

## Implementasi Data Mining Dalam Menentukan Penjualan Alat Perabot Dengan Menggunakan Metode K-Means Clustering Pada PT.XYZ

Silvia Lestari<sup>1</sup>; Rahmatun Nazila<sup>2</sup>; Lukna Aulia Ulhar<sup>3</sup>; Muhammad Zidane<sup>4</sup>

Universitas Potensi Utama

E-mail : [silvialestai.via96@gmail.com](mailto:silvialestai.via96@gmail.com)<sup>1</sup>; [rahmatunnazila05@gmail.com](mailto:rahmatunnazila05@gmail.com)<sup>2</sup>;  
[luknaulhar@gmail.com](mailto:luknaulhar@gmail.com)<sup>3</sup>; [muhhammadzidane@gmail.com](mailto:muhhammadzidane@gmail.com)<sup>4</sup>

**Abstract:** The increasing development of technology and increasing buying and selling activities require every business owner to adapt to technological developments, for business owners selling or processing sales data is very important as is done by PT.XYZ household furniture items such as tables, chairs, dressers, wardrobes, sofas and many more, where sales data is still done manually, such as a lack of reviewing what products consumers need and ineffective data storage. To overcome this problem the researcher tried to implement it using one of the methods available in data mining, namely the K-Means Algorithm. –certain data groups (clusters). So, by grouping this data, the company can find out which items are selling best and which are not selling well. So that the goods in the warehouse do not pile up. From this research, the resulting output is 5 of the best-selling items, and 5 of the least-selling items. With the data processing carried out, it is hoped that it can provide solutions to the company so that they can find out which items are the best-selling and best-selling items.

**Keywords:** data mining,clustering,furniture,k-means

**Abstrak:** PT.XYZ merupakan salah satu distributor yang menjual segala jenis barang perabot rumah tangga seperti meja,kursi,meja rias,lemari baju,sofa dan masih banyak lagi,yang dimana data penjualan masih dilakukan dengan cara manual seperti kurangnya dalam peninjauan produk-produk apa saja yang dibutuhkan konsumen dan penyimpanan data-data yang kurang efektif.Untuk mengatasi permasalahan tersebut peneliti mencoba melakukan implementasi dengan menggunakan salah satu metode yang ada pada data mining yaitu Algoritma K-Means.algoritma k-menas Clustering merupakan salah satu teknik dari salah satu fungsionalitas data mining,algoritma clustering merupakan algoritma pengelompokkan sejumlah data menjadi kelompok–kelompok data tertentu (cluster). Sehingga dengan adanya pengelompokan data ini pihak perusahaan dapat mengetahui barang paling laris dan,kurang laris. Sehingga barang yang ada digudang tidak menumpuk. Dari penititan ini output yang dihasilkan yaitu, barang paling laris sebanyak 5, dan barang yang kurang laris sebanyak 5 Dengan adanya pengolahan data yang dilakukan diharapkan dapat memberikan solusi kepada Pihak perusahaan agar dapat mengetahui mana barang yang paling laris, dan laris .

**Kata kunci:** data mining,clustering,perabot,k-means

### LATAR BELAKANG

Di era globalisasi ini, kebutuhan alat rumah tangga di kalangan masyarakat semakin meningkat sehingga penjualan untuk alat perabot rumah tangga pun jugak meningkat.sehingga masyarakat memiliki tingkat penjualan yang tinggi untuk melengkapi kebutuhan di rumah masing-masing.Pentingnya peralatan rumah tangga dan potensi toko perabot sebagai tempat penjualan produk tersebut. Semakin tingginya persaingan bisnis dari Persaingan bisnis yang semakin ketat memaksa para pengusaha untuk lebih berupaya dalam penyediaan jasa khususnya dalam pengelolaan kesesuaian produk khususnya dalam proses pelayanan penjualan dan pembelian furniture rumah.adapun alat perabot rumah tangga yang tersedia yaitu :meja,kursi,meja rias,lemari baju,sofa dan masih banyak lagi.untuk memuaskan

pelanggan. Seperti kesulitan toko perabot dalam meningkatkan penjualan dan pendapatan perusahaan. Oleh karena itu peneliti mencoba melakukan penerapan algoritma data mining dengan menggunakan metode k-means clustering.

Data mining sendiri merupakan salah satu kegiatan penambangan data dengan mengekstraksi pola yang memerlukan data dalam jumlah besar, dan strategi menjadi menarik apabila tidak sederhana, implisit, dan tidak diketahui sebelumnya, serta strategi yang dihasilkan harus mudah dipahami dan berguna digunakan untuk data yang diprediksi dengan tingkat kepastian tertentu. (Cv & Konstruksi, 2021).

Clustering adalah teknik dalam kemampuan data mining dan merupakan algoritma untuk mengelompokkan sekumpulan data ke dalam kelompok data (cluster) tertentu. Di sisi lain, penambangan data sering disebut sebagai penemuan pengetahuan dalam basis data (KDD), dan merupakan aktivitas yang mengumpulkan dan menggunakan data historis untuk menemukan keteraturan, pola, atau hubungan dalam sejumlah besar data (. et al., 2020). Clustering juga merupakan algoritma analisis data yang bertujuan untuk mengelompokkan data dengan karakteristik yang sama dalam wilayah yang sama. Salah satu pendekatan yang digunakan untuk mengembangkan metode clustering adalah K-means. Teknik ini merupakan teknik pengelompokan data non-hierarki (kendala) yang berupaya mengelompokkan data menjadi dua kelompok atau lebih (cluster) (Darmi & Setiawan, 2016).

Penelitian ini dilakukan untuk menentukan produk penjualan yang termasuk kedalam kelompok barang paling laris dan barang yang tidak laris. Yang dimana data tersebut akan dapat menjadi acuan bagi pemilik usaha untuk mempresiksi strategi pemasaran untuk kedepannya dan juga penelitian ini melakukan analisa dengan menggunakan platform weka untuk menentukan tingkat keakuratan hasil akhir pengelompokkan penjualan mana yang paling laris dan tidak laris.

## **KAJIAN TEORITIS**

### **Data Mining**

Data mining merupakan salah satu proses yang dilakukan untuk penambangan data yang dimana didalam data mining sendiri mampu membaca data atau menambang data dalam jumlah yang besar untuk mencapai sebuah keputusan atau untuk menghasilkan pengetahuan yang baru. Data mining juga merupakan kombinasi informasi dari database besar, menggabungkan pola, statistik, database, dan visualisasi informasi (Perdana et al., 2023). Data mining merupakan bidang dari beberapa ilmu yang menggabungkan metode statistik dan matematika untuk menghasilkan informasi penting dari sejumlah besar data, sehingga

menghasilkan informasi dan pengetahuan yang menjadi dasar pengambilan keputusan (Clustering, 2023). Data mining juga merupakan proses mencari pola pada data. Pola-pola ini berasal dari berbagai jenis database, termasuk database relasional, gudang data, data transaksional, dan data berorientasi objek. Penambahan data memungkinkan pebisnis mengambil keputusan dengan cepat dan akurat (Nurajizah & Salbinda, 2021).

### **Clustering**

Analisis kluster merupakan analisis yang mengelompokkan objek ke dalam kelompok kelompok yang relatif serupa. Cluster yang terbentuk cenderung berisi objek serupa, sedangkan cluster lainnya sangat bervariasi. Berbeda dengan klasifikasi, analisis cluster tidak menemukan variabel target. Variabel target tidak diklasifikasikan, diprediksi, atau diprediksi secara clustering (Widodo & Indonesia, 2022). Cluster adalah sekelompok objek data yang serupa satu sama lain dalam cluster yang sama tetapi berbeda dengan objek dalam cluster yang berbeda. Objek-objek dikelompokkan menjadi satu atau lebih cluster sehingga objek-objek yang berada dalam suatu cluster mempunyai tingkat kemiripan yang tinggi satu sama lain (Darmi & Setiawan, 2016).

### **Algoritma K-MEANS**

Algoritma K-Means merupakan salah satu algoritma yang termasuk dalam data mining yang menggunakan teknik clustering yang menggunakan perhitungan jarak untuk mengelompokkan data ke dalam cluster. Algoritma K-Means membagi data menjadi kelompok, data dengan karakteristik serupa tetap berada dalam kelompok yang sama, dan data yang kontras diklasifikasikan ke dalam kelompok berbeda (Aminah & Susanti, 2023). K-means adalah teknik pengelompokan data non-hierarki yang berupaya membagi data yang ada menjadi dua kelompok atau lebih (Annur, 2019). Cara ini hanya bisa digunakan untuk atribut yang sifatnya numerik atau numerik. Algoritma ini juga dapat meminimalkan jarak antara data dan clusternya. Langkah-langkah algoritma K-Means adalah (Nurzaman & Sari, 2023):

1. Tentukan jumlah k yang akan digunakan sebagai cluster yang akan di bentuk
2. Inisialisasi k pusat cluster ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat cluster diberi nilai awal dengan angka-angka random.
3. Tempatkan semua data/objek ke cluster terdekat. kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut.

Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat cluster.

Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam clustermana. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:

$$D(i, j) = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2}$$

$D(i, j)$  = Jarak data ke i ke pusat cluster j

$X_{ki}$  = Data ke i pada atribut data ke k

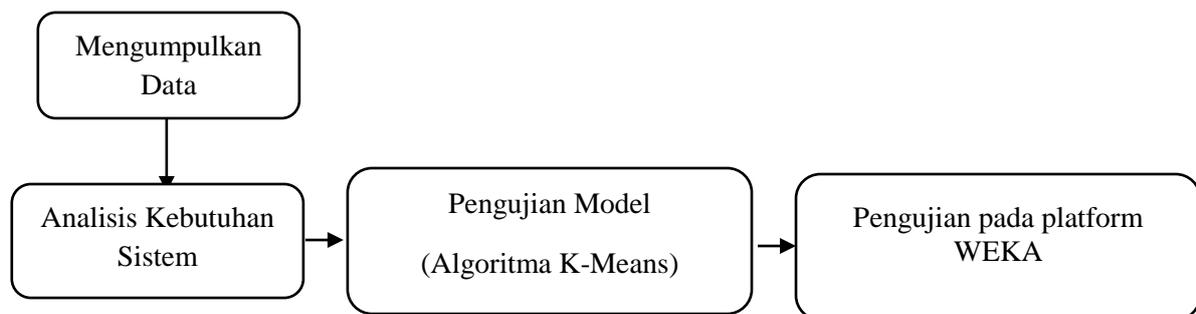
$X_{kj}$  = Data ke j pada atribut data ke k

Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata-rata dari semua data/ objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.

Tugaskan lagi setiap objek memakai pusat cluster yang baru. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat cluster tidak berubah lagi.

## METODE PENELITIAN

Metodologi merupakan serangkaian proses atau langkah yang harus dilakukan sebelum menjalankan sebuah penelitian. Adapun alur penelitian ini dapat dilihat pada Gambar 1. Yang ada dibawah ini untuk menjelaskan alur penelitian sebagai berikut:



**Gambar 1.** Kerangka kerja penelitian

### Pengumpulan Data

Penelitian diawali dengan penulis mengumpulkan beberapa referensi untuk mencari data dan informasi mengenai penelitian yang sedang diteliti. Referensi tersebut dapat berupa buku, jurnal, makalah penelitian yang berkaitan dengan pengembangan teknik clustering dan algoritma K-means, atau artikel yang berkaitan dengan kasus yang sama dengan kasus penelitian ini.

## Analisa Kebutuhan

Pada langkah ini penulis akan melakukan analisa dan perancangan terhadap sistem agar aplikasi yang akan digunakan sesuai dengan hasil yang diinginkan dan dapat mempermudah proses analisa data dalam menentukan barang yang laris dan tidak laris secara terstruktur.

## Pengujian Model (Algoritma K-means)

Pengujian model dengan menggunakan Algoritma K-Means yang menggunakan teknik clustering yang menggunakan perhitungan jarak untuk mengelompokkan data ke dalam cluster. Cara ini hanya bisa digunakan untuk atribut yang sifatnya numerik atau numerik.

## Pengujian pada platform WEKA

Pada langkah ini penulis menggunakan platform yaitu weka untuk mengimplementasikan metode K-means. Cara ini digunakan untuk membuat grup atau cluster penjualan yang berisi produk terlaris dan tidak terjual.

## HASIL DAN PEMBAHASAN

Penelitian ini didapat dari toko perabot xyz yang dimana data yang digunakan adalah data penjualan sebanyak 10 sample data. Variabel atau atribut digunakan dalam pengelompokan data penjualan ini terdiri dari Stok Awal, dan Terjual dan kemudian akan diolah dengan platform weka, seperti yang terlihat pada Gambar 2:

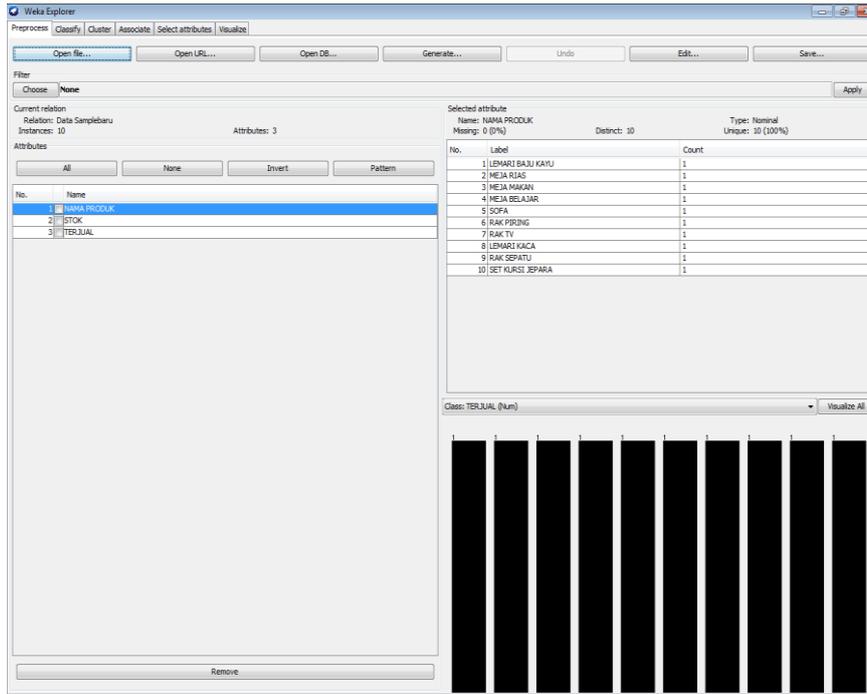


**Gambar 2.** Tampilan Weka

*Sumber: Silvia Lestari, 2024*

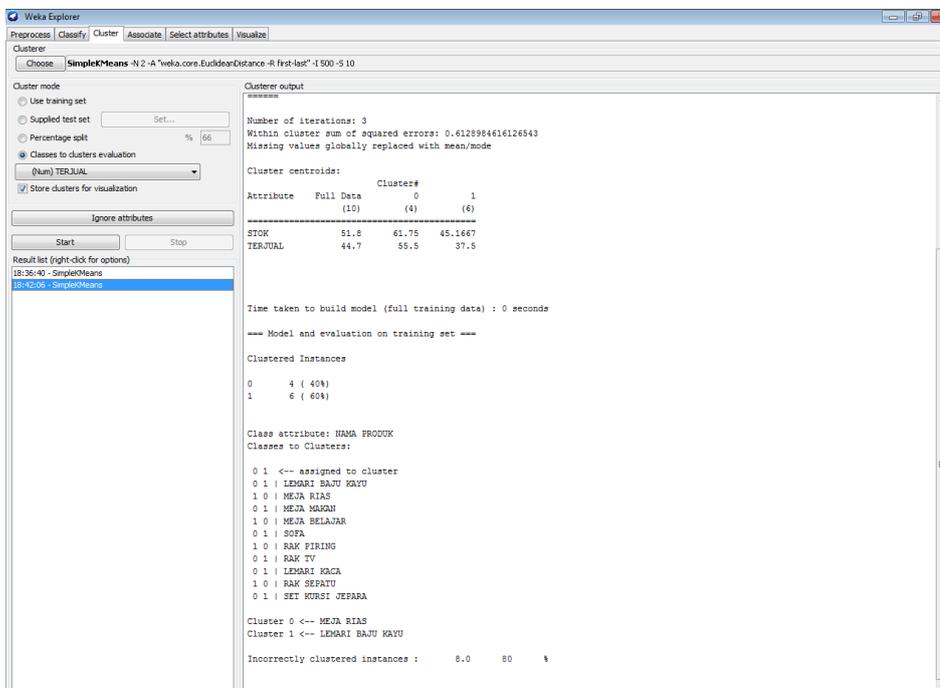
Setelah weka dibuka lalu pilih menu explorer untuk menginput data penjualan yang akan di analisis cluster terlaris dan tidak laris dari data penjualan perabot rumah tangga yang dapat dilihat pada gambar 3 berikut ini:

**IMPLEMENTASI DATA MINING DALAM MENENTUKAN PENJUALAN ALAT PERABOT DENGAN MENGGUNAKAN METODE K-MEANS CLUSTERING PADA PT.XYZ**



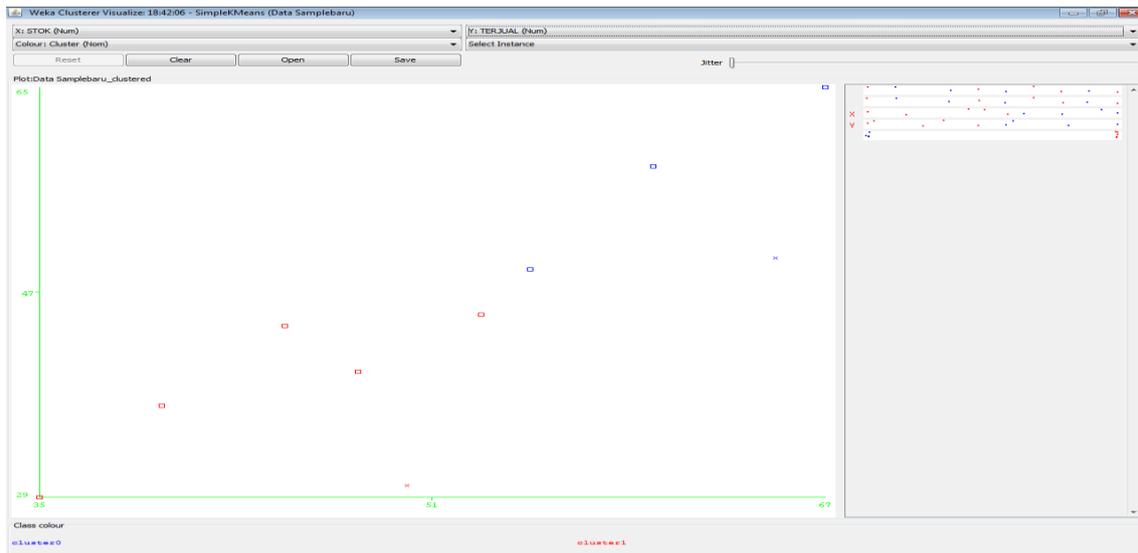
**Gambar 3.** Inputan data penjualan  
*Sumber: Silvia Lestari, 2024*

Dari gambar 3 kita dapat melihat setiap data menunjukkan masing-masing grafik penjualannya yang berasal dari data penjualan, kemudian pilih menu cluster untuk menghitung dengan algoritma k-means dan untuk menghitung nilai dari kedua cluster yang dapat dilihat pada Gambar 4, sebagai berikut :



**Gambar 4.** Tampilan hasil cluster  
*Sumber: Silvia Lestari, 2024*

Dari gambar diatas dapat kita lihat cluster yang dihasilkan memiliki 2 cluster yaitu dengan nilai  $C1 =$  Meja Rias dengan nilai centroid sebesar 61.75 pada Stok kemudian penjualan sebesar =55.5 dan  $C2=$ lemari Baju Kayu dengan nilai centroid 45.1667 pada stok kemudian penjualan sebesar=37.5,yang mengartikan bahwa nilai cluster pertama dijadikan barang terlaris dan cluster kedua sebagai barang yang tidak laris, yang dapat dilihat bentuk visual modelnya pada Gambar 5, sebagai berikut :



**Gambar 5.** Visualize cluster assignments  
*Sumber: Silvia Lestari, 2024*

## KESIMPULAN DAN SARAN

Kesimpulan dari penelitian ini menunjukkan bahwa PT.XYZ, distributor perabot rumah tangga, menghadapi tantangan dalam pengelolaan data penjualan yang masih dilakukan secara manual. Kurangnya tinjauan produk yang dibutuhkan oleh konsumen dan penyimpanan data yang kurang efektif menjadi masalah utama. Untuk mengatasi permasalahan tersebut, peneliti telah mencoba menerapkan algoritma K-Means, sebuah metode dalam data mining, untuk mengelompokkan data penjualan menjadi kelompok-kelompok tertentu atau cluster. Tujuannya adalah untuk mengidentifikasi barang-barang yang paling laris dan kurang laris, sehingga manajemen persediaan dapat dioptimalkan. Hasil dari penelitian ini menunjukkan bahwa terdapat 5 barang yang paling laris dan 5 barang yang kurang laris. Dengan demikian, pengolahan data yang dilakukan diharapkan dapat memberikan solusi kepada perusahaan dalam mengelola stok barang dengan lebih efisien, memastikan bahwa barang yang paling laris tersedia dengan cukup, sementara mengurangi stok barang yang kurang laris yang dapat mengakibatkan penumpukan di gudang. Hal ini diharapkan dapat meningkatkan efisiensi dan produktivitas operasional perusahaan serta

meningkatkan kepuasan pelanggan dengan memastikan ketersediaan barang yang sesuai dengan kebutuhan mereka.

## DAFTAR REFERENSI

- A. F., Kesuma, F. T., & Tamba, S. P. (2020). Penerapan Data Mining Untuk Menentukan Penjualan Sparepart Toyota Dengan Metode K-Means Clustering. *Jurnal Sistem Informasi Dan Ilmu Komputer Prima(JUSIKOM PRIMA)*, 2(2), 67–72. <https://doi.org/10.34012/jusikom.v2i2.376>
- Aminah, S., & Susanti, T. (2023). *Implementasi Algoritma K-Means Clustering Penerima Bantuan Beasiswa UKT Pada Institut Teknologi Pagar Alam*. 9(1).
- Annur, H. (2019). *Penerapan Data Mining Menentukan Strategi Penjualan Variasi Mobil Menggunakan Metode K-Means Clustering ( Studi Kasus Toko Luxor Variasi Gorontalo )*. 5(1).
- Clustering, M. M. K. (2023). *Implementasi Data Mining Dalam Mengelompokkan Data penduduk Kurang Mampu*. 2(November), 945–955.
- Cv, D. I., & Konstruksi, I. (2021). *ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI PEMASARAN*. 3(1), 30–36.
- Darmi, Y., & Setiawan, A. (2016). *PENERAPAN METODE CLUSTERING K-MEANS DALAM*. 12(2), 148–157.
- Nurajizah, S., & Salbinda, A. (2021). *Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten*. 7(2), 158–163. <https://doi.org/10.31294/jtk.v4i2>
- Nurzaman, M. Y., & Sari, B. N. (2023). *Implementasi K-Means Clustering Dalam Pengelompokan Banyaknya Jumlah Petani Berdasarkan Kecamatan Di Provinsi Jawa Barat*. 10(3).
- Perdana, R. W., Meri, R., Studi, P., Komputer, T., Studi, P., Informatika, M., Solok, K., & Barat, S. (2023). *IMPLEMENTASI DATA MINING PADA PENJUALAN SEPRAI MENGGUNAKAN Abstrak*. 7(1), 144–154.
- Widodo, E., & Indonesia, U. I. (2022). *Implementasi Clustering K-Medoids dalam Pengelompokan Kabupaten*. 5(2), 121–130.