



Simulasi dalam Optimalisasi Pengadaan Barang menggunakan Metode K-Mean Clustering

Indah Savitri Hidayat^{1✉}, Sarjon Defit², Gunadi Widi Nurcahyo³

^{1,2,3}Universitas Putra Indonesia YPTK Padang

indahsavitri748@gmail.com

Abstract

Products provided by a store have an influence on store sales. Consumers will be attracted to stores that provide products according to their wants and needs. The purpose of this research is to find out what ornamental flower products are most in demand by consumers, in demand by consumers and less desirable to consumers. Keywords: inventory of goods, K-Mean Clustering, Data Mining, cluster, optimal. Store managers can get information about goods that have been depleted of inventory stock to be updated immediately. The method used in this study is the K-Mean Clustering method which belongs to one of the branches of Data Mining. The data used in the study is data from January 2020 to December 2020 as many as 100 pieces taken from Naafilah official shop, Padang. The data variables used in the entry of goods are the year, product name, price and amount sold. Furthermore, the data is processed using Rapid Miner software. The first stage of processing is to determine the value of clusters randomly, in this study researchers divided the cluster values into 3 groups. Next, the centroid value of each group will be determined. Centroid is derived from the minimum value, middle value and maximum value of the data provided. Then, the cluster process is calculated using the euclidean distance formula. Cluster calculations are done by calculating the closest distance to the data. The final result of this study is to find out the best-selling, best-selling and less-selling ornamental flowers, so that sellers can optimize the provision of ornamental flowers for the future.

Keywords: Inventory of Goods, K-Mean Clustering, Data Mining, Cluster, Optimal Sales.

Abstrak

Produk yang disediakan oleh suatu toko memiliki pengaruh terhadap penjualan toko. Konsumen akan tertarik pada toko yang menyediakan produk sesuai dengan keinginan dan kebutuhannya. Tujuan penelitian ini adalah untuk mengetahui produk bunga hias apa saja yang paling diminati konsumen, diminati konsumen dan kurang diminati konsumen. Pengelola toko bisa mendapatkan informasi mengenai barang yang telah habis stok persediaannya agar segera di *update*. Metode yang digunakan pada penelitian ini adalah metode K-Mean Clustering yang termasuk dalam salah satu cabang Data Mining. Data yang digunakan pada penelitian merupakan data dari Januari 2020 hingga Desember 2020 sebanyak 100 buah yang diambil dari toko Naafilah Official Shop, Padang. Variabel data yang digunakan pada pemasukan barang yaitu tahun, nama produk, harga dan jumlah terjual. Selanjutnya, data tersebut diolah menggunakan *software Rapid Miner*. Tahapan pengolahan yang pertama adalah menentukan nilai *cluster* secara acak, dalam penelitian ini peneliti membagi nilai *cluster* menjadi 3 kelompok. Selanjutnya, nilai *centroid* dari masing-masing kelompok tersebut akan ditentukan. Nilai *centroid* diperoleh dari nilai minimum, nilai tengah dan nilai maksimum dari data yang telah disediakan. Lalu, proses *cluster* dihitung menggunakan rumus *euclidean distance*. Perhitungan *cluster* dilakukan dengan menghitung jarak yang paling dekat dengan data. Hasil akhir penelitian ini adalah mengetahui bunga hias yang paling laris, laris dan kurang laris, sehingga penjual dapat mengoptimalkan penyediaan bunga hias untuk masa yang akan datang.

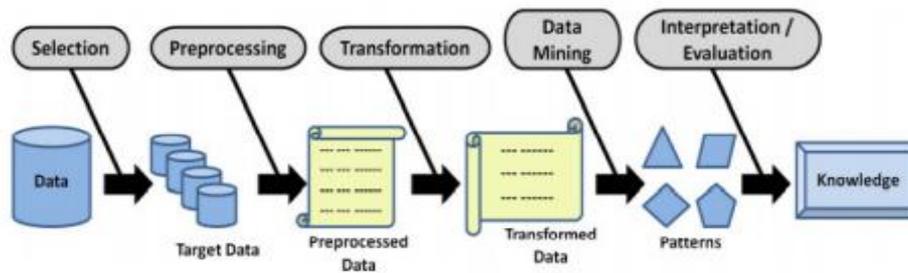
Kata Kunci : persediaan barang, K-Mean Clustering, Data Mining, *cluster*, penjualan optimal.

© 2021 JSisfotek

1. Pendahuluan

Data yang berukuran besar dapat disimpan dalam *database*. *Database* bisa diartikan juga sebagai koleksi data. Kumpulan dari *database* bisa menjadi Big Data. Big Data merupakan data yang memiliki volume dan kapasitas yang sangat besar sehingga tidak bisa diolah hanya dengan manual dan tradisional, namun butuh proses yang sedemikian rupa agar data tersebut dapat diubah menjadi informasi yang berharga dan dapat dimanfaatkan sesuai yang dibutuhkan [1]. Jadi, Big Data adalah kumpulan data yang sangat besar yang pengolahannya tidak bisa dilakukan secara manual, sehingga membutuhkan cara dan alat agar data tersebut bisa berguna. Dalam perkembangannya, terdapat

beberapa metodologi yang dapat digunakan dalam melakukan ekstraksi data salah satunya yaitu *Knowledge Discovery in Database* (KDD). KDD merupakan suatu proses persimpangan dari beberapa disiplin ilmu seperti statistik, *database*, *Artificial Intelligent* (AI), visualisasi dan komputer paralel yang mempengaruhi pengetahuan "*interdisciplinary knowledge*". KDD memiliki arti suatu proses pemilihan data, pembersihan data, transformasi data, presentasi, temuan dan evaluasi [2]. KDD adalah suatu metode yang dapat kita gunakan untuk menghasilkan pengetahuan dari *database* yang ada. Hasil pengetahuan yang telah didapatkan tersebut akan digunakan untuk basis pengetahuan atau *knowledge base* untuk pengambilan keputusan [3].

Gambar 1. Proses *Knowledge Discovery in Database* (KDD)

Menurut Gustientiedina dkk, 2019, adapun proses *Knowledge Discovery in Database* (KDD) adalah *Selection* diperlukan untuk menentukan variabel yang akan digunakan agar kesamaan dan perulangan tidak terjadi pada Data Mining; *Preprocessing* terdapat dua tahap, yang pertama *Data Cleaning*, yaitu menghilangkan data yang tidak diperlukan seperti menangani *missing value*, *noise* data dan menangani data – data yang tidak konsisten serta relevan, kedua *Data Integration*, yaitu dilakukan terhadap atribut yang mengidentifikasi entitas yang unik; *Transformation*; data dirubah sesuai format *ekstention* yang sesuai dalam pengolahan Data Mining, karena beberapa metode pada Data Mining memerlukan format khusus sebelum dilakukan pemrosesan; Data Mining; merupakan proses utama pada metode yang diterapkan untuk mendapatkan pengetahuan baru dari data yang diproses; *Evaluation/Interpretation* yang dilakukan identifikasi terhadap pola-pola yang menarik kedalam *knowledge base* yang akhirnya akan menghasilkan pola-pola khas maupun model prediksi yang dievaluasi untuk menilai kajian yang ada, apakah sudah memenuhi target yang diinginkan atau belum; *Knowledge* yaitu pengetahuan baru yang dihasilkan bisa dipahami semua orang sehingga pengetahuan tersebut bisa dijadikan sebagai acuan pengambilan keputusan dan pola-pola yang dihasilkan akan dipresentasikan kepada pengguna.

Salah satu tahapan dalam keseluruhan proses KDD adalah Data Mining [4]. Mining diartikan sebagai usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar [1]. Mining merupakan proses informasi yang dilakukan dari *database* yang dapat dilakukan meskipun tidak terlihat [5]. Data Mining adalah suatu rangkaian proses yang mengolah kumpulan data yang berlimpah untuk menggali informasi yang belum diketahui dengan cara mengenali pola-pola data dari informasi tersebut [6]. Salah satu metode yang termasuk dalam Data Mining adalah *clustering*. *Clustering* atau pengelompokan adalah alat yang digunakan dalam ilmu data [7], [8], [9]. Secara umum, *cluster* adalah sekelompok objek yang lebih mirip satu sama lain daripada anggota *cluster* lain [10]. Salah satu algoritma yang digunakan dalam *clustering* adalah K-Mean. Algoritma K-Means adalah metode iterasi sederhana untuk mempartisi set data yang

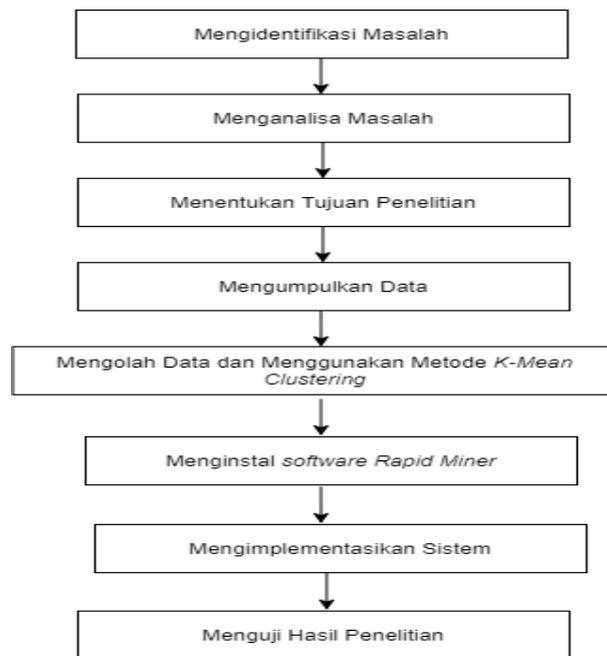
diberikan ke dalam jumlah *cluster* yang ditentukan pengguna [6]. Tujuan pengelompokan data adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, pada umumnya berusaha meminimalkan variasi suatu kelompok dan memaksimalkan variasi antar kelompok. Algoritma K-Mean *Cluster* menggunakan analisa pengklasteran dengan *Non-Hierarchical Cluster*, metode ini dimulai dari pengambilan *cluster* yang diinginkan terlebih dahulu, *cluster* yang diambil bisa saja dua *cluster*, tiga *cluster* atau *cluster* yang lain hingga terbentuknya *cluster* [11].

Usaha Mikro, Kecil dan Menengah (UMKM) memiliki pelanggan dan memiliki stok produk dalam jumlah yang banyak, namun UMKM tidak memiliki pemetaan pelanggan dan stok produk sehingga mereka tidak tahu produk apa saja yang menjadi favorit pelanggan [9]. Konsumen akan tertarik pada toko yang menyediakan produk sesuai dengan keinginan dan kebutuhannya [4].

2. Metodologi Penelitian

Data yang digunakan pada penelitian ini adalah data penjualan bunga hias yang diperoleh dari UMKM Nafiihah *Official Shop*, Padang. Pada data ini terdapat informasi yaitu hari, tanggal, bulan, tahun, nama bunga, harga, stok dan jumlah terjual.

Metodologi pengolahan data yang digunakan adalah Data Mining dengan metode K-Mean *Clustering*. Pada tahap analisa data, data yang dibutuhkan adalah data dalam bentuk *Microsoft Excel*, namun data yang didapat dari toko Nafiihah *Official Shop* adalah data dalam bentuk pembukuan sehingga perlu dilakukan perekapan ulang ke bentuk *Microsoft Excel* lalu menggunakan rumus *euclidean distance* dan *software Rapid Miner*. Setelah melakukan pengolahan dan perhitungan data, maka didapatlah hasil *cluster* yang diinginkan. Maka dilakukan penelitian ini dengan tahapan yang disajikan pada kerangka penelitian pada Gambar 2.



Gambar 2. Kerangka Kerja Penelitian

Langkah-langkah yang dilakukan untuk membentuk *clustering* pada metode K-Mean *Clustering* adalah:

- Memilih nilai *cluster* awal secara acak.
- Setiap data dalam *dataset* dibagi kedalam beberapa kelompok *cluster* antara setiap titik dan pusat *cluster* yang didapatkan berdasarkan jarak *euclidean distance*. Adapun rumus *euclidean distance* disajikan pada Persamaan (1).

$$J = \sqrt{(ni - mi)^2 + (nj - mj)^2} \quad (1)$$

Di mana J sebagai jarak data ke pusat *cluster*, n sebagai data ke- pada atribut ke-, dan m sebagai titik pusat ke-, pada atribut.

- Setiap pusat *cluster* dihitung ulang berdasarkan dari nilai rata-rata dalam *cluster* yang didapatkan. Adapun rumus untuk menentukan nilai *centroid* baru disajikan pada Persamaan (2).

$$cd = \frac{c1+cn}{dc} \quad (2)$$

Di mana Cd sebagai nilai *centorid* yang ingin ditentukan, c1 sebagai nilai *cluster* pada data 1, cn sebagai nilai *cluster* pada data ke 2, 3, 10, 50, dst dan dc sebagai jumlah data per *cluster*. Dengan kata lain, untuk menentukan nilai *centroid*, peneliti dapat mencari nilai rata-rata per *cluster*.

- Ulangi langkah 2 dan 3 sampai kelompok tersebut sesuai. Perulangan dapat didefinisikan secara berbeda tergantung pada implementasi, tetapi biasanya pada proses langkah 2 dan 3 dapat diulangi apabila kelompok *cluster* masih mengalami perubahan dan proses akan terhenti apabila tidak ada perubahan nilai antar *cluster*.

3. Hasil dan Pembahasan

3.1 Analisa Data

Jumah data yang digunakan dalam penelitian ini sebanyak 100 data. *Output* yang diinginkan dari penelitian ini adalah 3 *cluster*, yaitu barang yang paling laris, laris dan kurang laris. Atribut yang digunakan dalam perhitungan data adalah nama barang, stok dan jumlah terjual.

3.2 Perhitungan Metode K-Mean *Clustering*

Untuk melakukan *cluster* data, hal pertama yang dilakukan adalah menentukan jumlah *cluster*. *Cluster* pertama ditentukan secara acak, pada penelitian ini, peneliti menagambil dari data ke 5, 46 dan 80. Adapun *cluster* pertama yang digunakan seperti Tabel 1.

Tabel 1. *Cluster* pertama

<i>Cluster</i>	Stok	Jumlah Terjual
<i>Cluster</i> 1	10	85
<i>Cluster</i> 2	10	412
<i>Cluster</i> 3	12	199

Cluster pertama akan diolah menggunakan rumus *euclidean distance* seperti persamaan 1, lalu hitung jarak dari data 1 hingga 100. Namun, dalam perhitungan ini peneliti hanya memberikan sampel data ke 1 hingga data ke 3, sehingga didapatlah hasil *cluster* data untuk *cluster* 1, *cluster* 2 dan *cluster* 3.

- Menghitung Jarak Dari Data Ke 1 Terhadap Pusat *Cluster*

$$C1 = \sqrt{(12 - 10)^2 + (10 - 85)^2} = 75$$

$$C2 = \sqrt{(12 - 10)^2 + (10 - 412)^2} = 402$$

$$C3 = \sqrt{(12 - 12)^2 + (10 - 199)^2} = 189$$

- Menghitung Jarak Dari Data Ke 2 Terhadap Pusat *Cluster*

$$C1 = \sqrt{(10 - 10)^2 + (2 - 85)^2} = 83$$

$$C2 = \sqrt{(10 - 10)^2 + (2 - 412)^2} = 410$$

$$C3 = \sqrt{(10 - 12)^2 + (2 - 199)^2} = 197$$

- Menghitung Jarak Dari Data Ke 3 Terhadap Pusat *Cluster*

$$C1 = \sqrt{(10 - 10)^2 + (5 - 85)^2} = 80$$

$$C2 = \sqrt{(10 - 10)^2 + (5 - 412)^2} = 407$$

$$C3 = \sqrt{(10 - 12)^2 + (5 - 199)^2} = 194$$

Berdasarkan hasil perhitungan jarak data dengan pusat *cluster*, maka hasil yang didapatkan oleh masing-masing *cluster*. Hasil *cluster* dari iterasi 1 adalah:

C1= 1, 2, 3, 4, 5, 7, 8, 9,10, 11, 12, 13, 14, 15, 16, 17, 19, 22, 23, 24, 25, 26, 27, 28, 29, 31, 32, 33, 34, 35, 37, 38, 40, 41, 42, 43, 44, 45, 47, 48, 49, 50, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 69, 71, 72, 73, 75, 76, 77, 79, 81, 82, 83, 85, 87, 88, 90, 91, 92, 93, 95, 96, 97, 98, 99, 100.

C2= 6, 46, 55, 68, 94.
 C3= 18, 20, 21, 30, 36, 39, 51, 52, 70, 78, 80, 84, 86, 89.

Selanjutnya adalah menghitung nilai pusat *cluster* atau nilai *centroid* baru untuk masing-masing menggunakan rumus sesuai dengan Persamaan (2). Nilai *centroid* didapatkan dari nilai rata-rata masing-masing *cluster* yang telah dilakukan pada iterasi 1 sebelumnya. Berikut hasil nilai *centroid* seperti Tabel 2.

Tabel 2. *Cluster* Kedua

<i>Cluster</i>	Stok	Jumlah Terjual
<i>Cluster</i> 1	8,89	29,38
<i>Cluster</i> 2	11,4	513
<i>Cluster</i> 3	9,86	208

Cluster kedua diolah menggunakan rumus *euclidean distance*, lalu hitung jarak dari data 1 hingga 100. Namun, dalam perhitungan ini peneliti hanya memberikan sampel data ke 1 hingga data ke 3, sehingga didapatlah hasil *cluster* data untuk *cluster* 1, *cluster* 2 dan *cluster* 3.

a. Menghitung Jarak Dari Data Ke 1 Terhadap Pusat *Cluster*

$$C1 = \sqrt{(12 - 8,89)^2 + (10 - 29,38)^2} = 20$$

$$C2 = \sqrt{(12 - 11,4)^2 + (10 - 513)^2} = 503$$

$$C3 = \sqrt{(12 - 9,86)^2 + (10 - 208)^2} = 198$$

b. Menghitung Jarak Dari Data Ke 2 Terhadap Pusat *Cluster*

$$C1 = \sqrt{(10 - 8,89)^2 + (2 - 29,38)^2} = 27$$

$$C2 = \sqrt{(10 - 11,4)^2 + (2 - 513)^2} = 511$$

$$C3 = \sqrt{(10 - 11,4)^2 + (2 - 513)^2} = 206$$

c. Menghitung Jarak Dari Data Ke 3 Terhadap Pusat *Cluster*

$$C1 = \sqrt{(10 - 8,89)^2 + (5 - 29,38)^2} = 24$$

$$C2 = \sqrt{(10 - 11,4)^2 + (5 - 513)^2} = 508$$

$$C3 = \sqrt{(10 - 11,4)^2 + (5 - 513)^2} = 203$$

Hasil perhitungan yang didapatkan dari masing-masing *cluster* menghasilkan *cluster* dari iterasi 2 yaitu:

C1= 1, 2, 3, 4, 5, 7, 8, 9,10, 11, 13, 14, 15, 16, 17, 19, 22, 23, 24, 25, 26, 27, 28, 29, 31, 33, 34, 35, 37, 38, 40, 41, 42, 43, 44, 45, 47, 48, 49, 50, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 69, 71, 72, 73, 75, 76, 77, 79, 81, 82, 83, 85, 87, 88, 90, 91, 92, 93, 95, 96, 97, 98, 99, 100.
 C2= 46, 55, 68, 94.
 C3= 6, 12, 18, 20, 21, 30, 32, 36, 39, 51, 52, 70, 74, 78, 80, 81, 84, 86, 89, 96.

Setelah didapatkan hasil iterasi 2, peneliti dapat membandingkan hasil *cluster* iterasi 1 dan iterasi 2,

pada penelitian ini menghasilkan nilai *cluster* yang berbeda, yaitu pada data 6, 12, 32, 74, 81 dan 96. Maka, peneliti harus melakukan perulangan iterasi agar nilai *cluster* per data tidak lagi berubah. Untuk menentukan nilai *centroid* baru iterasi ketiga, cara yang digunakan sama dengan menentukan nilai *centroid* pada iterasi 2. Pada iterasi ketiga, peneliti menentukan nilai rata-rata per *cluster* dari iterasi 2. Berikut hasil nilai *centroid* seperti Tabel 3.

Tabel 5. *Cluster* Ketiga

<i>Cluster</i>	Stok	Jumlah Terjual
<i>Cluster</i> 1	11	31
<i>Cluster</i> 2	10	511
<i>Cluster</i> 3	10	202

Setiap data akan dihitung menggunakan rumus *euclidean distance* seperti persamaan 1, data dihitung dari data 1 hingga data 100.

a. Menghitung Jarak Dari Data Ke 1 Terhadap Pusat *Cluster*

$$C1 = \sqrt{(12 - 11)^2 + (10 - 31)^2} = 21$$

$$C2 = \sqrt{(12 - 10)^2 + (10 - 551)^2} = 541$$

$$C3 = \sqrt{(12 - 10)^2 + (10 - 202)^2} = 192$$

b. Menghitung Jarak Dari Data Ke 2 Terhadap Pusat *Cluster*

$$C1 = \sqrt{(10 - 11)^2 + (2 - 31)^2} = 29$$

$$C2 = \sqrt{(10 - 10)^2 + (2 - 551)^2} = 549$$

$$C3 = \sqrt{(10 - 10)^2 + (2 - 202)^2} = 200$$

c. Menghitung Jarak Dari Data Ke 3 Terhadap Pusat *Cluster*

$$C1 = \sqrt{(10 - 11)^2 + (5 - 31)^2} = 26$$

$$C2 = \sqrt{(10 - 10)^2 + (5 - 551)^2} = 546$$

$$C3 = \sqrt{(10 - 10)^2 + (5 - 202)^2} = 197$$

Hasil perhitungan *cluster* dari iterasi 3 adalah:

C1= 1, 2, 3, 4, 5, 7, 8, 9,10, 11, 13, 14, 15, 16, 17, 19, 22, 23, 24, 25, 26, 27, 28, 29, 31, 33, 34, 35, 37, 38, 40, 41, 42, 43, 44, 45, 47, 48, 49, 50, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 69, 71, 72, 73, 75, 76, 77, 79, 81, 82, 83, 85, 87, 88, 90, 91, 92, 93, 95, 96, 97, 98, 99, 100.
 C2 = 46, 55, 68, 94.
 C3 = 6, 12, 18, 20, 21, 30, 32, 36, 39, 51, 52, 70, 74, 78, 80, 81, 84, 86, 89, 96.

Iterasi 1 dan iterasi 2 menghasilkan nilai *cluster* yang sama. Maka, proses K-Mean selesai. Jadi, kesimpulan

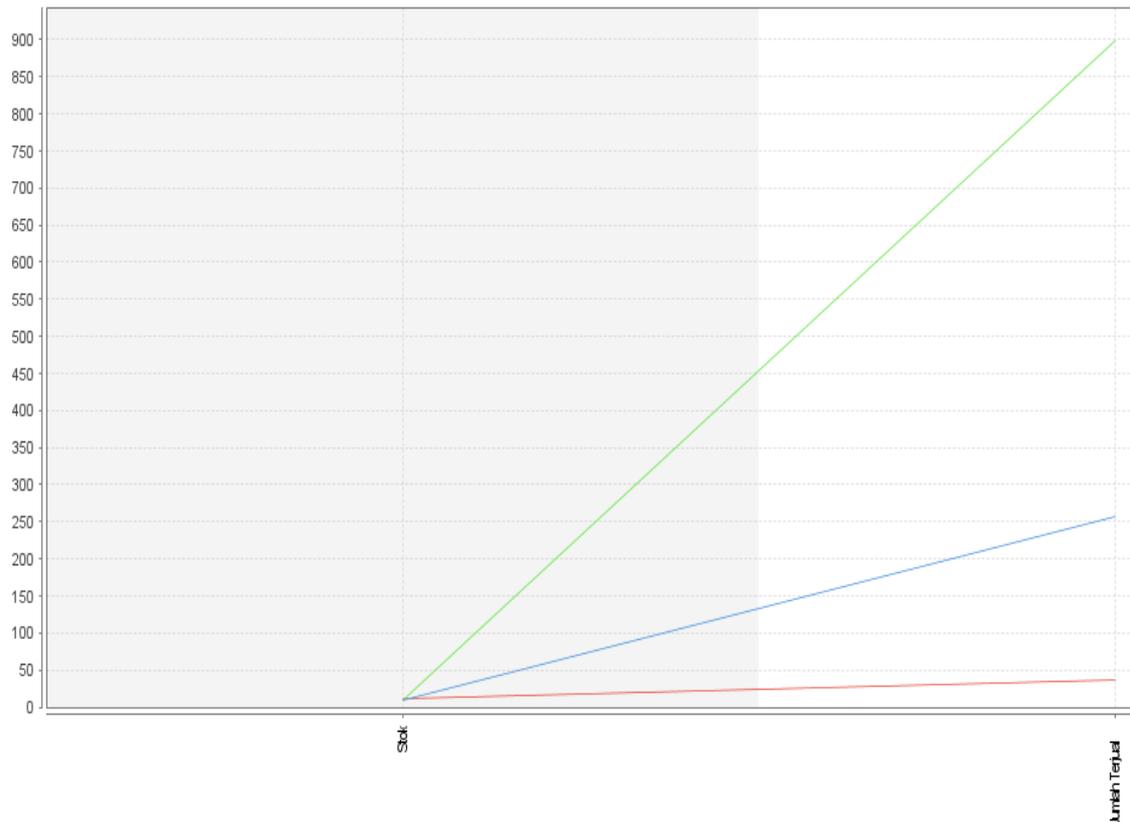
dari hasil *clustering* menggunakan *Microsoft Excel* yaitu sebagai berikut:

- Bunga hias yang paling laris adalah pada *cluster* 2, yaitu data ke 46, 55, 68 dan 94.
- Bunga hias yang laris adalah pada *cluster* 3, yaitu data ke 6, 12, 18, 20, 21, 30, 32, 36, 39, 51, 52, 70, 74, 78, 80, 81, 84, 86, 89 dan 96.
- Bunga hias yang kurang laris adalah pada *cluster* 1, yaitu data ke 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 13, 14, 15,

16, 17, 19, 22, 23, 24, 25, 26, 27, 28, 29, 31, 33, 34, 35, 37, 38, 40, 41, 42, 43, 44, 45, 47, 48, 49, 50, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 69, 71, 72, 73, 75, 76, 77, 79, 81, 82, 83, 85, 87, 88, 90, 91, 92, 93, 95, 96, 97, 98, 99, 100.

3.3. Hasil Pengujian Data.

Hasil pengujian data ditampilkan *plot* yang disajikan pada Gambar 3.



Gambar 3. Hasil Plot

Gambar 3. menguraikan bahwa garis ditarik dari arah sebelah kiri (stok) menuju sebelah kanan (jumlah terjual). Garis berwarna hijau menunjukkan *cluster* 1, garis berwarna biru menunjukkan *cluster* 0 dan garis berwarna merah menunjukkan *cluster* 2. Pada tampilan *Plot*, garis berwarna hijau disebut sebagai *cluster* 1 pada *Microsoft Excel*, garis berwarna biru disebut sebagai *cluster* 3 pada *Microsoft Excel* dan garis warna merah disebut *cluster* 2.

Hasil *clustering* adalah berikut:

- Bunga hias yang paling laris terdapat pada *cluster* 1, yaitu data ke 94.
- Bunga hias yang laris terdapat pada *cluster* 0, yaitu data ke 6, 18, 20, 21, 30, 36, 39, 46, 52, 55, 68, 70, 74, 78, 80, 84, 86 dn 89.
- Bunga hias yang kurang laris terdapat pada *cluster* 2, yaitu data ke 1, 2, 3, 4, 5, 7, 8, 10, 12, 13, 14, 15, 16,

17, 19, 22, 23, 24, 25, 26, 27, 28, 29, 31, 32, 33, 34, 35, 37, 38, 40, 41, 42, 43, 44, 45, 47, 48, 49, 50, 51, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 67, 68, 71, 72, 73, 75, 76, 77, 79, 81, 82, 83, 85, 87, 88, 90, 91, 92, 93, 95, 97, 98, 99 dan 100.

4. Kesimpulan

Metode K-Mean *Clustering* dapat digunakan untuk simulasi pengadaan penjualan produk bunga hias. Metode K-Mean *Clustering* dapat memberikan rekomendasi bunga hias yang paling laris, laris dan kurang laris, sehingga bisa dijadikan pedoman kepada pemilik toko untuk persediaan barang di masa yang akan datang.

Daftar Rujukan

- Mirantika, N. (2021). Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Covid-19 di Provinsi Jawa Barat. *Nuansa Informatika*, 15(2), 92–98. doi:10.25134/nuansa.v15i2.4321

- [2] Ependi, U., & Putra, A. (2019). Solusi Prediksi Persediaan Barang dengan Menggunakan Algoritma Apriori (Studi Kasus: Regional Part Depo Auto 2000 Palembang). *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 5(2), 139. doi:10.26418/jp.v5i2.32648
- [3] Gustientiedina, G., Adiya, M. H., & Desnelita, Y. (2019). Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 5(1), 17–24. doi:10.25077/teknosi.v5i1.2019.17-24
- [4] Kamila, I., Khairunnisa, U., & Mustakim, M. (2019). Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau. *Jurnal Ilmiah Rekayasa Dan Manajemen Sistem Informasi*, 5(1), 119. doi:10.24014/rmsi.v5i1.7381
- [5] Sinaga, K. P., & Yang, M.-S. (2020). Unsupervised K-Means Clustering Algorithm. *IEEE Access*, 8, 80716–80727. doi:10.1109/access.2020.2988796
- [6] Astuti, D. (2019). Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma K-Means Clustering. *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, 1(2), 60–72. doi:10.20895/inista.v1i2.71
- [7] Yang, M.-S., & Sinaga, K. P. (2019). A Feature-Reduction Multi-View k-Means Clustering Algorithm. *IEEE Access*, 7, 114472–114486. doi:10.1109/access.2019.2934179
- [8] Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering*, 336, 012017. doi:10.1088/1757-899x/336/1/012017
- [9] Singh, S., & Singh, P. (2020). Speaker specific feature based clustering and its applications in language independent forensic speaker recognition. *International Journal of Electrical and Computer Engineering (IJECE)*, 10(4), 3508. doi:10.11591/ijece.v10i4.pp3508-3518
- [10] Rachma, A., Aden, A., & Rusdiana, Y. (2019). Analisis Cluster Menggunakan Algoritma K-Means Cluster Untuk Culstering Jenis Penyakit Menular Pada Puskesmas Di Kecamatan Kota Tangerang. *Jurnal Sainika Unpam: Jurnal Sains Dan Matematika Unpam*, 2(1), 15. doi:10.32493/jsmu.v2i1.2915
- [11] Sinaga, L., Ahmad, A., & Safii, M. (2019). Penerapan Data Mining Pada Jumlah Pelanggan Perusahaan Air Bersih Menurut Provinsi Menggunakan Metode K-Means Clustering. *Jurnal RESISTOR (Rekayasa Sistem Komputer)*, 2(2), 119–125. doi:10.31598/jurnalresistor.v2i2.418.