

Implementasi Metode *K-Means Clustering* Untuk Menentukan Persediaan Barang Pada Toko SS BabyShop

Andre Prayoga Rizkyandri¹, Dr. Jasmir, S.Kom, M.Kom², Yulia Arvita, S.Kom, M.S.I³

^{1,2,3} Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Dinamika Bangsa, Jambi,

Indonesia Email: ¹andrerizkyandri@gmail.com, ²jasmir.skripsi@gmail.com, ³yulia_arvita@yahoo.co.id

Artikel Info :

Artikel History :

Submitted : 01-08-2023

Accepted : 14-08-2023

Published : 30-09-2023

Kata Kunci: Metode *K-Means*, Data Mining, Persediaan, *Clustering*, SPSS

Keywords: K-Means Method, Data Mining, Inventory, Clustering, SPSS

Abstrak– SS BABYSHOP merupakan salah satu toko usaha yang bergerak di bidang penjualan pakaian dan keperluan sehari-hari di jambi, menentukan stok yang dilakukan secara manual dan tidak akurat akan menyebabkan biaya simpan yang tinggi, karena bisa terjadi kekosongan atau kelebihan produk tertentu. Hal ini tentu akan sangat merugikan semua pelaku usaha seperti halnya pada toko SS BABYSHOP. Penelitian ini bertujuan untuk mengelompokkan produk yang di jual pada toko SSBABYSHOP menjadi 5 cluster guna untuk mengetahui item barang apa saja yang terjual dengan jumlah paling banyak, terjual dengan jumlah banyak, terjual dengan jumlah sedang, terjual dengan jumlah sedikit, dan terjual dengan jumlah paling sedikit, sehingga stok yang akan disediakan berdasarkan kelompok item barang pada cluster masing-masing yang telah ditentukan. Metode yang digunakan pada penelitian ini adalah *K-Means Clustering* mencari partisi yang optimal dari data dengan meminimalkan kriteria jumlah kesalahan kuadrat dengan prosedur iterasi yang optimal. Variabel yang digunakan adalah jumlah transaksi, jumlah stok dan rata-rata penjualan.

Abstract– SS BABYSHOP is one of the business shops engaged in selling clothing and daily necessities in Jambi, determining stock which is done manually and inaccurately will cause high storage costs, because there can be vacancies or excess of certain products. This will certainly be very detrimental to all business actors as is the case with the SS BABYSHOP store. This study aims to classify the products sold at the SS BABYSHOP store into 5 clusters in order to find out what items were sold in the most quantities, sold in large quantities, sold in moderate quantities, sold in small quantities, and sold in the highest quantities. few, so that the stock to be provided is based on the group of items in each predetermined cluster. The method used in this study is K-Means Clustering to find the optimal partition of the data by minimizing the criterion of the number of squared errors with the optimal iteration procedure. The variables used are the number of transactions, total stock and average sales.

1. PENDAHULUAN

Dunia usaha adalah dunia yang terus berkembang dari waktu ke waktu. Setiap individu yang menjalankan usaha, selalu mencari jalan untuk memperoleh sesuatu yang lebih menguntungkan dari sebelumnya. Toko SS Baby Shop adalah salah satu usaha yang bergerak dibidang penjualan pakaian dewasa, pakaian anak anak dan keperluan sehari-hari. Sebagai seorang pemilik usaha, tentunya harus memiliki strategi dalam mengelola persediaan barang agar tidak mengalami kerugian akibat barang yang menumpuk. Namun permasalahannya, persediaan barang pada Toko SS Baby Shop masih dilakukan secara manual, yaitu pemilik hanya melihat barang yang habis saja tanpa mempertimbangkan barang mana yang lebih laku terjual dan barang mana yang lakunya lambat. Padahal mengatur ketersediaan barang sesuai permintaan dan kebutuhan konsumen dapat mengurangi nilai kerugian dari menumpuknya stok barang tidak laku terjual.

Sehingga sebagai manajemen yang baik dalam proses mengatur persediaan stok barang sangat diperlukan, untuk menghindari penumpukan barang yang sama dan barang yang kurang diminati oleh konsumen, Untuk mengatasi permasalahan tersebut maka digunakan penerapan *K-Means Clustering*, karena permasalahan yang sering terjadi pada toko mengalami kesulitan pada saat menentukan stok minimum pada tiap barang yang harus dipenuhi berdasarkan keinginan konsumen. Cara mengatasi permasalahan yang dialami saat ini, maka dibutuhkan sebuah metode sehingga dapat menentukan barang mana yang harus di stok banyak, sedang atau bahkan sedikit agar tidak lagi mengalami kekurangan atau bahkan kelebihan dalam pemenuhan stok barang tertentu, metode yang digunakan pada penelitian ini adalah metode *K-Means clustering*

Menurut Benny Ginting dan Fristi Riandri [1] “Data mining adalah suatu proses pengerukan atau pengumpulan informasi penting dari suatu data yang besar. Proses data mining seringkali menggunakan metode statistika, matematika, hingga memanfaatkan teknologi artificial intelligence”. Muhammad Hilman Fakhri dan Khaerul Umam Fakh [2] menyatakan bahwa “sedangkan Data mining, sering juga disebut sebagai knowledge discovery in database (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar”. Menurut Fitri Yunita [3] “Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik

atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan". Deki Hari Nusti¹, Indra Kanedi dan Eko Prasetyo Rohmawan [4] "K-means merupakan suatu algoritma yang digunakan dalam pengelompokan secara partisi yang memisahkan data ke dalam kelompok yang berbeda-beda. Algoritma ini mampu meminimalkan jarak antara data ke cluster-nya. Pada dasarnya penggunaan algoritma ini dalam proses clustering tergantung pada data yang didapatkan dan konklusi yang ingin dicapai di akhir proses".

Yuyun Novianti [5] menyatakan bahwa "K-Means merupakan algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah K cluster yang sudah ditetapkan diawal". Sabrina Aulia Rahmah [6] menyatakan bahwa "Metode K-means merupakan suatu metode yang dapat melakukan pengelompokan data dalam jumlah yang cukup besar dengan perhitungan waktu yang relatif cepat dan efisien". Menurut Witten et al. [7] "Metode K-Means adalah metode yang termasuk dalam algoritma clustering berbasis jarak yang membagi data ke dalam sejumlah cluster dan algoritma ini hanya bekerja pada atribut numerik".

Menurut Hendro Priyatman, Fahmi Sajid dan Dannis Haldivany [8] "Clustering atau analisis cluster adalah proses pembentukan kelompok data (cluster) dari himpunan data yang tidak diketahui kelompok-kelompok atau kelas-kelasnya dan proses menentukan data-data termasuk ke dalam cluster yang mana. Netri Elizawati dan Lido Sabda Lesmana [9] menyatakan bahwa "Clustering yaitu proses pembentukan kelompok data dari himpunan data yang tidak diketahui kelas atau kelompoknya, dengan kata lain Clustering proses untuk mengetahui kelas-kelas atau taksonomi atau analisis topologi dari data-data yang ada".

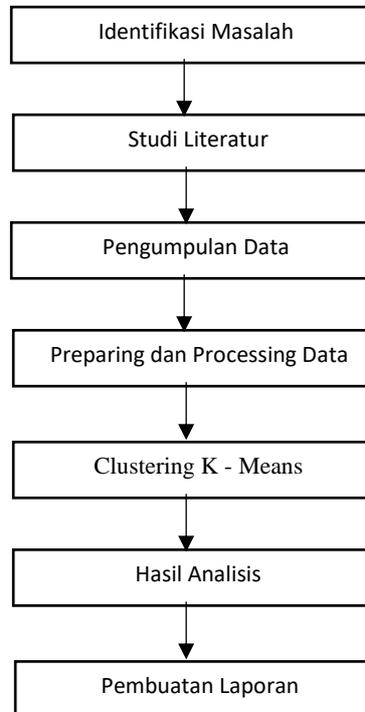
Prayogo Putra Tjaya, Rino dan Hartana Wijaya [10] menyatakan "Clustering merupakan proses pengelompokan sekumpulan objek ke dalam beberapa kelas. Tujuannya adalah mengelompokkan objek data menjadi cluster terpisah sehingga tiap objek pada cluster yang sama memiliki kemiripan karakteristik, sedangkan objek yang memiliki karakteristik berbeda akan ditempatkan pada cluster yang lain. Objek-objek dalam suatu cluster yang sama harus semirip mungkin satu sama lain dan objek-objek dalam suatu cluster harus berbeda dengan objek-objek dalam cluster lain".

Pada penelitian sejenis yaitu R. Andrean, S. Fendy, and A. Nugroho dengan hasil penelitian disimpulkan bahwa Dari hasil klasterisasi, perusahaan dapat melakukan pengendalian persediaan berdasarkan kelompok produk seperti penjelasan hasil dan pembahasan di atas yaitu produk klaster 0 dapat digunakan sebagai persediaan dengan jumlah sedikit, sedangkan kelompok produk pada klaster 1 digunakan sebagai persediaan dengan jumlah banyak, dan kelompok produk klaster 2 dapat digunakan sebagai persediaan dengan jumlah sedang [11]. Dan menurut N. Anggraini¹, jasmir, alam jusia.pareza, disimpulkan bahwa 1 jenis data menghasilkan sebuah hasil (output) yang berbeda-beda pula seperti yang telah dijelaskan sebelumnya. Dari hasil (output) yang telah diperoleh dari proses penerapan metode k-means clustering, maka hasil dari analisis dan evaluasi yang di dapat berupa sebuah kelompok data yang telah ter-cluster dari masing-masing kelompok, dan hasil (output) tersebut dapat dijadikan sebuah informasi yang berguna bagi pemilik Toko Pemsart jambi [12].

H. Sumarno, H. Priyono, and N. Hikmah dari hasil penelitian nya yaitu bahwa Proses pengelompokkan data pembelian tinta dengan menggunakan attribut barang masuk, barang keluar dan stok barang dari jumlah sampel sebanyak 42 dapat dilakukan perhitungan dari iterasi pertama sampai dengan iterasi kedua karena posisi cluster terdekat sudah sama. Dari hasil iterasi ke-2 dapat dikelompokkan barang yang paling banyak keluar pada Cluster 2 (C2) yang terdiri dari 3 barang, untuk barang yang keluarnya sedang ada pada Cluster 1 (C1) dan untuk barang yang sedikit keluarnya ada pada Cluster 3 (C3). Metode K-Means dapat digunakan untuk memudahkan bagian pembelian dalam menentukan pengelompokkan data pembelian tinta dengan tepat dan cepat [13]. A. Sani, J. Teknologika, and E. Muningsih dari hasil penelitian nya bahwa Nilai centroid ditentukan dengan membagi menjadi 4 kelompok dimana C1 merupakan level transaksi tertinggi, C2 merupakan level transaksi terendah, C3 dan C4 merupakan level transaksi rata rata 1 dan rata rata 2. Dari hasil clustering didapatkan untuk nilai C1 berjumlah 10 kode artikel, nilai untuk C2 berjumlah 4 kode artikel, nilai untuk C3 berjumlah 15 artikel dan untuk C4 berjumlah 8 artikel [14]. I. Nasution, A. Perdana Windarto, and M. Fauzan dari hasil penelitian disimpulkan bahwa jumlah record yang digunakanebanyak 34 provinsi dengan menghasilkan 2 cluster yakin cluster tinggi sebanyak 8 provinsi dan cluster rendahsebanyak 26 provinsi. Berdasarkan hasil pengujian k-means untuk kasus presentase data penduduk miskin menggunakan tools RapidMiner versi 5.3 diperoleh hasil yang sama dengan analisis perhitungan algoritma dimana diperoleh 8 provinsi dengan cluster tinggi yang menjadi pusat perhatian bagi pemerintah dalam melakukan sosialisasi dan pemetaan dalam pemberian bantuan pada provinsi tersebut [15].

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian



1. Identifikasi Masalah

Dalam hal ini, diharapkan dapat menghasilkan analisis permasalahan yang ada, yaitu bagaimana menentukan stok barang pada toko SS Baby Shop sesuai penilaian dan data yang diperoleh, agar tidak terjadi kesalahan dalam menentukan persediaan barang dan mana saja yang harus di stok ulang. Berdasarkan masalah diatas penulis mencari solusi dari permasalahan tersebut.

2. Studi Literatur

Studi literatur merupakan suatu kegiatan mencari referensi teori yang relevan dengan kasus atau permasalahan yang ditemukan. Pada tahapan ini dilakukan pencarian landasan-landasan teori yang diperoleh dari berbagai buku dan juga internet untuk melengkapi perbendaharaan konsep dan teori, sehingga memiliki landasan dan keilmuan yang baik dan sesuai.

3. Pengumpulan Data

Pada tahap ini penulis melakukan pengumpulan data dan informasi yang dibutuhkan dalam penelitian. Pengumpulan data ini dilakukan dengan menggunakan beberapa metode yaitu:

a. Pengamatan (*Observasi*)

Metode ini dilakukan dengan cara melakukan pengamatan secara langsung pada objek yang diteliti. Penulis melakukan pengamatan pada data-data penjualan pada toko SS Baby Shop KotaJambi, sehingga penulis memperoleh *softcopy* data-data yang dibutuhkan dalam proses penelitian ini.

b. Wawancara

Pada metode ini penulis melakukan wawancara singkat secara langsung dengan pihak yang

berwenang di toko SS Baby Shop Kota Jambi untuk memperoleh data dan informasi yang lebih akurat mengenai data-data Penjualan.

c. Penelitian Kepustakaan (*Library Research*)

Untuk menunjang penelitian yang akan dilakukan, penulis melakukan *library research* yakni dengan mencari data-data dari buku maupun jurnal penelitian sejenis yang berhubungan dengan metode yang penulis gunakan dalam penelitian ini adalah metode k-means.

4. *Preparing dan Processing Data*

Pada tahap ini penulis melakukan persiapan untuk data yang akan di proses. Lalu pada tahap *processing* penulis akan melakukan proses cleaning pada data penjualan karena tidak semua data penjualan akan sesuai dengan ketentuan. Lalu data akan diubah ke bentuk format csv agar bisa dilakukan proses perhitungan menggunakan aplikasi SPSS.

5. Clustering K-Means

Langkah-Langkah algoritma Clustering K-Means adalah sebagai berikut:

A. Langkah 1: Tentukan jumlah clustering yang diinginkan

B. Langkah 2: Pilih centroid awal secara acak. Pada langkah ini secara acak akan dipilih 3 buah data karyawan sebagai centroid.

C. Langkah 3: Hitung jarak dengan *centroid*.

Pada langkah ini setiap data penjualan akan ditentukan *centroid* terdekatnya, dan data tersebut akan ditetapkan sebagai anggota kelompok yang terdekat dengan *centroid*. Buat tabel hasil perhitungan jarak selengkapnya antara masing-masing data penjualan dengan *centroid*, maka di dapatkan keanggotaan dari masing-masing cluster, lalu hitung nilai *bcv*, *wes* & Rasio.

D. Langkah 4: pembaruan *centroid* dengan menghitung rata-rata nilai pada masing-masing *cluster*.

Setelah menghitung rata-rata nilai pada masing masing *cluster* didapatkan centroid baru.

E. Langkah 5: (iterasi ke 2) kembali kelangkah ke 3, jika masih ada data yang berpindah cluster dan nilairasio lebih besar dibandingkan nilai rasio pada iterasi sebelumnya, namun jika tidak ada data yang

berpindah cluster dan nilai rasio lebih kecil dibandingkan nilai rasio pada iterasi sebelumnya maka iterasi akan dihentikan.

6. Hasil Analisis

Dari hasil analisis akan dinilai keakuratan data model yang sudah di dapatkan dengan perhitungan metode Clustering dengan menggunakan algoritma K-Means.

7. Penulis Laporan

Setelah semua tahapan penelitian dilakukan, maka akan dibuat laporan sebagai dokumentasi penelitian agar dapat di manfaatkan pada waktu yang akan datang. Baik oleh penelitian sendiri maupun penelitian lainnya.

2.2 Langkah-Langkah K-Means

Langkah-Langkah algoritma K-Means menurut Tia Noviana et al. [16]

1. Tentukan K sebagai jumlah cluster yang akan dibentuk

2. Tentukan k Centroid (titik pusat cluster) awal secara random/acak.

$$V = \sum_{i=1}^n X ; i = 1, 2, 3, \dots n$$

Dimana; v : centroid pada cluster

X_i : objek ke -i

n : banyak objek/jumlah objek yang menjadi anggota cluster.

3. Hitung jarak setiap objek ke masing-masing centroid di masing-masing cluster. Untuk menghitung jarak antara objek dengan centroid dapat menggunakan Euclidian Distance

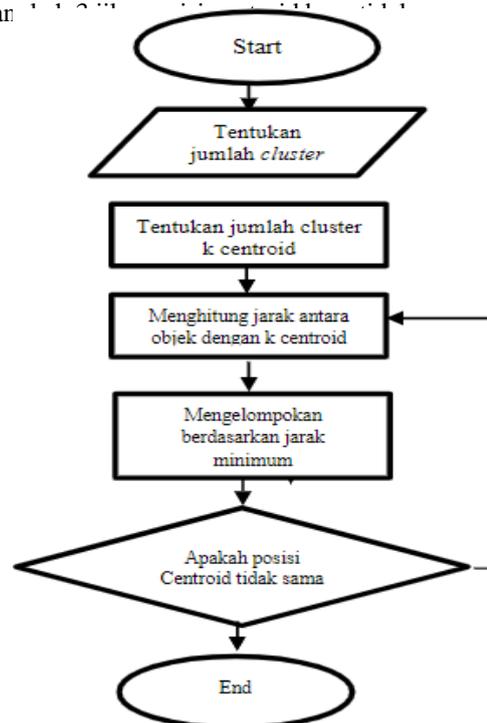
$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} ; 1, 2, 3, \dots, n$$

Dimana; x_i : objek x ke-i

y_i : daya y ke-i

n : banyaknya objek

4. Alokasikan masing-masing objek ke dalam centroid yang paling dekat.
5. Lakukan iterasi kemudian tentukan posisi centroid baru dengan menggunakan persamaan.
6. Ulangi langkah 4 dan 5 sampai posisi centroid tidak berubah.



Gambar 2.2 Flowchart Algoritma K-Means [16]

3. HASIL DAN PEMBAHASAN

Penulis melakukan penelitian terhadap Toko SS BABYSHOP ini guna memberikan solusi dari permasalahan yang ada dengan menggunakan teknik data mining, melakukan sebuah proses data cleaning dan penggabungan dari 3 jenis data yang ada. Data yang akan digunakan untuk diolah pada penelitian ini meliputi data penjualan barang, data persediaan barang, dan data transaksi yang ada pada Toko SS BABYSHOP.

Sehingga nantinya hasil dari olahan data-data tersebut akan menghasilkan output yang mana bisa dijadikan sebuah informasi yang sangat berguna bagi pemilik toko, informasi tersebut berupa kelompok data yang telah ter-cluster menjadi 5 bagian yaitudapat diketahui item barang apa saja yang terjual dengan jumlah paling banyak, terjual dengan jumlah banyak, terjual dengan jumlah sedang, terjual dengan jumlah sedikit dan terjual dengan jumlah paling sedikit. Kemudian pemilik Toko SS BABYSHOP pun dapat melakukan sebuah proses penyetokan barang dengan tepat, dan dapat terhindar dari proses penyetokan barang yang sia-sia.

4. KESIMPULAN

Penelitian ini menggunakan data penjualan yang diambil pada tokoh SS baby shop dengan jumlah data yang digunakan 203 dan mempunyai 3 atribut yaitu jumlah transaksi jumlah stok dan rata-rata penjualan. Metode yang digunakan dalam penelitian ini adalah K-Means Clustering dari perhitungan yang telah dilakukan maka direkomendasikan penulis 5 klaster untuk menentukan persediaan barang yaitu Cluster 1 yang dikategorikan paling banyak Cluster 2 yang dikategorikan terjual sedang Cluster 3 yang dikategorikan terjual paling sedikit kelas 4 yang di kategorikan terjual sedikit dan cluster 5 dikatagorikan terjual banyak, dan hasil perbandingan perhitungan manual dan SPSS berbeda Hal ini dikarenakan pemilihan pusat Cluster nya dipilih secara random hasil.

REFERENCES

- [1] B. Ginting and F. Riandari, "Implementasi Metode K-Means Clustering Dalam Pengelompokan Bibit Tanaman Kopi Arabika," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 3, no. 2, 2020.
- [2] M. H. Fakhriza and K. Umam, "ANALISIS PRODUK TERLARIS MENGGUNAKAN METODE K-MEANS CLUSTERING PADA 'PT.SUKANDA DJAYA,'" 2021.
- [3] F. Yunita, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING PADA PENERIMAAN MAHASISWA BARU (STUDI KASUS : UNIVERSITAS ISLAM INDRAGIRI)," 2018.
- [4] D. H. Nusti, I. Kanedi, Eko, and P. Rohmawan, "Application of K-Means Clustering Algorithm in Grouping Inventory Data at Putra Shop," *JURNAL KOMITEK*, vol. 1, no. 1, p. 1, 2021, doi: 10.53697/jkomitek.v1i1.
- [5] Yuyun Novianti, "IMPLEMENTASI K-MEANS CLUSTERING ALGORITHM UNTUK ANALISA SOAL UJIAN ONLINE PADA SMP ISLAM AL-FATH PARE," Aug. 2017.
- [6] S. Aulia, "KLASTERISASI POLA PENJUALAN PESTISIDA MENGGUNAKAN METODE K-MEANS CLUSTERING (STUDI KASUS DI TOKO JUANDA TANI KECAMATAN HUTABAYU RAJA)," 2020.
- [7] W. Mega, "CLUSTERING MENGGUNAKAN METODE K-MEANS UNTUK MENENTUKAN STATUS GIZI BALITA," 2015
- [8] H. Priyatman, F. Sajid, and D. Haldivany, "JEPIN (Jurnal Edukasi dan Penelitian Informatika) Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa," 2019.
- [9] N. Elizawati, D. Lido, and S. Lesmana, "Analisis Nilai Rapor Siswa Kelas X Jurusan Multimedia Terhadap Minat pada Pelajaran Produktif Di Kelas XII untuk Menentukan Kompetensi Siswa dengan Metode Clustering Algoritma K-Means (Studi Kasus Di SMKN 4 Padang)," 2017. [Online]. Available: <http://jurnal.pcr.ac.id>
- [10] P. Putra Tjaya and H. Wijaya, "IMPLEMENTASI METODE CLUSTERING K-MEANS UNTUK REKOMENDASI PENGADAAN STOK LAMPU DI PT GLOBAL LIGHTING INDONESIA," 2021. [Online]. Available: <https://jurnal.buddhidharma.ac.id/index.php/algor/index>
- [11] R. Andrean, S. Fendy, and A. Nugroho, "Klasterisasi Pengendalian Persediaan Aki Menggunakan

Metode K-Means,” JOINTECS) Journal of Information Technology and Computer Science, vol. 4, no. 1, pp. 2541–3619, 2019, doi: 10.31328/jo.

- [12] N. Anggraini¹, jasmir, alam jusia.pareza, “Penerapan Metode K-Means Clustering untuk Menentukan Persediaan StokBarang Pada Toko Pensmart Jambi.”
- [13] H. Sumarno, H. Priyono, and N. Hikmah, “Pengelompokan Data Pembelian Tinta Dengan Menggunakan Metode K-Means,” 2019. [Online]. Available: <http://tunasbangsa.ac.id/ejurnal/index.php/jsakti>
- [14] A. Sani, J. Teknologika, and E. Muningsih, “PENERAPAN METODE K-MEANS CLUSTERING PADA PERUSAHAAN.”
- [15] I. Nasution, A. Perdana Windarto, and M. Fauzan, “Penerapan Algoritma K-Means Dalam Pengelompokan Data PendudukMiskin Menurut Provinsi,” *Technology and Science (BITS)*, vol. 2, no. 2, pp. 76–83, 2020, [Online]. Available: <https://www.bps.go.id>.
- [16] Tia Noviana, Jasmir, and Yudi Novianto, P. Studi Teknik Informatika, S. Dinamika Bangsa, and J. Jl Jendral Sudirman Thekok -Jambi, “PENERAPAN DATA MINING MENENTUKAN KELOMPOK PRIORITAS PENERIMA BANTUAN BERAS RASTRA DENGAN CLUSTERING K-MEANS.” Dec. 2019