalam pengelolaan logistik adalah data mining, khususnya metode Clustering. Salah satu metode yang paling umum digunakan adalah K-Means Clustering, yang mengelompokkan data menurut karakteristik untuk mengidentifikasi pola yang dimaksud. Menurut penelitian (Putra et al., 2022), teknik clustering sangat efektif dalam mengidentifikasi pola dalam data yang kompleks dan beragam. Pengelompokan K- Means memungkinkan data tentang penggunaan barang dikelompokkan menurut frekuensi, waktu penggunaan, atau atribut relevan lainnya.

Studi sebelumnya, seperti penelitian Neva (2023) mengenai k-means, k-medoids, dan x-means untuk menilai kinerja kerja karyawan. Tujuan dari penelitian ini adalah untuk membandingkan berbagai metode clustering untuk menemukan metode yang paling efektif untuk menganalisis cluster dengan mempertimbangkan kinerja kerja karyawan. Menurut hasil penelitian, metode k-means memiliki Indeks Davies Bouldin (DBI) yang rendah. (-0,377), sehingga lebih unggul dalam pengelompokan data. Penelitian serupa oleh Hartama & Sapriyaldi (2023) Tujuan penelitian ini adalah untuk mengumpulkan informasi tentang kelompok data kecelakaan lalu lintas yang didasarkan pada waktu kejadian. Hasil validasi dengan matrix davies bouldin index menunjukkan bahwa 4 cluster memiliki nilai yang cukup untuk mengelompokkan data dengan baik. Hasil evaluasi menunjukkan bahwa masing-masing dari 4 cluster yang dibentuk memiliki nilai sebesar 0,134.

Selanjutnya penelitian yang dilakukan oleh Sugianto et al., (2020) Dalam penelitian ini, algoritma k-medoids digunakan dalam penelitian ini untuk mengelompokkan data penyakit pasien. Hasil penelitian menunjukkan bahwa cluster dengan nilai silhouette coefficient sebesar 0,409373 adalah yang terbaik. Selanjutnya penelitian yang dilakukan oleh Pujiono et al., (2024) dengan judul implementasi data mining untuk menentukan pola penjualan produk menggunakan algoritma k-means clustering. Penelitian ini bertujuan untuk mengidentifikasi strategi dan teknik yang dapat digunakan untuk mengelola stok barang dengan lebih efisien, mengidentifikasi barang yang diminati, dan menghindari stok yang berlebihan. Hasil penelitian menunjukkan bahwa cluster 2 memiliki nilai terbaik dengan nilai DBI sebesar -0,310.

Melihat keberhasilan penerapan metode clustering dalam berbagai studi sebelumnya, maka diperlukan penerapan serupa pada instansi pemerintah yang memiliki sistem logistik kompleks dan dinamis. Salah satu instansi yang memiliki sistem logistik kompleks dan dinamis adalah Komisi Pemilihan Umum (KPU). Sebagai lembaga negara yang bertanggung jawab atas penyelenggaraan pemilu dan pilkada, KPU harus mengelola logistik dalam jumlah besar, menjamin akurasi, ketepatan waktu, serta efisiensi dalam pendistribusiannya ke seluruh wilayah Indonesia (Biroroh, 2021). KPU tidak hanya mengelola barang dalam jumlah besar, tetapi juga dituntut untuk menjamin akurasi, ketepatan waktu, serta efisiensi dalam pendistribusiannya. Kompleksitas kebutuhan logistik yang tinggi dan dinamika pelaksanaan pemilu menjadi alasan kuat mengapa KPU relevan untuk dijadikan objek penelitian, terutama dalam konteks penerapan teknologi analitik untuk meningkatkan proses pengambilan keputusan (Dewi et al., 2022).

Komisi Pemilihan Umum (KPU) adalah salah satu lembaga negara yang paling penting dalam mengelola logistik skala besar, terutama terkait dengan pemilu dan Pilkada. Komisi Pemilihan Umum (KPU) juga bertanggung jawab atas pengadaan, penyimpanan, dan distribusi berbagai jenis logistik pemilu ke seluruh wilayah Indonesia, termasuk daerah terpencil yang menghadapi tantangan geografis dan infrastruktur (Sudrajat et al., 2024). Data Pilkada Serentak tahun 2020 mencatat bahwa lebih dari 270 daerah terlibat dalam proses ini, yang mencakup distribusi alat peraga kampanye, formulir, kotak suara, dan logistik lainnya (Khalyubi et al., 2020). Namun, menurut penelitian Badan Pengawas Pemilu (Pribadi et al., 2022), sekitar 15% kendala Pilkada berasal dari permasalahan distribusi dan pengelolaan

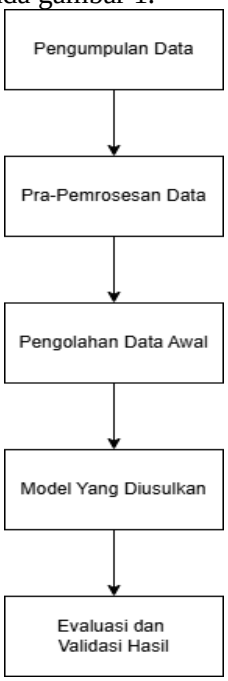
logistik. Di tingkat daerah, seperti yang dilaporkan Komisi Pemilihan Umum (KPU) Provinsi Jawa Timur (Pamungkas et al., 2023), sekitar 20% logistik di gudang Komisi Pemilihan Umum (KPU) Kab/Kota tidak terpakai, sementara 20% lainnya mengalami kekurangan saat dibutuhkan. Hal ini menunjukkan adanya ketidaktepatan dalam perencanaan dan manajemen logistik.

Urgensi dari penelitian ini terletak pada perlunya mengoptimalkan pengelolaan logistik di gudang Komisi Pemilihan Umum (KPU) Kabupaten Banyuwangi dengan pendekatan berbasis data. Penggunaan data historis untuk mengidentifikasi pola penggunaan barang secara sistematis diharapkan mampu memberikan landasan dalam pengambilan keputusan logistik yang lebih tepat. Dengan demikian, proses pengadaan, penyimpanan, dan distribusi dapat dilakukan secara lebih efisien dan akurat. Hal ini sejalan dengan amanat Peraturan Komisi Pemilihan Umum (PKPU) Nomor 15 Tahun 2023 menekankan pentingnya efisiensi, efektivitas, dan akuntabilitas dalam setiap tahapan pemilihan (Titania & Nawangsari, 2025). Penelitian ini diharapkan dapat membantu dalam pengembangan penerapan algoritma Clustering K-Means di institusi publik. Secara praktis, penelitian ini dapat memberikan saran berbasis data untuk meningkatkan perencanaan logistik, mengurangi biaya penyimpanan, dan mengoptimalkan distribusi barang di KPU Kabupaten Banyuwangi.

Berdasarkan pemaparan tersebut, penelitian ini bertujuan mengidentifikasi pola penggunaan barang di gudang KPU Kabupaten Banyuwangi selama penyelenggaraan Pilkada Tahun 2024 dengan menganalisis data historis penggunaan barang. Metode clustering akan digunakan untuk mengelompokkan data historis penggunaan barang guna menghasilkan pola yang dapat dijadikan dasar dalam pengambilan keputusan logistik. Serta memberikan rekomendasi untuk meningkatkan efisiensi manajemen gudang KPU Kabupaten Banyuwangi berdasarkan hasil analisis clustering, termasuk perbaikan dalam perencanaan pengadaan, penyimpanan dan distribusi barang. Dengan demikian, diharapkan hasil penelitian ini dapat meningkatkan efisiensi manajemen gudang, serta mendukung kelancaran penyelenggaraan Pilkada yang transparan dan akuntabel.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif berbasis data *mining*, dengan fokus pada metode *clustering* untuk menemukan pola dalam data historis penggunaan barang di gudang KPU Kabupaten Banyuwangi. Jenis penelitian yang digunakan adalah *clustering* karena menggunakan algoritma *K-Means* untuk mengelompokkan data berdasarkan fitur tertentu (Nur et al., 2024). Maka dalam penelitian ini dilakukan tahapan pada gambar 1.



Gambar 1. Desain Penelitian
Sumber: (Fajriansyah, 2021)

Gambar 1 menggambarkan alur metodologi penelitian yang digunakan untuk menganalisis pola penggunaan barang logistik di Gudang Komisi Pemilihan Umum (KPU) Kabupaten Banyuwangi dengan menerapkan algoritma K-Means Clustering. Tahap pertama adalah pengumpulan data, di mana penelitian ini menggunakan pendekatan kuantitatif berbasis data mining dengan fokus pada metode clustering untuk menemukan pola tersembunyi dalam data historis penggunaan barang logistik. Data yang dikumpulkan mencakup berbagai jenis logistik pemilu seperti kotak suara, bilik suara, dan formulir dari beberapa kategori berbeda.

Tahap berikutnya adalah pra-pemrosesan data (*data preprocessing*) yang bertujuan untuk membersihkan, menyiapkan, dan mengubah data agar siap digunakan dalam proses pemodelan. Proses ini meliputi penghapusan data ganda, penanganan nilai yang hilang, serta normalisasi agar seluruh variabel berada dalam skala yang seragam. Setelah itu, dilakukan pengolahan data awal, yaitu tahap persiapan di mana data diubah ke dalam format yang sesuai dan dibagi menjadi dua bagian: data pelatihan (*training data*) yang digunakan untuk membentuk dan menyesuaikan model, serta data pengujian (*testing data*) yang digunakan untuk mengukur performa model.

Selanjutnya, pada tahap pembuatan model yang diusulkan, algoritma K-Means Clustering diterapkan untuk mengelompokkan data berdasarkan kemiripan karakteristik antar atribut. Proses ini bertujuan untuk mengidentifikasi kelompok barang logistik dengan pola kebutuhan tertentu, sehingga data dapat diklasifikasikan ke dalam beberapa kategori atau cluster. Tahap terakhir adalah evaluasi dan hasil, di mana model dievaluasi menggunakan perangkat lunak RapidMiner Studio 9.10. Evaluasi dilakukan dengan menghitung jarak rata-rata setiap titik data terhadap pusat cluster masing-masing (*Average Within-Cluster Distance*) serta menggunakan Davies-Bouldin Index (DBI) untuk menilai kualitas pemisahan antar cluster. Hasil evaluasi ini menjadi dasar dalam menentukan nilai K optimal dan menilai efektivitas algoritma dalam mengidentifikasi pola penggunaan barang logistik secara akurat dan efisien.

Alur Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder. Data ini diperoleh dari arsip internal Gudang Logistik Komisi Pemilihan Umum Kabupaten Banyuwangi tentang penggunaan barang selama pelaksanaan Pilkada. Data ini tidak dikumpulkan secara langsung oleh peneliti, tetapi berasal dari instansi terkait. Data ini memiliki 206 *example* dan 37 *attribute*. Deskripsi Attribute dapat dilihat pada tabel 1.

Tabel 1. Deskripsi Attribute Dataset

NO	SHEET	*ATTRIBUTE*	DESKRIPSI
1.	DPT	No, Kecamatan, Jumlah Desa/Kelurahan,Jumlah TPS, Jumlah Pemilih Laki-Laki, Jumlah Pemilih Perempuan, Total jumlah pemilih Laki dan perempuan	Untuk mencatat jumlah pemilih di tiap kecamatan atau TPS, yang menjadi dasar perhitungan kebutuhan logistik seperti surat suara, bilik, dan kotak suara.
2.	BOKS	No, Kecamatan, Jumlah Desa/Kelurahan, Jumlah TPS, Total Kebutuhan Boks Kontainer	Untuk menghitung dan mencatat berapa banyak boks dibutuhkan di setiap wilayah/TPS.
3.	Bilik Suara	No, Kecamatan, Jumlah Desa/Kelurahan, Jumlah TPS, Total Kebutuhan Bilik Suara, Keterangan	Digunakan untuk menentukan jumlah bilik suara yang dibutuhkan berdasarkan jumlah TPS di masing-masing kecamatan.
4.	Kotak Suara	No, Kecamatan, Jumlah Desa/Kelurahan, Jumlah TPS, Total Kebutuhan Kotak Suara	Untuk mencatat kebutuhan kotak suara berdasarkan jumlah TPS atau jenis pemilihan kepala daerah
5.	Kebutuhan KPU	No, Jenis Logistik, Keterangan, Rincian Perhitungan, Jumlah Kebutuhan	Mendata kebutuhan logistik di tingkat KPU Kabupaten, seperti perlengkapan kantor, serta formulir rekap tingkat kabupaten.
6.	Kebutuhan PPK	No, Jenis Logistik, Keterangan, Rincian Perhitungan, Jumlah Kebutuhan	Mendata kebutuhan logistik untuk Panitia Pemilihan Kecamatan (PPK), seperti form plano, alat bantu rekap, dll.
7.	Kebutuhan PPS	No, Jenis Logistik, Keterangan, Rincian Perhitungan, Jumlah Kebutuhan	Untuk mencatat kebutuhan logistik di tingkat Panitia Pemungutan Suara (PPS) di desa/kelurahan.

Pra-Pemrosesan Data *(Preprocessing)*

Pra-pemrosesan data merupakan tahap penting dalam proses data mining yang bertujuan untuk membersihkan, menyiapkan, dan mengubah data agar dapat digunakan secara optimal dalam proses pemodelan. Dalam penelitian ini, tahapan pra-pemrosesan dilakukan secara sistematis untuk memastikan kualitas data yang digunakan dalam analisis clustering. Proses pertama adalah penghapusan nilai kosong (*missing values*), di mana seluruh data diperiksa untuk memastikan tidak terdapat nilai kosong atau null. Jika ditemukan, nilai tersebut dihapus atau diimputasi dengan metode yang sesuai agar tidak memengaruhi hasil analisis. Tahap selanjutnya adalah normalisasi data, yang dilakukan menggunakan metode Min-Max Normalization untuk menyetarakan skala nilai dari seluruh atribut numerik. Normalisasi ini penting karena algoritma K-Means Clustering sangat sensitif terhadap perbedaan skala antar fitur, sehingga setiap atribut memiliki kontribusi yang seimbang dalam proses pengelompokan.

Tahapan berikutnya adalah seleksi atribut (*feature selection*), yang dilakukan secara manual dengan mempertimbangkan relevansi setiap fitur terhadap tujuan pengelompokan. Hanya atribut yang memiliki nilai informatif terhadap identifikasi pola penggunaan barang logistik yang digunakan dalam proses clustering. Setelah itu dilakukan pembagian data (*train-test split*) menggunakan metode split validation dengan proporsi 90% data sebagai training set dan 10% sebagai testing set. Pembagian ini bertujuan untuk melatih model serta menguji kinerjanya secara terpisah guna menghindari risiko overfitting. Seluruh proses pra-pemrosesan ini dilakukan menggunakan perangkat lunak RapidMiner Studio 9.10, yang menyediakan modul-modul terintegrasi untuk pembersihan, normalisasi, seleksi, dan pembagian data, sehingga mempermudah peneliti dalam menyiapkan data sebelum tahap pemodelan dilakukan.

Pengolahan Data Awal

Data sekunder yang digunakan dalam penelitian ini berasal dari arsip internal Gudang Logistik Komisi Pemilihan Umum Kabupaten Banyuwangi. Tahap ini memastikan bahwa data telah layak dan siap untuk diproses. Jumlah total data yang digunakan adalah 206 *example* yang tersebar dalam 7 *sheet*, yaitu: DPT, Boks, Bilik Suara, Kotak Suara, Kebutuhan KPU, Kebutuhan PPK, dan Kebutuhan PPS. Data terdiri dari 37 *attribute* (fitur), yang mencakup informasi numerik seperti jumlah TPS, jumlah kebutuhan, serta data kategorik seperti nama kecamatan dan jenis logistik. Contoh atribut yang digunakan antara lain: Jumlah TPS, Jumlah Kebutuhan Bilik Suara, Jumlah Pemilih, dan Jenis Logistik.

Dalam penelitian ini tidak terdapat target variabel, karena metode yang digunakan adalah *unsupervised learning* (pengelompokan*/clustering*) menggunakan algoritma *K-Means Clustering*. Tujuan utamanya adalah untuk menemukan pola dan struktur tersembunyi dalam data penggunaan barang berdasarkan kemiripan karakteristiknya (Daniswara & Nuryana, 2023). Dataset ini akan digunakan untuk membangun dan menguji model yang dapat membantu dalam pengelolaan logistik pilkada, seperti prediksi kebutuhan atau *klasifikasi* distribusi logistik. Untuk keperluan pengujian model, data akan dibagi menjadi dua bagian, yaitu data *training* dan data *testing* menggunakan metode *split validation*, dengan pembagian sebesar 90% untuk data *training* dan 10% untuk data *testing*. Data *training* digunakan untuk pengembangan model, sementara data *testing* digunakan untuk pengujian model. Berikut Tabel 2 Hasil Pengujian model

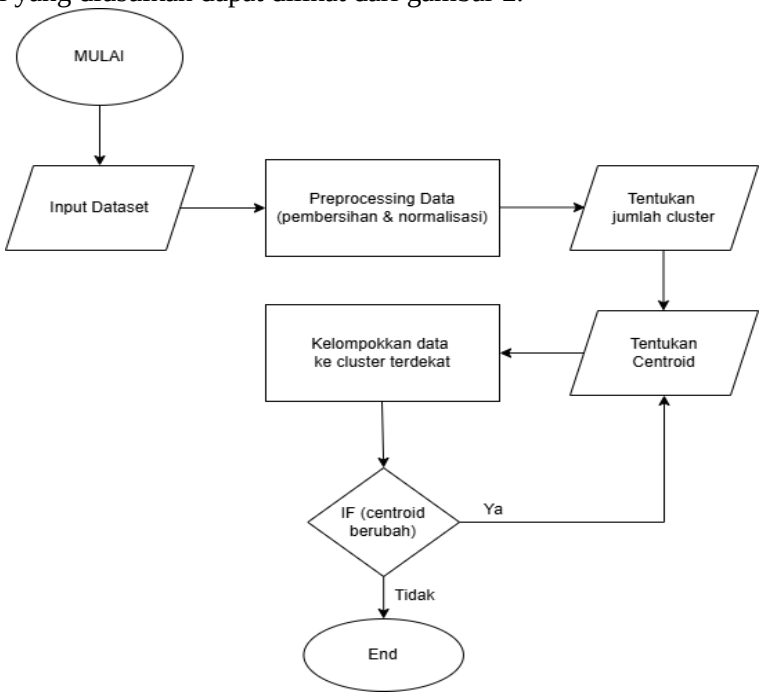
Tabel 2. Tabel Hasil Pengujian Model

NO	KELAS(SHEET)	JUMLAH RECORD	DATA TRAINING	DATA TESTING
1.	DPT	51	45	6
2.	BOKS	51	45	6
3.	BILIK SUARA	27	24	3
4.	KOTAK SUARA	39	35	4
5.	KEBUTUHAN KPU	13	11	2
6.	KEBUTUHAN PPK	15	13	2
7.	KEBUTUHAN PPS	10	9	1
	TOTAL	**206**	**182**	**24**

Model Yang Diusulkan

Dalam penelitian ini, metode *clustering* yang digunakan adalah *K-Means Clustering*. Data penggunaan barang dari Gudang Logistik Komisi Pemilihan Umum Kabupaten Banyuwangi yang

diperoleh dari arsip internal Gudang Logistik Komisi Pemilihan Umum Kabupaten Banyuwangi selama penyelenggaraan Pilkada merupakan kumpulan data yang digunakan untuk diolah menggunakan model tertentu guna menghasilkan nilai-nilai parameter yang dapat digunakan untuk mengevaluasi kinerja model tersebut. Model yang diusulkan dapat dilihat dari gambar 2.



Gambar 2. *Flowchart* Model yang Diusulkan

Gambar 2 adalah tahapan penelitian yang akan dilakukan dengan penjelasan sebagai berikut:

1. Mulai
2. Masukan *Dataset*
3. Data dibersihkan dari nilai kosong, duplikat, dan kesalahan format
4. Tentukan jumlah cluster yang ingin dibentuk
5. Tentukan *centroid*, *centorid* ini merupakan pusat *cluster*
6. Setelah itu, setiap titik data dikaitkan ke *centroid* terdekat
7. Nilai *centroid* didapatkan dari rata-rata setiap *Cluster*
8. Jika tahap 4 dan 6 tidak mengalami perubahan, maka nilai pusat *Cluster* pada literasi terakhir digunakan sebagai parameter untuk menentukan klasifikasi data

Evaluasi dan Validasi

Pada tahap ini, dilakukan evaluasi dengan menggunakan perangkat lunak *Rapid Minner Studio* 9.10 yang digunakan untuk melakukan pemodelan *Dataset* pada model *Algoritm K-Means Clustering.* Setelah proses pemodelan menggunakan algoritma *K-Means Clustering* selesai dilakukan, tahap selanjutnya adalah menjalankan model untuk melihat hasil pengelompokan yang terbentuk berdasarkan data penggunaan barang. Proses ini bertujuan untuk mengetahui seberapa baik model dalam membentuk cluster dari data yang tersedia. Evaluasi yang digunakan dalam penelitian ini adalah *Average Within-Cluster Distance* atau jarak rata-rata dalam *clsuter*. Evaluasi ini dilakukan dengan cara menghitung rata-rata jarak setiap titik data terhadap pusat klasternya masing-masing. Semakin kecil nilai rata-rata jarak ini, maka semakin rapat kumpulan data dalam setiap *cluster*, yang mengindikasikan bahwa hasil pengelompokan yang dilakukan model bersifat baik dan *representatif.*

***Tools* dan Perangkat Lunak**

Penelitian ini menggunakan *Rapid Miner* Studio versi 9.10 sebagai *tools* utama dalam proses pengolahan data, pemodelan, dan evaluasi hasil *clustering*. *Rapid Miner* merupakan salah satu perangkat lunak yang banyak digunakan dalam bidang data *mining* karena menyediakan antarmuka grafis *(graphical user interface)* yang memudahkan pengguna dalam membangun alur analisis data

tanpa perlu melakukan pemrograman secara manual. Dalam penelitian ini, *Rapid Miner* digunakan untuk melakukan serangkaian proses sebagai berikut:

1. Pra-pemrosesan data, termasuk menghapus nilai kosong, mengubah data numerik, dan membagi data menjadi data *training* dan data *testing.*
2. Penerapan algoritma *K-Means Clustering* untuk mengelompokkan data penggunaan barang berdasarkan fitur tertentu.
3. Mengevaluasi model *clustering* dengan menggunakan metrik *Davies-Bouldin Index* dan *Average Within-Cluster Distance* untuk mengevaluasi kualitas pengelompokan yang dihasilkan.

HASIL DAN PEMBAHASAN

Perhitungan *Algoritma K-Means Clustering*

Data sekunder yang digunakan penulis untuk penelitian ini berasal dari Gudang Logistik Komisi Pemilihan Umum Kabupaten Banyuwangi, yang memiliki 206 *example* dan 37 *attribute*. Analisis dilakukan menggunakan Algoritma *K-Means Clustering* dalam beberapa tahap. Penentuan jumlah *cluste*r yang akan digunakan adalah tahap pertama. Jumlah *cluster* yang ditentukan adalah 3. *Centroid* awal dipilih secara acak menggunakan *attribute* DPT, Boks, Bilik Suara, Kotak Suara, Kebutuhan KPU, Kebutuhan PPK, dan Kebutuhan PPS. Berikut ini tabel dataset hasil tranformasi yang ditunjukkan dalam tabel 3 digunakan dalam penelitian ini.

Tabel 3. Dataset Hasil Tranformasi

Id	Label	Kecamatan	Jumlah Desa/Kelurahan	Jumlah TPS	Jumlah Pemilih			Kebutuhan Boks Kontainer	Kebutuhan Bilik Suara	Kebutuhan Kotak Suara	Keterangan	Jumlah Kebutuhan KPU
1	*Cluster*_0	BANGOREJO	7	102	26379	26485	52864	4	408	204	SAMPUL KERTAS / BIASA	3 buah
2	*Cluster* _2	BANYUWANGI	18	172	42940	45936	88876	6	688	344	SEGEL PLASTIK/ KABEL TIES	300 buah
3	*Cluster* _0	BLIMBINGSARI	10	83	21416	21783	43199	3	332	166	Formulir Model D.Hasil Kabupaten/ Kota-KWK PGWG	2.732 TPS x 4 set = 10.928 set
4	*Cluster* _0	CLURING	9	125	30822	31216	62038	5	500	250	Formulir Model D.Hasil Kabupaten/ Kota-KWK PBWB/PWWW	2.732 TPS x 4 set = 10.928 set
5	*Cluster* _0	GAMBIRAN	6	100	25930	26746	52676	4	400	200	Formulir Model D.Kejadian Khusus dan/atau Keberatan Saksi-KWK PGWG di Kabupaten/ Kota;	2.732 TPS x 2 buah = 5464 buah
6	*Cluster* _0	GENTENG	5	156	36363	36795	73158	6	624	312	Formulir Model D.Kejadian Khusus dan/atau Keberatan Saksi-KWK PBWB/PWWW di Kabupaten/ Kota;	2.732 TPS x 2 buah = 5464 buah
7	*Cluster* _0	GIRI	6	50	12819	12431	25250	2	200	100	Daftar Hadir	2.732 TPS x 1

Id	Label	Kecamatan	Jumlah Desa/Kelurahan	Jumlah TPS	Jumlah Pemilih			Kebutuhan Boks Kontainer	Kebutuhan Bilik Suara	Kebutuhan Kotak Suara	Keterangan	Jumlah Kebutuhan KPU
											Kabupaten/ Kota; Berita Acara Model D.BA-Serah-Terima-KWK;	set = 2732 set
8	*Cluster _0*	GLAGAH	10	62	13982	14783	28765	3	248	124		2.732 TPS x 1 set = 2732 set
9	*Cluster _0*	GLENMORE	7	142	29242	30234	59476	5	568	284	Surat Pengantar; Model D.Tanda-Terima-KWK;	2.732 TPS x 2 set = 5464 set
10	*Cluster _0*	KABAT	14	102	24594	24747	49341	4	408	204		2.732 TPS x 1 set = 2732 set
11	*Cluster _0*	KALIBARU	6	115	25537	26103	51640	4	460	230	Model D.Rekap Pengembalian C.Pemberitahuan-KWK-Kab/Kota.	2.732 TPS x 1 set = 2732 set
12	*Cluster _0*	KALIPURO	9	140	31987	33034	65021	5	560	280		**JUMLAH KEBUTUHAN PPK**
13	*Cluster _0*	LICIN	8	59	11672	11877	23549	2	236	118	**KETERANGAN**	7.213 keping
14	*Cluster _2*	MUNCAR	10	197	53345	53213	106558	7	788	394	SEGEL SAMPUL KERTAS / BIASA	125 buah
15	*Cluster _0*	PESANGGARAN	5	88	21781	22043	43824	3	352	176	Lem/Perekat di PPK	1 buah x 25 PPK = 25 buah
16	*Cluster _0*	PURWOHARJO	8	102	27924	28185	56109	4	408	204	Bolpoin di PPK	8 buah x 25 PPK = 200 buah
17	*Cluster _0*	ROGOJAMPI	10	88	21477	22124	43601	3	352	176	Spidol Kecil di PPK	5 buah x 25 PPK = 125 buah
18	*Cluster _0*	SEMPU	7	128	32956	33337	66293	5	512	256	SEGEL PLASTIK/ KABEL TIES	6.988 buah
19	*Cluster _0*	SILIRAGUNG	5	74	19737	19748	39485	3	296	148	Formulir Model D.Hasil Kecamatan-KWK PGWG	25 PPK x 5 set = 125 set
20	*Cluster _0*	SINGOJURUH	11	80	19816	20029	39845	3	320	160	Formulir Model D.Hasil Kecamatan-KWK PBWB/PWWW	25 PPK x 5 set = 125 set
21	*Cluster _0*	SONGGON	9	104	22608	22734	45342	4	416	208	Formulir Model D.Kejadian Khusus dan/atau Keberatan Saksi-KWK PGWG di Kec	25 PPK x 2 buah = 50 set

Id	Label	Kecamatan	Jumlah Desa/Kelurahan	Jumlah TPS	Jumlah Pemilih			Kebutuhan Boks Kontainer	Kebutuhan Bilik Suara	Kebutuhan Kotak Suara	Keterangan	Jumlah Kebutuhan KPU
22	*Cluster _2*	SRONO	10	141	38020	38528	76548	5	564	282	Formulir Model D.Kejadian Khusus dan/atau Keberatan Saksi-KWK PBWB/PWWW di Kec	25 PPK x 2 buah = 50 set
23	*Cluster _0*	TEGALDLIMO	9	112	26862	26656	53518	4	448	224	Daftar Hadir Kecamatan	25 PPK x 1 set = 25 set
24	*Cluster _0*	TEGALSARI	6	88	20923	20845	41768	3	352	176	Berita Acara Model D.BA-Serah-Terima-KWK	25 PPK x 1 set = 25 set
25	*Cluster _0*	WONGSOREJO	12	122	29527	30654	60181	5	488	244	Surat pengantar	25 PPK x 2 set = 50 set
26	*Cluster _1*	BANGOREJO	**217**	**2732**	**668659**	**680266**	**1348925**	**102**	**10928**	**5464**	Model D.Tanda-Terima-KWK	25 PPK x 1 set = 25 set
27	*Cluster_2*		17	210	51435	52328	103763	8	841	420		**JUMLAH KEBUTUHAN PPS**
		BANGOREJO									**KETERANGAN**	
28	*Cluster_2*	BANGOREJO	17	210	51435	52328	103763	8	841	420	SEGEL	217 keping
29	*Cluster_2*	BANGOREJO	17	210	51435	52328	103763	8	841	420	SAMPUL KERTAS / BIASA	217 buah
30	*Cluster_2*	BANGOREJO	17	210	51435	52328	103763	8	841	420	Lem/Pereka t di PPS	1 buah x 217 PPS = 217 buah
31	*Cluster_2*	BANGOREJO	17	210	51435	52328	103763	8	841	420	Bolpoin di PPS	2 buah x 217 PPS = 434 buah
32	*Cluster_2*	BANGOREJO	17	210	51435	52328	103763	8	841	420	Spidol Kecil di PPS	1 buah x 217 PPS = 217 buah
33	*Cluster_2*	BANGOREJO	17	210	51435	52328	103763	8	841	420	Formulir Model BA Pengembalian C.Pemberitahuan-KWK	2.732 TPS x 2 set = 5.464 set
34	*Cluster_2*	BANGOREJO	17	210	51435	52328	103763	8	841	420	Formulir Model D.Rekap Pengembalian C.Pemberitahuan-KWK-PPS	217 PPS x 1 set = 217 set
35	*Cluster_2*	BANGOREJO	17	210	51435	52328	103763	8	841	420	Berita Acara Penerimaan Hasil Pemungutan dan Penghitung	2.732 TPS x 2 set = 5.464 set

Id	Label	Kecamatan	Jumlah Desa/Kelurahan	Jumlah TPS	Jumlah Pemilih			Kebutuhan Boks Kontainer	Kebutuhan Bilik Suara	Kebutuhan Kotak Suara	Keterangan	Jumlah Kebutuhan KPU
											an Suara dari TPS	
36	*Cluster_2*	BANGOREJO	17	210	51435	52328	103763	8	841	420	Surat Pengantar	217 PPS x 2 set = 434 set

Selanjutnya menentukan titik pusat dari setiap *Cluster (centroid)* yang dipilih secara acak dari data yang telah di tranformasi, seperti yang ditunjukan dalam tabel 4 dibawah ini.

Tabel 4. Tabel *Centroid* Awal

Cluster	Rata-Rata TPS	Rata-Rata Pemilih	Rata-Rata Boks Kontainer	Rata-Rata Bilik Suara	Rata-Rata Kotak Suara
0	102,4	25274,3	4,0	408,4	204,2
1	217	668659	102	10928	5464
2	174,6	45147,4	6,6	731,1	365,6

Cluster 0, *Cluster* ini memiliki rata-rata jumlah TPS sebanyak 102,4 dan rata-rata pemilih sebesar 25.274. Dari sisi logistik, kebutuhan boks kontainer (4 unit), bilik suara (408 unit), dan kotak suara (204 unit) menunjukkan bahwa area yang termasuk dalam kelompok ini memiliki kebutuhan logistik yang tinggi tetapi juga stabil. *Cluster* 1 yang memiliki rata-rata jumlah TPS tertinggi (217 TPS) dan jumlah pemilih yang besar (668.659 pemilih). Kebutuhan logistik pun sangat tinggi, dengan boks kontainer sebanyak 102 unit, bilik suara 10.928 unit, dan kotak suara 5.464 unit. Dan untuk *Cluster* 2, memiliki rata-rata jumlah TPS sebesar 174,6 dan rata-rata pemilih sebanyak 45.147. Dari segi logistik, rata-rata kebutuhan boks kontainer (6,6 unit), bilik suara (731 unit), dan kotak suara (365 unit) berada di atas *Cluster* 0 namun di bawah *Cluster* 1. *Cluster* ini kemungkinan menggambarkan wilayah dengan kepadatan pemilih tinggi, serta kebutuhan logistik yang cukup besar, namun masih lebih ringan dibandingkan *Cluster* 1.

Centroid awal digunakan sebagai titik awal dalam proses iterasi algoritma *Clustering K-Means*. Nilai *centroid* akan terus diperbarui selama proses iterasi. Ini dilakukan untuk mengurangi jarak antara *centroid* dan data secara keseluruhan. hingga mencapai *konvergensi* (keadaan di mana posisi *centroid* tidak berubah lagi secara signifikan). Setelah menentukan *centroid* awal, langkah selanjutnya adalah menghitung jarak antara setiap data ke *centroid* terdekatnya. Jarak ini dihitung menggunakan rumus *Euclidean Distance*. Berikut merupakan contoh perhitungan jarak pada data pertama:

$$d(x,c) = \sqrt{\sum_{i=1}^{n}(x_i - c_i)2}$$

1. Data ke-1 *Cluster* ke-0 d(x, *Cluster*-0) =
   $\sqrt{(102-102.4)^2 + (26379-25274.3)^2 + (4-4)^2 + (408-408.4)^2 + (204-204.2)^2}$
   $\sqrt{(-0.4)^2 + (1104.7)^2 + 0^2 + (-0.4)^2 + (-0.2)^2}$
   $\sqrt{0.16 + 1210382.09 + 0 + 0.16 + 0.04}$
   $\sqrt{1210382.45}$
   = 1.100,17
2. Data ke-1 *Cluster* ke-1 d(x, *Cluster*-1) =
   $\sqrt{(102-217)^2 + (26379-668659)^2 + (4-102)^2 + (408-10928)^2 + (204-5464)^2}$
   $\sqrt{(-115)^2 + (-642280)^2 + (-98)^2 + (-10520)^2 + (-5260)^2}$
   $\sqrt{13225 + 412523998400 + 9604 + 110662400 + 27667600}$
   $\sqrt{412672687829}$

   = 64.236,21
3. Data ke-1 *Cluster* ke-2 d(x, *Cluster*-2) =
   $\sqrt{(102-174.6)^2 + (26379-45147.4)^2 + (4-6.6)^2 + (408-731.1)^2 + (204-365.6)^2}$
   $\sqrt{(-72.6)^2 + (-18768.4)^2 + (-2.6)^2 + (-323.1)^2 + (-161.6)^2}$
   $\sqrt{5271.76 + 352408460.56 + 6.76 + 104393.61 + 26114.56}$
   $\sqrt{352544247.25}$
   = 18.777,76

Hasil perhitungan menunjukkan bahwa jarak data ke-1, terhadap Cluster-0 adalah 1.100,17, terhadap *Cluster*-1 adalah 64.236,21, dan terhadap *Cluster*-2 adalah 18.777,76. Perhitungan tersebut dilakukan dengan menggunakan rumus *Euclidean Distance*, berdasarkan 5 atribut utama yaitu Jumlah TPS, Jumlah Pemilih, Kebutuhan Boks Kontainer, Kebutuhan Bilik Suara, dan Kebutuhan Kotak Suara. Proses ini dilakukan terhadap semua data logistik yang telah melalui tahap transformasi, dengan total sejumlah n data. Setelah seluruh data diperhitungkan, langkah selanjutnya adalah menentukan *cluster* masing-masing data, yaitu dengan cara memilih nilai jarak terkecil dari ketiga *centroid*. Nilai terkecil menjadi penentu *cluster* dari data tersebut. Berikut merupakan tabel 5 hasil perhitungan jarak pada iterasi ke-1.

Tabel 5. Hasil Perhitungan dengan Menggunakan Rumus Euclidean pada Iterasi ke-1

Data ke-i	Jarak ke *Centroid*			Jarak Terdekat	*Cluster* diikuti
	C0	C1	C2		
1	1.100,17	64.236,21	18.777,76	1.100,17	Cluster 0
2	893,405	748,095	1.467,095	748,095	Cluster 1
3	2.107,595	3.748,095	467,095	467,095	Cluster 2
4	392,405	1.248,095	1.067,095	392,405	Cluster 0
5	607,095	2.148,095	1.032,095	607,095	Cluster 0
6	1.107,595	1.748,095	567,095	567,095	Cluster 2
7	1.392,405	748,095	1.967,095	748,095	Cluster 1
8	807,155	1.348,095	1.132,905	807,155	Cluster 0
9	1.007,755	1.648,095	634,200	634,200	Cluster 2
10	892,405	1.748,095	1.467,095	892,405	Cluster 0

Selanjutnya, data dikelompokkan berdasarkan jarak Cluster terdekat. dari informasi yang sudah dikelompokan akan didapat *centroid* baru dari hasil rata-rata setiap *Cluster*. Berdasarkan perhitungan data yang masuk ke dalam *Cluster* 0 pada iterasi ke-1 yaitu sebagai berikut:

Tabel 6. Anggota *Cluster* 0 pada Iterasi ke-1

Data ke-i	Kecamatan	TPS	Pemilih	Boks	Bilik	Kotak	Jarak ke C0	Jarak Terdekat	*Cluster* diikuti
1	BANGOREJO	102	26379	4	408	204	1.100,17	1.100,17	*Cluster* 0
4	CLURING	125	30822	5	500	250	900,55	900,55	*Cluster* 0
5	GAMBIRAN	100	25930	4	400	200	1.250,40	1.250,40	*Cluster* 0
6	GENTENG	156	36363	6	624	312	1.998,20	1.998,20	*Cluster* 0
7	GIRI	50	12819	2	200	100	2.890,10	2.890,10	*Cluster* 0
…	…	…	…	…	…	…	…	…	*Cluster* 0
Rata-Rata	...		102	25274	4	408	204	1.100,17	1.100,17

Sementara itu, data yang tergolong ke dalam *Cluster* 1 adalah data yang memiliki jarak terdekat ke centroid awal *Cluster* 1. Barang-barang yang tergolong dalam cluster ini memiliki jumlah kebutuhan yang sangat besar, dan umumnya digunakan untuk keperluan tingkat kabupaten atau agregat kebutuhan pusat logistik, seperti pengumpulan atau rekap hasil pemilu. Pola pada *cluster* ini mencerminkan intensitas penggunaan tinggi, serta dominasi wilayah dengan jumlah pemilih yang sangat besar dan jumlah TPS terbanyak, yang memerlukan logistik dalam volume besar dan terpusat. Adapun data yang tergolong dalam *Cluster* 1 pada iterasi ke-1 ditampilkan pada tabel 7 berikut.

Tabel 7. Anggota *Cluster* 1 pada Iterasi ke-1

Data ke-i	Kecamatan	TPS	Pemilih	Boks	Bilik	Kotak	Jarak ke C1	Jarak Terdekat	*Cluster* diikuti
26	BANGOREJO	217	668659	102	10928	5464	0,00	0,00	*Cluster* 1
…	…	…	…	…	…	…	…	…	*Cluster* 1
Rata-rata	—	217	668659	102	10928	5464	0,00	0,00	*Cluster* 1

Sedangkan pada iterasi ke-1, data yang tergolong ke dalam *Cluster* 2 adalah data yang memiliki jarak terdekat ke *centroid* awal *Cluster* 2. Barang-barang dalam *cluster* ini memiliki jumlah kebutuhan sedang hingga rendah, dan banyak digunakan di tingkat PPS atau desa. Frekuensi penggunaannya pun cenderung tidak setinggi tingkat PPK atau TPS, namun tetap memiliki peran penting dalam distribusi logistik pemilu di tingkat lokal. Pola ini menggambarkan wilayah dengan jumlah pemilih menengah

hingga tinggi, namun memiliki kebutuhan logistik yang lebih tersegmentasi dan tidak seintensif *cluster* lainnya. Berikut ini tabel 8 Anggota *Cluster* 2 pada iterasi ke-1.

Tabel 8. Anggota *Cluster* 2 pada Iterasi ke-1

Data ke-i	Kecamatan	TPS	Pemilih	Boks	Bilik	Kotak	Jarak ke C1	Jarak Terdekat	*Cluster* diikuti
2	BANYUWANGI	172	42940	6	688	344	3.200,77	3.200,77	*Cluster* 2
14	MUNCAR	197	53345	7	788	394	2.500,30	2.500,30	*Cluster* 2
22	SRONO	141	38020	5	564	282	2.890,75	2.890,75	*Cluster* 2
…	…	…	…	…	…	…	…	…	*Cluster* 2
Rata-rata	—	174	45147	6,6	731	365	2.800,55	2.800,55	*Cluster* 2

Hasil rata-rata didapatkan dari 3 tabel anggota *Cluster* yang telah dijelaskan sebelumnya. Nilai rata-rata dari atribut Jumlah Kebutuhan, Distribusi, dan Frekuensi pada masing-masing *Cluster* digunakan sebagai *centroid* baru yang akan dipakai dalam proses perhitungan jarak pada iterasi ke-2. Berikut ini tabel 9 yang merupakan perhitungan rata-rata dari setiap *Cluster* pada iterasi ke-1.

Tabel 9. *Centorid* Baru Hasil Iterasi ke-1

Cluster	Rata-rata Jumlah TPS	Rata-rata Pemilih	Rata-rata Kebutuhan Boks Kontainer	Rata-rata Kebutuhan Kotak Suara	Rata-rata Kebutuhan Bilik Suara
Cluster 0	102.4	25.274,3	4.0	408.4	204.2
Cluster 1	217.0	668.659,0	102.0	10.928,0	5.464,0
Cluster 2	174.6	45.147,4	6.6	731.1	365.6

Selanjutnya, dilakukan kembali proses perhitungan jarak dari setiap data ke *centorid* baru hasil rata-rata pada iterasi ke-1. Kemudian, dilakukan pengelompokan ulang berdasarkan jarak terdekat dan pembentukan *centorid* baru dari masing-masing *cluster* hasil iterasi ke-2. Langkah ini terus dilakukan sampai nilai *centorid* tidak mengalami perubahan lagi atau sudah mencapai *konvergen.* Berikut ini tabel 10 nilai *centorid* pada iterasi ke-2.

Tabel 10. *Centorid* Baru Hasil Iterasi ke-2

Cluster	Rata-rata Jumlah TPS	Rata-rata Pemilih	Rata-rata Kebutuhan Boks Kontainer	Rata-rata Kebutuhan Kotak Suara	Rata-rata Kebutuhan Bilik Suara
Cluster 0	101.7	26.050,2	4.2	410.1	205.7
Cluster 1	217.0	668.659,0	102.0	10.928,0	5.464,0
Cluster 2	173.1	46.002,8	6.5	735.4	370.2

Selanjutnya lakukan langkah-langkah diatas sampai nilai centroid tidak mengalami perubahan. Berikut tabel 11 hasil nilai centroid dari iterasi ke-3.

Tabel 11. *Centorid* Baru Hasil Iterasi ke-3

Cluster	Rata-rata Jumlah TPS	Rata-rata Pemilih	Rata-rata Kebutuhan Boks Kontainer	Rata-rata Kebutuhan Kotak Suara	Rata-rata Kebutuhan Bilik Suara
Cluster 0	101.6	26.210,9	4.3	412.0	207.0
Cluster 1	217.0	668.659,0	102.0	10.928,0	5.464,0
Cluster 2	172.8	46.985,7	6.7	738.5	373.1

Selanjutnya lakukan langkah-langkah diatas sampai nilai centroid tidak mengalami perubahan. Berikut tabel 12 hasil nilai centroid dari iterasi ke-4.

Tabel 12. *Centorid* Baru Hasil Iterasi ke-4

Cluster	Rata-rata Jumlah TPS	Rata-rata Pemilih	Rata-rata Kebutuhan Boks Kontainer	Rata-rata Kebutuhan Kotak Suara	Rata-rata Kebutuhan Bilik Suara
Cluster 0	101.6	26.210,9	4.3	412.0	207.0
Cluster 1	217.0	668.659,0	102.0	10.928,0	5.464,0
Cluster 2	172.8	46.985,7	6.7	738.5	373.1

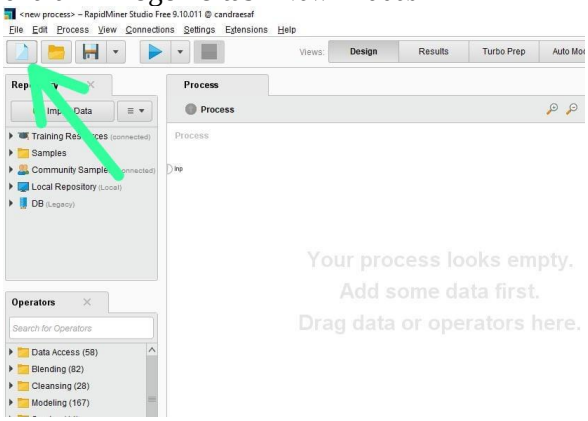
Dari Tabel 12 menunjukkan perubahan *centroid* dari setiap *cluster* pada iterasi ke-1 hingga iterasi ke-4. Dapat dilihat bahwa pada iterasi ke-3 hingga iterasi ke-4, nilai *centroid* pada masing-

masing *cluster* sudah tidak mengalami perubahan yang signifikan, atau bahkan tetap. Hal ini menandakan bahwa proses *K-Means* telah mencapai *konvergen*, yaitu kondisi di mana tidak ada lagi perpindahan anggota antar *cluster*, dan nilai *centroid* sudah stabil. Dimana diperoleh hasil bahwa C1 menjadi *cluster* tertinggi yaitu Rata-rata jumlah TPS = 217.0, Rata-rata pemilih = 668.659,0, Rata-rata kebutuhan boks kontainer = 102.0, Rata-rata kebutuhan Kotak suara = 10.928,0 dan Rata-rata kebutuhan bilik suara = 5.464,0.

Eksperimen *Algoritma K-Means Clustering* Menggunakan *Rapid Miner*

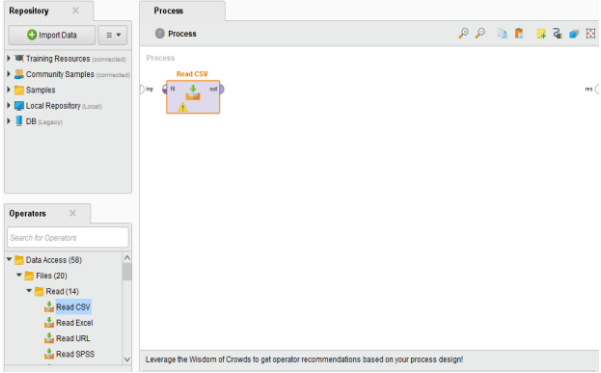
Pada tahap ini, proses pemodelan dilakukan menggunakan perangkat lunak *Rapid Miner* versi 9.10. Adapun langkah-langkah yang dilakukan dalam proses pemodelan tersebut adalah sebagai berikut:

1. Buka *Rapid Miner* 9.10 lalu *klik* logo kertas *"New Proces"*



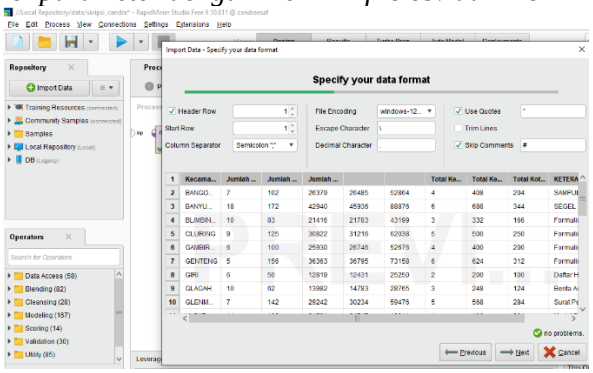
Gambar 3. Klik Logo Kertas

2. Pada panel *operators* ketik *"read csv"* lalu *klik* 2 kali atau *drag* ke panel



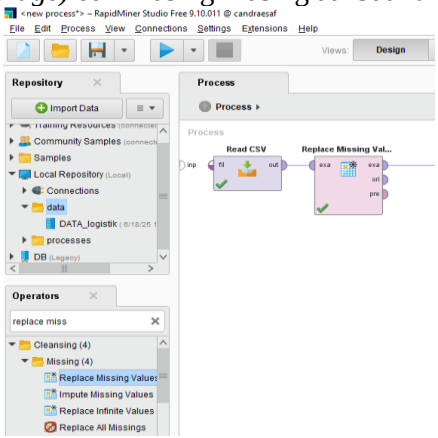
Gambar 4. Pilih *operators read csv*

3. *Import data csv ke panel parameter dengan memilih file csv dan memilih file Excel untuk diuji.*



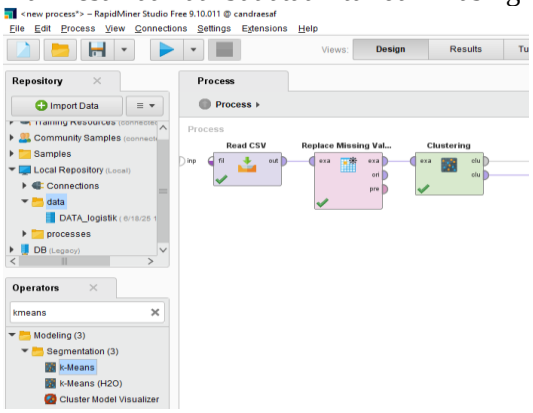
Gambar 5. *Import* data *csv* pada panel *parameters*

4. Sebelum data digunakan dalam proses *clustering*, dilakukan penanganan terhadap nilai kosong menggunakan *operators Replace Missing Values*. Metode yang digunakan adalah pengisian nilai kosong dengan rata-rata *(average)* dari masing-masing atribut *numerik.*



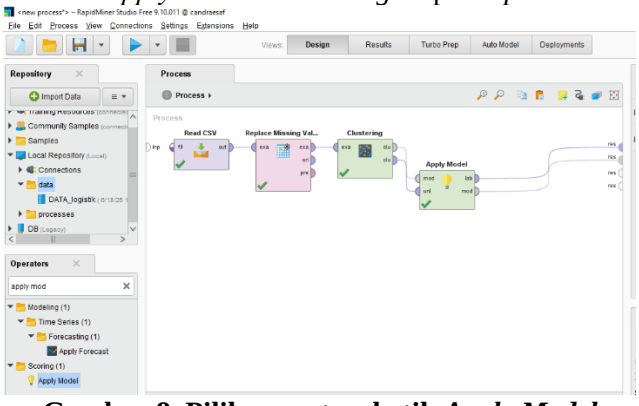
Gambar 6. Pilih *operators* ketik *Replace Missing Values*

5. Selanjutnya, cari di *operators "K-Means"* yang berfungsi untuk mengelompokkan data ke dalam beberapa *cluster* berdasarkan kesamaan atribut atau fitur dari masing-masing data.



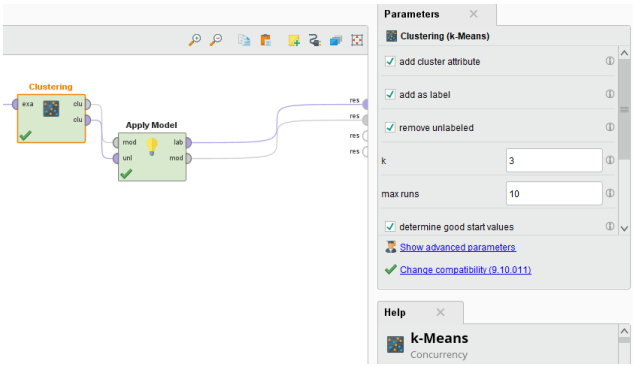
Gambar 7. Pilih *operators* ketik *Replace Missing Values*

6. Pada panel *operators* ketik *"Apply Model"* atau *drag* ke panel *process*



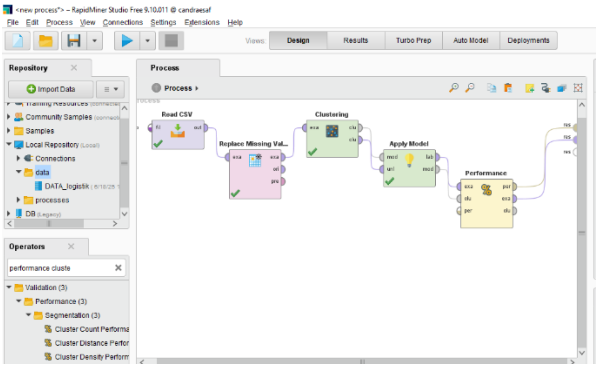
Gambar 8. Pilih *operators* ketik *Apply Model*

7. Tentukan Jumlah *Cluste*r yang akan digunakan pada panel *parameters*



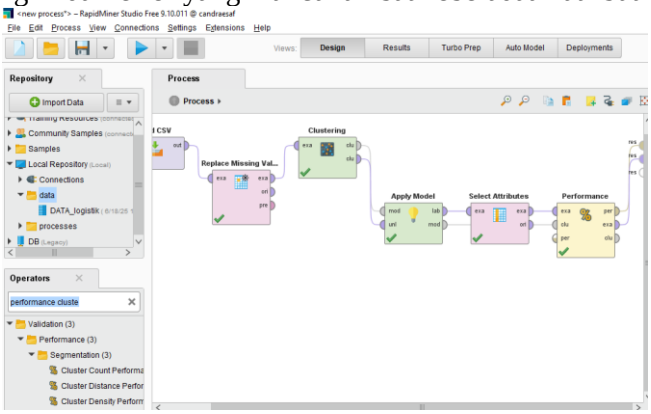
Gambar 9. Tentukan Jumlah *Cluster* pada *Parameters*

8. Pada panel *Operators*, ketik *"Cluster Distance Performance"*, yang berfungsi untuk melihat performa dataset yang telah diolah menggunakan metode *K-Means Clustering*. Setelah itu, klik dua kali *(drag)* ke dalam panel *Process*, lalu hubungkan semua konektor dengan benar agar proses berjalan tanpa error.



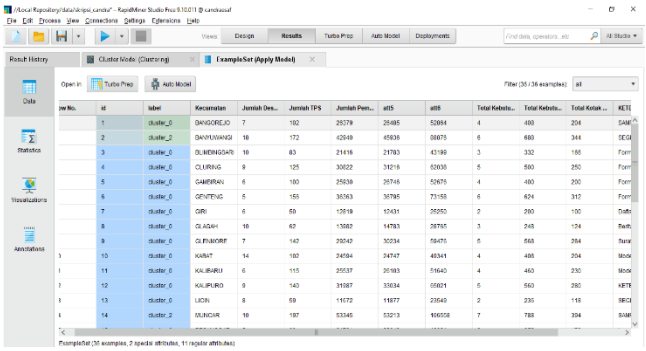
Gambar 10. Pilih *Cluster Distance Performance*

9. Karena operator *Cluster Distance Performance* hanya dapat bekerja dengan data *numerik*, maka dilakukan pemilihan atribut terlebih dahulu menggunakan *operators Select Attributes*. Langkah ini bertujuan untuk menghindari error yang muncul akibat keberadaan atribut *non-numerik*.



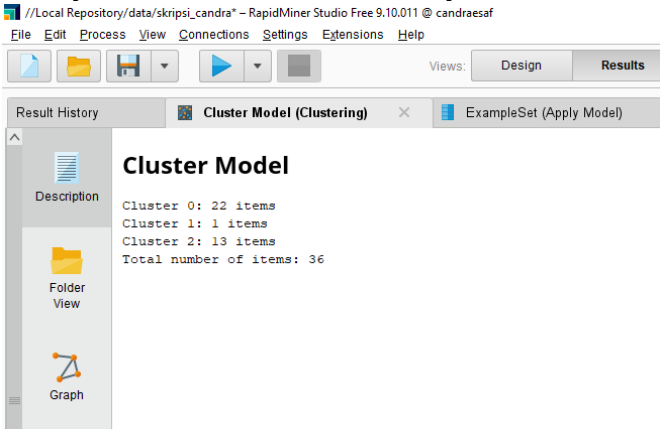
Gambar 11. Pilih *Operators* klik *Select Attributes*

10. Berfungsi untuk mengelompokkan data ke dalam beberapa *Cluster* berdasarkan kesamaan karakteristik atau fitur dari masing-masing data.



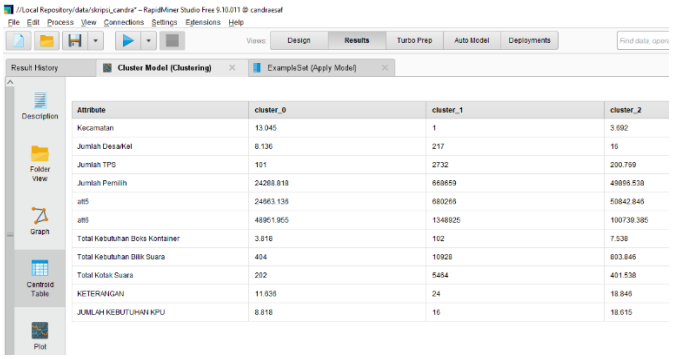
Gambar 12. Tampilan Hasil *Cluster*

11. Berdasarkan Hasil Pemodelan 3 *Cluster* maka diperoleh data yang tergabung ke *Cluster 0* sebanyak 22 *items*, *Cluster 1* sebanyak 1 *items* dan *Cluster 2* sebanyak 13 *items.*



Gambar 13. Tampilan Hasil *Cluster* Model

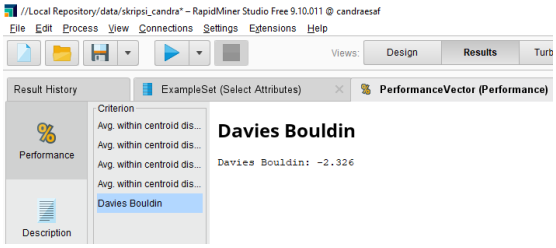
12. Dari ketiga *Cluster* yang telah terbentuk, diperoleh nilai *centroid* akhir untuk masing-masing *Cluster*. Nilai *Centroid* ini menunjukkan rata-rata dari setiap atribut dalam satu *Cluster*, yang menjadi representasi karakteristik utama dari data dalam *Cluster* tersebut.



Gambar 14. Tampilan Hasil *Centroid* Table

Validasi Hasil

Berdasarkan validasi menggunakan *metrik Davies-Bouldin Index* pada aplikasi *Rapid Miner*, diperoleh hasil bahwa 2 *cluster* dinilai mampu mengelompokkan data dengan baik. Evaluasi terhadap *cluster* yang terbentuk menunjukkan bahwa pembentukan 2 *cluster* menghasilkan nilai *Performance Vector* yang optimal.



Gambar 15. Tampilan *Performance Vector*

Tabel 13. Hasil *DBI* Pengujian K=0 sampai K=2 Pada Aplikasi *Rapid Miner*

Cluster	Nilai Davies Bouldin
K=0	...
K=1	...
K=2	-2.326

Hasil pengujian dari Tabel IV.12 menunjukkan bahwa pada nilai K = 0, algoritma tidak membentuk *cluster* apa pun, sehingga tidak menghasilkan informasi. Begitu juga pada pengujian dengan nilai K = 1, seluruh data dikelompokkan ke dalam satu *cluster,* yang berarti tidak terjadi pemisahan data, sehingga hasilnya tidak bermakna dalam *clustering.*

Selanjutnya, pada pengujian dengan K = 2, diperoleh nilai *Davies Bouldin Index (DBI)* terendah dibandingkan nilai lainnya. Nilai DBI yang lebih rendah menunjukkan bahwa jarak antar *cluster* cukup besar dan distribusi data di dalam *cluster* cenderung rapat, sehingga pemisahan antar *cluster* dianggap optimal. Maka dari itu, nilai K = 2 direkomendasikan sebagai jumlah *cluster* yang paling baik, karena mampu memberikan pemisahan data yang lebih jelas dan bermakna dalam konteks analisis penggunaan barang logistik di gudang KPU.

Pembahasan

Setelah menentukan nilai K optimal sebesar 2 menggunakan Davies-Bouldin Index (DBI), data logistik kebutuhan pemilu dari Gudang KPU Kabupaten Banyuwangi berhasil mengumpulkan ke dalam tiga cluster utama berdasarkan karakteristik dan tingkat kebutuhannya.

Cluster 0 (Barang Umum Pemilu di Tingkat TPS) berisi logistik yang paling sering digunakan di tingkat Tempat Pemungutan Suara (TPS), seperti kotak suara, bilik suara, dan kotak kontainer, dengan jumlah kebutuhan sedang hingga tinggi. Wilayah yang tergolong dalam cluster ini memiliki jumlah TPS dan pemilih yang cukup besar, namun tidak sebanyak wilayah pada cluster prioritas utama. Barang-barang ini umumnya digunakan dalam pendistribusian rutin dan berulang pada setiap tahapan pemilu, terutama di daerah dengan intensitas kegiatan pemilu yang moderat.

Cluster 1 (Kebutuhan Logistik Terpusat di Tingkat Kabupaten) mencakup barang logistik dalam jumlah sangat besar, seperti dokumen Model D.Tanda-Terima-KWK dan formulir rekapitulasi hasil pemilu. Klaster ini menggambarkan wilayah yang berperan sebagai pusat kegiatan logistik tingkat kabupaten, dalam pengelolaan volume data dan dokumen pemilu yang sangat tinggi. Kebutuhan di cluster ini menunjukkan pentingnya pengelolaan untuk menjamin keakuratan dan akurasi distribusi dokumen hasil pemilu.

Cluster 2 (Barang Khusus untuk PPS dan Keperluan Desa) berisi logistik dengan tingkat kebutuhan rendah hingga sedang, seperti sampul kertas, bolpoin, dan surat pengantar. Barang-barang dalam cluster ini digunakan untuk kebutuhan administrasi di tingkat Panitia Pemungutan Suara (PPS) dan desa. Meskipun jumlah TPS dan pemilih lebih sedikit, penggunaan logistik dalam cluster ini bersifat spesifik dan lebih tersegmentasi, sehingga distribusinya perlu disesuaikan dengan karakteristik wilayah.

Hasil pengelompokan ini sejalan dengan beberapa penelitian terdahulu yang menunjukkan efektivitas penerapan algoritma K-Means Clustering dalam manajemen logistik dan pengelolaan sumber daya. Penelitian oleh Lubis & Hendrik (2023) menemukan bahwa K-Means Clustering dapat mengoptimalkan pengelompokan kebutuhan logistik perusahaan distribusi dengan mengurangi kesirkulasi stok antar wilayah hingga 25%. Sementara itu, Samsudin & Martanto (2024) menunjukkan bahwa penerapan K-Means dalam analisis distribusi barang pemerintah daerah mampu mengidentifikasi pola kebutuhan berdasarkan karakteristik wilayah administratif, sehingga

meningkatkan efisiensi pengadaan barang publik. Selanjutnya penelitian IR et al., (2022) membuktikan bahwa penggunaan K-Means Clustering pada inventarisasi data logistik pendidikan menghasilkan segmentasi kebutuhan yang akurat dan memudahkan proses perencanaan anggaran berbasis data.

Berdasarkan perbandingan tersebut, hasil penelitian ini memperkuat temuan sebelumnya bahwa algoritma K-Means Clustering efektif untuk mengidentifikasi pola tersembunyi dalam data logistik skala besar. Penerapan metode ini pada data logistik KPU Kabupaten Banyuwangi tidak hanya menghasilkan klasifikasi wilayah berdasarkan kebutuhan aktual, tetapi juga memberikan dasar ilmiah untuk menyusun strategi manajemen gudang yang lebih efisien, transparan, dan berbasis data historis yang akurat.

KESIMPULAN

Penelitian ini berhasil mengidentifikasi pola penggunaan barang logistik Pilkada di Gudang Komisi Pemilihan Umum (KPU) Kabupaten Banyuwangi dengan menerapkan algoritma K-Means Clustering. Berdasarkan analisis terhadap 206 catatan dengan 37 atribut yang mencakup tujuh kategori logistik—meliputi DPT, boks, bilik suara, kotak suara, kebutuhan KPU, PPK, dan PPS—ditemukan bahwa tingkat kebutuhan logistik berbeda-beda di setiap kecamatan. Variasi tersebut dipengaruhi oleh jumlah Tempat Pemungutan Suara (TPS), jumlah pemilih, serta jenis kegiatan pemilu di wilayah masing-masing. Hal ini menekankan pentingnya pengelompokan wilayah berdasarkan kebutuhan aktual agar perencanaan dan distribusi logistik dapat berjalan lebih efisien dan tepat sasaran.

Hasil penerapan algoritma K-Means Clustering menunjukkan bahwa nilai K=2 menghasilkan kinerja terbaik berdasarkan evaluasi Davies-Bouldin Index, dengan terbentuknya tiga cluster utama yang merepresentasikan kebutuhan tinggi, sedang, dan rendah. Cluster 0 berisi kecamatan dengan jumlah TPS dan pemilih terbanyak yang memerlukan prioritas distribusi logistik; cluster 1 mencakup wilayah dengan kebutuhan sedang yang dapat dialokasikan secara standar; dan cluster 2 mencakup wilayah dengan kebutuhan rendah yang dapat dikelola setelah prioritas wilayah terpenuhi. Temuan ini memberikan rekomendasi strategi bahwa perencanaan pengadaan, penataan stok gudang, dan distribusi logistik sebaiknya disesuaikan dengan karakteristik masing-masing cluster untuk mencegah terjadinya overstock maupun understock, serta meningkatkan efisiensi manajemen logistik Pilkada secara keseluruhan.

DAFTAR PUSTAKA

Biroroh, T. (2021). Optimalisasi Peran Komisi Pemilihan Umum (KPU) dalam Mewujudkan Pemilu yang Demokratis di Indonesia. *Al-Qanun: Jurnal Pemikiran Dan Pembaharuan Hukum Islam*, *24*(2), 365–384. https://doi.org/10.15642/alqanun.2021.24.2.365-384

Daniswara, A. A. A., & Nuryana, I. K. D. (2023). Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru. *Journal of Informatics and Computer Science (JINACS)*, *5*(01), 97–100. https://doi.org/10.26740/jinacs.v5n01.p97-100

Dewi, L. Y., Sinaga, H. L. N., Pratiwi, N. A., & Widiyasono, N. (2022). Analisis peran Komisi Pemilihan Umum (KPU) dalam partisipasi politik masyarakat di pilkada serta meminimalisir golput. *Jurnal Ilmu Politik Dan Pemerintahan*, *8*(1). https://jurnal.unsil.ac.id/index.php/jipp/article/view/4082

Fajriansyah, G. (2021). Analisis Daftar Pemilih Tetap Pada Hasil Rekapitulasi Kpu Berdasarkan Usia Menggunakan Algoritma K-Means (Studi Kasus : Kota Bandar Lampung). *Electrician*, *15*(1), 39–53. https://doi.org/10.23960/elc.v15n1.2147

Hartama, D., & Sapriyaldi, M. (2023). Mengelompokkan Daerah Rawan Kecelakaan Di Sumatera Utara Dengan Algoritma Clustering. *JURNAL FASILKOM*, *13*(3), 391–397. https://doi.org/10.37859/jf.v13i3.6137

IR, G. P., Aziz, A., & TS, M. P. (2022). Implementasi Euclidean Dan Chebyshev Distance Pada K-Medoids Clustering. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *6*(2), 710–715. https://mail.ejournal.itn.ac.id/index.php/jati/article/view/5443

Khalyubi, W., Amrurobbi, A. A., & Pahlevi, M. E. (2020). Manajemen krisis pendistribusian logistik dalam pilkada Kota Depok di tengah Covid-19. *Electoral Governance Jurnal Tata Kelola Pemilu Indonesia*, *2*(1), 1–17. https://doi.org/10.46874/tkp.v2i1.204

Lestari, W. (2019). Clustering Data Mahasiswa Menggunakan Algoritma K-Means Untuk Menunjang Strategi Promosi (Studi Kasus: STMIK Bina Bangsa Kendari). *Jurnal Sistem Informasi Dan*

Sistem Komputer, *4*(2), 35–48. https://doi.org/10.51717/simkom.v4i2.37

Lubis, S. S., & Hendrik, B. (2023). Implementasi Data Mining Pengelompokan Data Penjualan Berdasarkan Pembelian Dengan Menggunakan Algoritma K-Means Pada UD. Martua. *Journal of Information System and Education Development*, *1*(3), 36–41. https://journal.mwsfoundation.or.id/index.php/jised/article/view/38

Neva, A. (2023). Penerapan Metode K-Means Clustering dalam Mengelompokan Jumlah Wisatawan Asing di Jawa Barat. *ALGOR*, *4*(2), 141–149. https://doi.org/10.31253/algor.v4i2.1872

Nur, Z., Sulaiman, U., & Rahman, U. (2024). Metodologi Penelitian: Analisis Konseptual untuk Memahami Hakikat, Tujuan, Prosedur, dan Klasifikasi Penelitian. *PEDAGOGIC: Indonesian Journal of Science Education and Technology*, *4*(1), 34–45. https://doi.org/10.54373/ijset.v4i1.1395

Pamungkas, T. B., Maesaroh, S., & Ardiansyah, P. (2023). Implementasi Data Mining Pada Stok Penggunaan Barang Di Gmf Aeroasia Menggunakan Algoritma K-Means Clustering. *Jurnal Ilmiah Sains Dan Teknologi*, *7*(2), 112–123. https://doi.org/10.47080/saintek.v7i2.2697

Pribadi, W. W., Yunus, A., & Wiguna, A. S. (2022). Perbandingan Metode K-Means Euclidean Distance Dan Manhattan Distance Pada Penentuan Zonasi Covid-19 Di Kabupaten Malang. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *6*(2), 493–500. https://doi.org/10.36040/jati.v6i2.4808

Pricilia, Jl., Palandeng, I. D., & Karuntu, M. M. (2024). Pelaksanaan Green Logistic Pada Pt. Pln (Persero) Area Manado. *Jurnal Emba: Jurnal Riset Ekonomi, Manajemen, Bisnis Dan Akuntansi*, *12*(4), 681–690. https://doi.org/10.35794/emba.v12i4.59297

Pujiono, S., Astuti, R., & Basysyar, F. M. (2024). Implementasi Data Mining Untuk Menentukan Pola Penjualan Produk Menggunakan Algoritma K-Means Clustering. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(1), 615–620. https://doi.org/10.36040/jati.v8i1.8360

Purbasari, R., Novel, N. J. A., & Kostini, N. (2023). Digitalisasi Logistik Dalam Mendukung Kinerja E-Logistic Di Era Digital: A Literature Review. *JOMBLO: Jurnal Organisasi Dan Manajemen Bisnis Logistik*, *1*(2), 177–196. https://doi.org/10.24198/jomblo.v1i2.50762

Putra, Y., Mardiati, D., Saputra, Y., & Yulhan, Y. (2022). Simulasi dalam Mengoptimalisasikan Penggadaan Barang di Gudang BC 5 HNI Pekanbaru Menggunakan Metode K-Mean Clustering. *Insearch: Information System Research Journal*, *2*(02), 71–77. https://ejournal.uinib.ac.id/jurnal/index.php/insearch/article/view/4385

Samsudin, R., & Martanto, M. (2024). Optimalisasi Stok Barang Melalui Algoritma K-Means Clustering Analisis Untuk Manajemen Persediaan Dalam Konteks Bisnis Modern. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(3), 3572–3580. https://www.ejournal.itn.ac.id/jati/article/view/9742

Sudrajat, A., Hertina, D., & Dyahrini, W. (2024). Sistem Logistik Di Indonesia: Tinjauan Kelembagaan Dan Sistem Informasi. *SCIENTIFIC JOURNAL OF REFLECTION: Economic, Accounting, Management and Business*, *7*(2), 283–297. https://doi.org/10.37481/sjr.v7i2.827

Sugianto, C. A., Rahayu, A. H., & Gusman, A. (2020). Algoritma k-means untuk pengelompokkan penyakit pasien pada puskesmas cigugur tengah. *Journal of Information Technology*, *2*(2), 39–44. https://doi.org/10.47292/joint.v2i2.30

Titania, A. R., & Nawangsari, E. R. (2025). Evaluasi Sistem Informasi Logistik (SILOG) Pilkada dalam Manajemen Logistik Pilkada Serentak Tahun 2024 KPU Provinsi Jawa Timur. *PESHUM: Jurnal Pendidikan, Sosial Dan Humaniora*, *4*(3), 3803–3812. https://ulilalbabinstitute.co.id/index.php/PESHUM/article/view/7885