

# Clustering Data Persediaan Barang Menggunakan Metode Elbow dan DBSCAN

Trisia Intan Berliana\*, Elvia Budianita, Alwis Nazir, Fitri Insani

Fakultas Sains dan Teknologi, Prodi Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: <sup>1,\*</sup>11950125232@students.uin-suska.ac.id, <sup>2</sup>elvia.budianita@uin-suska.ac.id, <sup>3</sup>alwis.nazir@uin-suska.ac.id, <sup>4</sup>fitri.insani@uin-suska.ac.id

Email Penulis Korespondensi: 11950125232@students.uin-suska.ac.id

Submitted: 07/12/2023; Accepted: 23/12/2023; Published: 26/12/2023

**Abstrak**—Dalam dunia bisnis dan manajemen persediaan, pengelolaan persediaan yang efisien merupakan hal yang sangat penting. Jika suatu perusahaan tidak memiliki persediaan, maka tidak mungkin dapat memenuhi keinginan konsumen. Mengelola persediaan barang memerlukan manajemen persediaan yang cermat, analisis data yang baik. Tantangan dalam persediaan barang melibatkan fluktuasi permintaan yang tidak terduga, menyulitkan penentuan tingkat persediaan optimal. Diversifikasi produk dengan berbagai karakteristik juga menjadi hambatan, menghambat pengelompokan dan perumusan strategi manajemen persediaan. Kurangnya segmentasi produk yang jelas menambah faktor penghambat, sulit mengidentifikasi kelompok barang serupa. Penumpukan stok yang tidak efisien dapat merugikan bisnis secara keseluruhan, sehingga penerapan clustering diperlukan untuk mengoptimalkan strategi persediaan berdasarkan karakteristik produk. Dengan menganalisis kelompok produk, perusahaan dapat mengembangkan strategi manajemen persediaan yang lebih efisien dan efektif. Pada penelitian ini menggunakan metode clustering dengan menggunakan metode elbow dan DBSCAN (Density-Based Spatial Clustering of Applications with Noise). Metode elbow yang digunakan untuk menentukan nilai eps dan minpts yang paling optimal. Tujuan penelitian ini untuk mengelompokkan data persediaan barang dengan menggunakan atribut Initial quantity (stok awal), quantity sold (stok terjual), dan quantity available (stok produk tersedia). Sehingga dengan data yang dikelompokkan dapat memudahkan perusahaan untuk mengoptimalkan persediaan barang yang paling banyak terjual dan peminatnya. Berdasarkan hasil pengujian elbow dan DBSCAN diperoleh 144 cluster dan 0 data noise dengan cluster 2 merupakan produk dengan jumlah penjualan dan persediaan yang paling banyak. Metode DBSCAN yang diuji tanpa menggunakan elbow memperoleh hasil cluster 3 dan 959 data noise.

**Kata Kunci:** Persediaan Barang ; Data Mining ; Clustering; Elbow ; DBSCAN

**Abstract**— In the world of business and inventory management, efficient inventory management is very important. If a company does not have inventory, it is impossible to fulfill consumer desires. Managing inventory requires careful inventory management and good data analysis. Challenges in inventory involve unpredictable fluctuations in demand, making it difficult to determine optimal inventory levels. Product diversification with various characteristics is also an obstacle, hindering grouping and formulating inventory management strategies. The lack of clear product segmentation adds to the inhibiting factor, making it difficult to identify groups of similar goods. Inefficient stockpiling can be detrimental to the business as a whole, so implementing clustering is necessary to optimize inventory strategies based on product characteristics. By analyzing product groups, companies can develop more efficient and effective inventory management strategies. This research uses a clustering method using the elbow method and DBSCAN (Density-Based Spatial Clustering of Applications with Noise). The elbow method is used to determine the most optimal EPS and Minpts values. The aim of this research is to group goods inventory data using the attributes Initial quantity (initial stock), quantity sold (stock sold), and quantity available (available product stock). So that grouped data can make it easier for companies to optimize the inventory of the most sold goods. and fans. Based on the elbow and DBSCAN test results, 144 clusters and 0 noise data were obtained, with cluster 2 being the product with the largest number of sales and inventory. The DBSCAN method which was tested without using elbows obtained cluster 3 results and 959 noise data.

**Keywords:** Inventory; Clustering; Data Mining; Elbow; DBSCAN

## 1. PENDAHULUAN

Dalam zaman pertumbuhan teknologi yang pesat, teknologi dapat mempermudah aktivitas manusia di berbagai sektor sehingga lebih efisien dan efektif. Teknologi mampu menyediakan informasi yang akurat dan membantu mengorganisasi data perusahaan dalam jumlah besar. Selain itu, teknologi juga mendukung perusahaan dalam pengendalian persediaan, yang dilaksanakan untuk memenuhi kebutuhan konsumen dan menyediakan barang yang diperlukan[1]. Pengambilan keputusan yang tepat dalam menetapkan kebijakan perusahaan. Salah satu manfaat teknologi yang dapat dimanfaatkan dalam industri adalah dalam mengelola persediaan barang. Persediaan adalah sejumlah bahan atau barang yang perusahaan sediakan, termasuk barang jadi, bahan mentah, dan barang dalam proses[2]. Semua ini disiapkan untuk menjaga kelancaran operasi perusahaan dan memenuhi permintaan konsumen kapan pun dibutuhkan. Penyimpanan barang adalah tindakan penting bagi perusahaan dagang, karena itu adalah elemen kunci dalam sektor perdagangan [3]. Kesalahan kecil terkait dengan persediaan barang dapat menyebabkan berbagai masalah, baik itu dalam hal akumulasi barang di gudang maupun ketidaktersediaan barang[4].

Pencatatan persediaan barang menggunakan perangkat lunak, namun masih mengalami kendala, terutama terkait dengan fluktuasi permintaan pelanggan yang menyebabkan ketidakstabilan persediaan. Perubahan permintaan yang dinamis dapat mengakibatkan penumpukan stok barang yang kurang diminati atau kekurangan

stok barang yang diminati. Keragaman produk juga menyebabkan manajemen persediaan yang kurang optimal, menghambat pemenuhan kebutuhan konsumen.

Kesulitan utama muncul dalam pengelompokan persediaan, terutama pada produk yang memiliki penjualan tinggi. Tanpa pengelompokan yang akurat, toko kesulitan mengelola persediaan dengan efektif. Untuk dapat mengatasi permasalahan tersebut, maka dibutuhkan suatu metode dan sistem perencanaan stok barang yang lebih baik sehingga memungkinkan perusahaan dapat menyusun strategi pengelompokan persediaan serta mengoptimalkan persediaan untuk memastikan ketersediaan produk yang cukup tanpa mengalami kekurangan atau penumpukan yang tidak diinginkan. Oleh karena itu, diperlukan suatu teknik pengolahan data dengan teknik data mining clustering untuk mengelompokkan persediaan barang, melihat pola permintaan dengan mengelompokkan produk dengan permintaan serupa. Teknik data mining yang akan digunakan pada penelitian ini adalah pengelompokan atau clustering menggunakan metode elbow dan metode DBSCAN.

Data mining adalah suatu proses statistik yang bertujuan untuk mengidentifikasi dan mengelompokkan data. Proses data mining ini bermanfaat untuk memberikan solusi kepada para pengambil keputusan di dunia bisnis, dengan tujuan untuk meningkatkan kinerja perusahaan [5].

Data mining memiliki kecenderungan dalam sejumlah data yang disimpan dalam penyimpanan yang menggunakan teknik pola seperti matematika dan statistic [6]. Proses clustering adalah proses di mana objek-objek dikelompokkan berdasarkan prinsip kesamaan antara kelas-kelasnya [7]. Clustering dikelompokkan ke dalam suatu kelompok yang mempunyai tingkat kemiripan/kesamaan yang tinggi dan sebaliknya yang mempunyai tingkat ketidaksamaan yang tinggi (kesamaan yang rendah) dari kelompok berbeda [8].

Clustering melibatkan upaya untuk memaksimalkan kesamaan di antara anggota dalam satu kelompok dan meminimalkan kesamaan di antara kelompok atau cluster [9]. Terdapat beberapa metode yang dapat digunakan untuk mencari atau menguji cluster terbaik, salah satunya metode elbow. Metode elbow merupakan metode yang digunakan untuk mencari jumlah cluster atau nilai yang optimal untuk suatu dataset [10]. DBSCAN merupakan sebuah algoritma clustering yang melakukan cluster atau pengelompokan data dengan prinsip mengelompokkan data yang berdekatan pada satu kelompok [11]. DBSCAN mengelompokkan bersama titik-titik yang saling dekat berdasarkan pengukuran jarak (biasanya jarak Euclidean) dan jumlah minimum titik. Ini juga menandai sebagai outlier titik-titik yang berada di daerah kepadatan rendah [12].

Penelitian sebelumnya yang relevan dengan topik ini telah dilakukan oleh Ramdhan, Dwilestari, Dana, Ajiz, dan Kasiani pada tahun 2022 dengan judul "Clustering Data Persediaan Barang Menggunakan Metode K-Means." Fokus penelitian ini adalah menentukan jumlah persediaan barang, mengingat fluktuasi permintaan barang yang selalu berubah dari waktu ke waktu. Untuk mengatasi tantangan tersebut, diperlukan manajemen data persediaan barang yang efisien, yang melibatkan pengolahan data historis transaksi menggunakan teknik data mining dengan metode algoritma K-Means Clustering. Dengan mengelompokkan data menggunakan metode K-Means, diharapkan perusahaan dapat lebih mudah menentukan persediaan barang dan dengan demikian meningkatkan efisiensi dalam pengadaan produk atau barang [13].

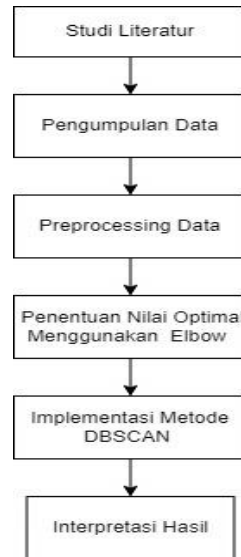
Selanjutnya, Dedy (2022) melakukan penelitian berjudul "Analisis Metode DBSCAN (Density-Based Spatial Clustering Of Application with Noise) dalam Mendeteksi Data Outlier." Fokus utama penelitian ini adalah menganalisis data mining untuk mengidentifikasi data outlier melalui metode pengelompokan (clustering). Metode clustering yang digunakan adalah DBSCAN (Density-Based Spatial Clustering Of Application with Noise). Hasil penelitian menyimpulkan bahwa nilai epsilon memiliki pengaruh signifikan terhadap jumlah cluster yang terbentuk. Semakin tinggi nilai epsilon, maka jumlah cluster yang mungkin terbentuk akan semakin sedikit [14]. Kemudian penelitian sebelumnya yang relevan dengan topik ini telah dilakukan oleh Sari, Oktavianto, dan Sulistyono pada tahun 2022 dengan judul "Algoritma K-Means dengan Metode Elbow untuk Mengelompokkan Kabupaten/Kota di Jawa Tengah Berdasarkan Komponen Pembentuk Indeks Pembangunan Manusia." Penelitian ini difokuskan pada upaya pengelompokan wilayah berdasarkan empat komponen yang membentuk Indeks Pembangunan Manusia, dengan tujuan meningkatkan indeks pembangunan manusia berdasarkan faktor-faktor tersebut. Penelitian ini menggunakan metode K-Means dan menerapkan metode elbow untuk menemukan jumlah cluster yang optimal guna menentukan ukuran cluster terbaik. Hasil penelitian menunjukkan bahwa algoritma K-Means dikembangkan berdasarkan komponen pembentuk Indeks Pembangunan Manusia, dan nilai jarak SSE (Sum of Squares Error) dari 1 cluster ke 2 cluster pada titik siku Elbow adalah sebesar 1022987468. Hasil pengelompokan menunjukkan bahwa terdapat 29 kabupaten/kota dalam cluster 1 dan 6 kabupaten/kota dalam cluster 2 [15].

Penelitian ini menggunakan metode elbow untuk dapat memberikan penentuan jumlah kelompok yang optimal pada data persediaan barang. Penelitian ini dapat membuka peluang untuk mengevaluasi ketepatan metode Elbow dan mempertimbangkan variabilitas hasilnya terhadap jenis dan karakteristik data persediaan yang berbeda. Serta penggunaan metode DBSCAN yang dapat mengatasi tantangan spesifik dalam pengelompokan data persediaan barang. Penelitian ini dapat memberikan wawasan tentang sejauh mana DBSCAN dapat memberikan keunggulan dibandingkan dengan metode clustering lainnya dalam mengatasi noise dan pola-pola kompleks dalam data persediaan. Skalabilitas dan kemampuannya untuk mengidentifikasi cluster dengan kepadatan yang berbeda-beda menjadi alasan utama yang mendukung penggunaan DBSCAN dalam konteks manajemen persediaan.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Pada penelitian ini akan dilakukan analisa pengelompokan menggunakan metode elbow dan DBSCAN, Adapun tahapan-tahapan yang dilakukan terdapat pada Gambar 1 dibawah ini.



**Gambar 1.** Tahapan Penelitian

Dari tahapan penelitian diatas dapat menjelaskan beberapa tahapan yaitu :

1. Studi Literatur

Studi literatur yang digunakan sebagai referensi atau rujukan untuk penelitian bersumber dari jurnal-jurnal atau penelitian terkait, buku dan e-book yang berkaitan dengan metode elbow dan DBSCAN.

2. Pengumpulan Data

Pengumpulan data yang digunakan bersumber dari kaggle. Data yang digunakan merupakan data persediaan toko kelontong. Data terdiri dari 1000 data. Dataset memiliki atribut 9 atribut. Adapun atribut yang digunakan pada dataset kaggle untuk penelitian adalah kuantitas awal kuantitas terjual dan kuantitas tersedia.

3. Preprocessing Data

Pra-pemrosesan data merupakan suatu proses yang melibatkan beberapa tahap untuk mengubah data awal menjadi bentuk yang lebih sederhana dan dapat dipahami [16].

a. Cleaning Data

Tahap pembersihan (cleaning data) merupakan bagian dari preprocessing data. Tujuannya untuk membersihkan data mentah dari masalah nilai kosong (missing values), data duplikat, dan lainnya yang dapat mengganggu keakuratan hasil analisis

b. Transformasi Data

Pada tahapan transformasi data dilakukan perubahan struktur ataupun format dari data mentah agar lebih sesuai untuk analisis atau pemodelan. Jika terdapat data yang berbentuk teks, maka data dapat diubah menjadi representasi numerik. Namun jika sudah berbentuk numerik maka tidak perlu dilakukan perubahan atau transformasi data

c. Normalisasi Data

Normalisasi data merupakan proses transformasi data dalam rangkaian sehingga semua nilai datanya berada dalam rentang tertentu. Normalisasi pada data set yang digunakan pada penelitian kali ini dilakukan menggunakan rumus scalling yaitu :

$$X_{\text{normalize}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \tag{1}$$

Keterangan:

X = Nilai atribut

Xmin = nilai minimum dari atribut.

Xmax = nilai maximum dari atribut.

4. Penentuan Nilai Menggunakan Elbow

Pada tahap ini dilakukan pengujian menggunakan python untuk menentukan kombinasi dari nilai eps dan minpts yang akan digunakan untuk metode DBSCAN. Penentuan jumlah cluster optimal dilakukan dengan memperhatikan nilai SSE (Sum Square Error) yaitu perhitungan jumlah kuadrat jarak antara centroid dan

anggota cluster digunakan untuk menentukan jumlah cluster yang ideal untuk membentuk siku [17]. Perhitungan nilai SSE dapat dilakukan menggunakan rumus :

$$SEE = \sum_{k=1}^k \sum_{x_i \in S_k} \|X_i - C_k\|^2 \tag{2}$$

Keterangan:

K = jumlah cluster

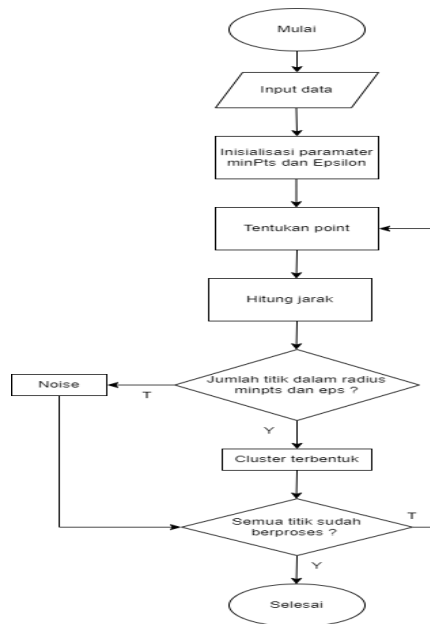
X<sub>i</sub> = nilai atribut dari data ke-I

C<sub>k</sub> = jumlah cluster i pada cluster ke-k"

"" = perhitungan jarak Euclidean

5. Analisa Metode DBSCAN

Metode DBSCAN merupakan algoritma clustering berdasarkan kepadatan data dengan menentukan jumlah data (MinPts) yang terdapat dalam radius Epsilon ( eps ) setiap data[18]. Kepadatan dalam DBSCAN mengakibatkan tiga jenis status pada setiap data, yakni core (inti), border (batas), dan noise (gangguan) [19]. DBSCAN menggunakan kepadatan data untuk menentukan klaster tanpa memerlukan definisi jumlah klaster sebelumnya, ini memungkinkan algoritma untuk menemukan klaster dengan bentuk dan ukuran yang berbeda dalam ruang data[20] . Setelah data melewati proses preprocessing data dan metode elbow maka data siap dipakai untuk tahap selanjutnya yaitu menggunakan metode DBSCAN.Tahap-tahap yang dapat dilakukan pada metode DBSCAN dijelaskan menggunakan flowchart seperti gambar 2 yaitu:



**Gambar 2.** Flowchart Metode DBSCAN

Dari flowchart diatas, menjelaskan tahapan dari algoritma DBSCAN yaitu :

- Input data
- Inisialisasi parameter minpts dan eps.Pilih nilai
- Tentukan point atau titik awak secara acak.
- Ulangi langkah 4-6 hingga semua titik telah diproses.
- Hitung jarak dari semua titik yang density reachable terhadap p. menggunakan rumus euclidean distance.

$$d_{xy} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \tag{3}$$

- Jika data tidak termasuk titik pusat atau yang memenuhi eps lebih dari minpts maka titik p adalah core point dan cluster terbentuk.
- Jika p border point dan tidak terdapat titik yang density reachable terhadap p, maka proses dilanjutkan ke titik yang lain.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Analisa Kebutuhan Data

Analisa kebutuhan data dilakukan pada data penelitian yang telah dikumpulkan.Data yang telah didapatkan yaitu data berumbber dari kaggle yang berupa data dari toko kelontong.Selanjutnya akan dilakukan pengolahan data menggunakan tahapan Preprocessing data.Adapun tahapannya sebagai berikut:

### 3.1.1 Preprocessing Data

Preprocessing data dilakukan untuk membersihkan, mengubah dan mempersiapkan data mentah dan memastikan tidak ada data kosong sebelum digunakan dalam analisis data. Tabel 1 merupakan dataset yang akan digunakan dalam penelitian.

**Tabel 1.** Data Set

Invent oryId	Store Name	City	Area	Product Name	Intial Quantity	Quantit y Sold	Quantity Available	Cost	Retail Price
1	Natio nal Stores	Ouro Branco	9 Springview Point	Chocolate Bar - Smarties	12	1	11	487 3,07	1500,3 9
2	Famil y Dolla r	Geraka rouí	5434 Daystar Circle	Pepper - Red Bell	7	6	1	308 9,77	2095,6 1
3	BJ's Whol esale Club	Radobo j	3 Darwin Drive	Chickensplit Half	21	10	11	459 1,63	364,02
4	Ocea n State Job Lot	Al Madíð	684 Bunting Lane	Zucchini - Green	4	3	1	33,1 3	2111,2 1
5	Ollie' s Barga in Outlet	Bo Phloi	50162 John Wall Drive	Cod - Salted, Boneless	12	9	3	375 0,18	934,82
6	BJ's Whol esale Club	Pieńsk	8 Fordem Pass	Flower - Dish Garden	10	1	9	407 1,16	366,89
...	...	...	...	...	...	...	...	...	...
999	Tuesd ay Morni ng	Gaya	496 Dapin Hill	Bread - Wheat Baguette	12	6	6	231 1,42	487,96
1000	Big Lots	Sexmo an	99 Granby Place	Soup - Knorr, French Onion	4	2	2	471 2,33	1769,6 1

#### a. Cleaning Data

Tahap ini dilakukan pembersihan data, dari 1000 data set, tidak terdapat missing values ( data kosong ) maupun data duplikat dan lainnya yang mengganggu analisis. Tabel 2 merupakan hasil dari cleaning data pada penelitian ini.

**Tabel 2.** Data Cleaning

Id	Intial Quantity	Quantity Sold	Quantity Available
1	12	1	11
2	7	6	1
3	21	10	11
4	4	3	1
5	12	9	3
6	10	1	9
7	11	7	4
...	...	...	...
...	...	...	...
998	15	9	6
999	12	6	6
1000	4	2	2

#### b. Transformasi Data

Pada tahapan transformasi data, data yang akan digunakan adalah initial quantity (kuantitas awal), quantity sold (kuantitas terjual), dan quantity available (kuantitas tersedia / akhir). Data yang digunakan sudah berbentuk

angka semua, jadi tidak terdapat atribut data perlu dilakukan proses transformasi. Tabel 3 dibawah ini merupakan hasil transformasi data pada penelitian ini.

**Tabel 3.** Transformasi Data

<b>Id</b>	<b>Intial Quantity</b>	<b>Quantity Sold</b>	<b>Quantity Available</b>
1	12	1	11
2	7	6	1
3	21	10	11
4	4	3	1
5	12	9	3
6	10	1	9
7	11	7	4
...	...	...	...
...	...	...	...
998	15	9	6
999	12	6	6
1000	4	2	2

c. Normalisasi Data

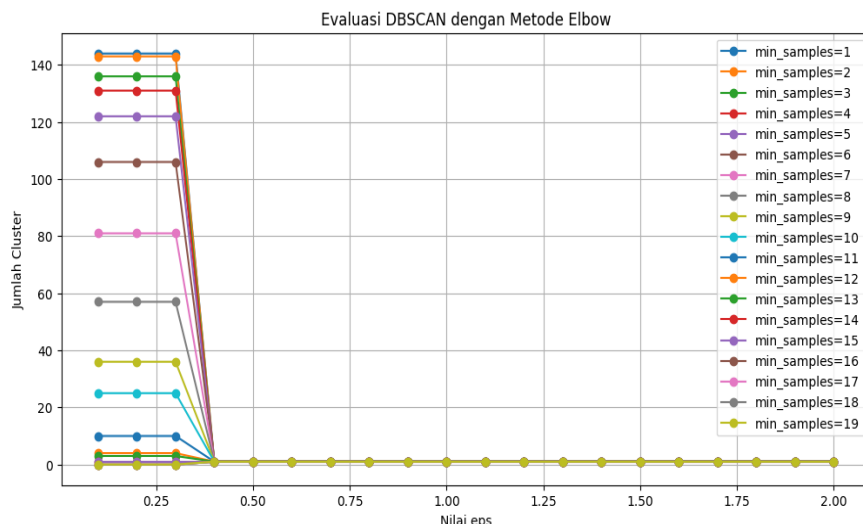
Normalisasi data merupakan proses transformasi data dalam rangkaian sehingga semua nilai datanya berada dalam rentang tertentu. Pada tabel 4 dibawah ini merupakan hasil normalisasi data yang telah dilakukan menggunakan rumus scalling seperti pada rumus 1.

**Tabel 4.** Normalisasi Data

<b>Id</b>	<b>Kuantitas Awal</b>	<b>Kuantitas Terjual</b>	<b>Kuantitas Tersedia</b>
0	0.873719	-0.122168	1.293041
1	0.041804	-1.305391	1.293041
2	0.041804	1.652667	-1.509388
3	0.041804	-0.122168	0.172069
4	-1.206070	-1.009585	-0.668659
...	...	...	...
995	0.873719	1.652667	-0.388417
996	-0.166175	-0.417974	0.172069
997	-0.166175	1.356861	-1.509388
998	-1.830006	-1.305391	-1.229145
999	0.249783	0.469444	-0.108174

**3.1.2 Penentuan Nilai Elbow**

Dalam penelitian ini elbow digunakan sebagai metode untuk menentukan nilai eps dan minPts yang optimal dalam pengelompokan data persediaan barang dengan 1000 data yang akan diproses pada metode clustering DBSCAN. Tools yang digunakan untuk menerapkan metode elbow ini menggunakan python. Pada gambar 3 dibawah ini merupakan grafik yang dihasilkan ketika data diproses menggunakan metode elbow .



**Gambar 3.** Grafik penentuan eps dan minPts terbaik menggunakan metode elbow

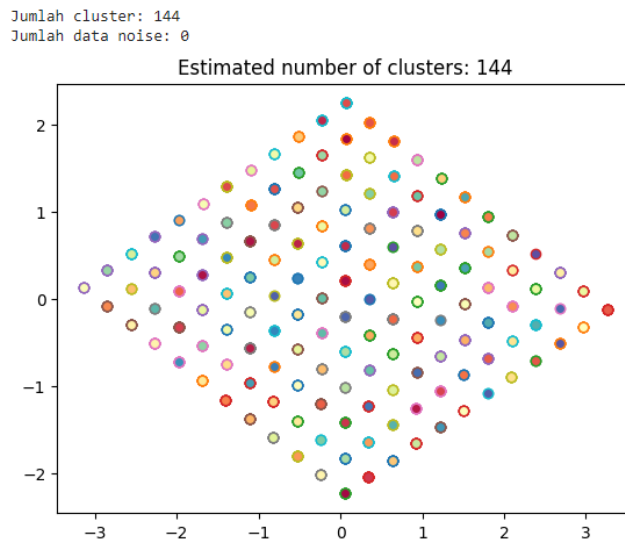
Penentuan nilai eps (epsilon) dan minPts dalam metode Elbow (siku) adalah langkah penting dalam penggunaan algoritma DBSCAN. Algoritma DBSCAN menggunakan parameter-parameter ini untuk mengidentifikasi cluster dalam data. Titik di mana penurunan skor elbow menjadi landai (elbow point) adalah pilihan yang baik, karena itu menunjukkan bahwa penambahan cluster tidak lagi signifikan. Eps dan minPts yang dipilih yaitu 0,01 dan 1.

### 3.1.3 Analisa Metode DBSCAN

Pada penelitian ini didapatkan nilai eps dan minPts yang optimal dengan menggunakan elbow adalah 0,01 dan 1. Secara berurutan pada gambar 4 dan gambar 5 merupakan inialisasi model DBSCAN dan hasil cluster yang didapatkan. Proses clustering menggunakan metode algoritma DBSCAN dengan menggunakan python yaitu :

```
# Inialisasi dan fitting model DBSCAN
dbscan = DBSCAN(eps=0.1, min_samples=1)
dbscan.fit(dataset)
```

**Gambar 4.** Inialisasi DBSCAN



**Gambar 5.** Hasil Cluster

Pada Gambar 5 diatas merupakan hasil cluster menggunakan DBSCAN yang menghasilkan 144 cluster dan 0 data noise. Hasil cluster berupa data yang telah dikelompokkan berdasarkan jumlah kuantitas awal, kuantitas terjual, kuantitas akhir dengan hasil jumlah persediaan atau stok yang sama atau mirip. Adapun beberapa contoh data dari cluster 0,1,2 dan 3 terdapat pada tabel 5 dibawah ini.

**Tabel 5.** Hasil Cluster

id	ProductName	Kuantitas Awal	Kuantitas Terjual	Kuantitas Tersedia	Cluster
1	Wine - White, Ej	17	6	11	0
2	Wine - Kwv Chenin Blanc South	17	6	11	0
3	Wine - Red, Concha Y Toro	17	6	11	0
4	Beans - Fava Fresh	17	6	11	0
5	Chicken Thigh - Bone Out	17	6	11	0
6	Bread - Mini Hamburger Bun	17	6	11	0
7	Towel Dispenser	17	6	11	0
8	Coke - Classic, 355 Ml	17	6	11	0
9	Chips - Doritos	17	6	11	0
10	Veal - Bones	17	6	11	0
11	Wine - Baron De Rothschild	13	2	11	1
12	Coconut - Shredded, Unsweet	13	2	11	1
13	Devonshire Cream	13	2	11	1
14	Pepper - Green, Chili	13	2	11	1
15	Kaffir Lime Leaves	13	2	11	1

id	ProductName	Kuantitas Awal	Kuantitas Terjual	Kuantitas Tersedia	Cluster
16	Beef Ground Medium	13	2	11	1
17	Beef - Tenderloin Tails	13	2	11	1
18	Veal - Heart	13	12	1	2
19	Wine - Montecillo Rioja Crianza	13	12	1	2
20	Beans - French	13	12	1	2
21	Cheese - Marble	13	12	1	2
22	Island Oasis - Magarita Mix	13	12	1	2
23	Five Alive Citrus	13	12	1	2
24	Wine - Ej Gallo Sierra Valley	13	12	1	2
25	Oil - Olive	13	12	1	2
26	Turkey - Breast, Double	13	6	7	3
27	Oregano - Dry, Rubbed	13	6	7	3
28	Rice Paper	13	6	7	3
29	Pea - Snow	13	6	7	3
30	Plaintain	13	6	7	3
31	Mustard - Seed	13	6	7	3
32	Wheat - Soft Kernal Of Wheat	13	6	7	3

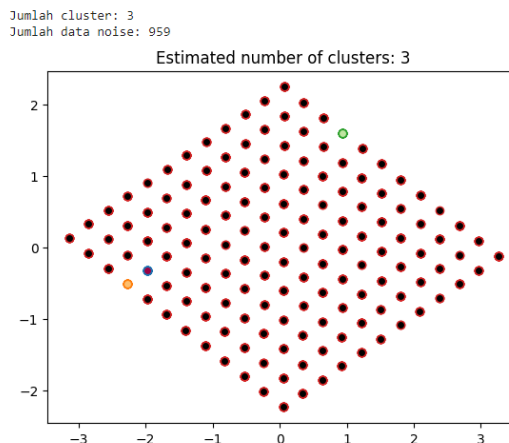
Berdasarkan pengujian yang telah dilakukan, algoritma DBSCAN dan elbow dapat mengelompokkan produk persediaan barang. Penelitian ini menggunakan atribut kuantitas awal, kuantitas terjual, dan kuantitas tersedia dengan menggunakan python. Metode elbow digunakan untuk menentukan nilai eps dan minpts yang paling optimal pada proses clustering DBSCAN, yang dimana nilai eps dan minpts didapat adalah 0,01 dan 1. Hasil cluster pada proses DBSCAN menggunakan python menghasilkan jumlah cluster 144 dan jumlah data noise 0. Hasilnya dapat mengelompokkan produk-produk dengan penjualan dan persediaan yang sama atau mirip. Seperti beberapa cluster dari 144 cluster yang tertulis pada table 4 yang menunjukkan perbedaan dari tiap-tiap cluster. Dari 144 cluster, Cluster 2 merupakan data dengan kuantitas penjualan terbanyak dibandingkan dengan cluster lain.

### 3.1.4 Analisa Metode DBSCAN Tanpa Elbow

Setelah melakukan pengolahan data menggunakan elbow dan DBSCAN dan diperoleh hasil 144 cluster. Untuk melihat hasil dengan cluster yang lebih kecil, dilakukan pengolahan data menggunakan metode DBSCAN tanpa metode elbow. Seperti pada gambar 7 dibawah ini, Eps dan minPts yang digunakan adalah 0,1 dan 13.

```
# Inisialisasi dan fitting model DBSCAN
dbscan = DBSCAN(eps=0.1, min_samples=13)
dbscan.fit(dataset)
```

Gambar 7. Inisialisasi DBSCAN



Gambar 8. Hasil Cluster

Berdasarkan proses metode DBSCAN tanpa menggunakan elbow yang telah dilakukan, diperoleh hasil seperti pada gambar 8 diatas dengan jumlah cluster 3 dan jumlah data noise 959. Adapun data clusternya seperti pada tabel 6 berikut:

**Tabel 6.** Hasil Cluster

id	ProductName	Kuantitas Awal	Kuantitas Terjual	Kuantitas Tersedia	Cluster
1	Coffee Cup 12oz 5342cd	6	4	2	0
2	Corn - Mini	6	4	2	0
3	Water - Mineral, Carbonated	6	4	2	0
4	Dr. Pepper - 355ml	6	4	2	0
5	Kumquat	6	4	2	0
6	Wine - Pinot Noir Stoneleigh	6	4	2	0
7	Squid - Tubes / Tentacles 10/20	6	4	2	0
8	Pepper - Green, Chili	6	4	2	0
9	Crab - Claws, Snow 16 - 24	6	4	2	0
10	Bulgar	6	4	2	0
11	Wine - Rhine Riesling Wolf Blass	6	4	2	0
12	Sauce - Vodka Blush	6	4	2	0
13	Chives - Fresh	6	4	2	0
14	Milk - Buttermilk	6	4	2	0
15	Muffin Hinge - 211n	6	4	2	0
16	Coffee - French Vanilla Frothy	5	4	1	1
17	Lid - 3oz Med Rec	5	4	1	1
18	Beef - Tongue, Fresh	5	4	1	1
19	Bandage - Fexible 1x3	5	4	1	1
20	Bacardi Raspberry	5	4	1	1
21	Fruit Mix - Light	5	4	1	1
22	Island Oasis - Ice Cream Mix	5	4	1	1
23	Capon - Breast, Double, Wing On	5	4	1	1
24	Mushroom Morel Fresh	5	4	1	1
25	Cheese - Brie,danish	5	4	1	1
26	Sauce - Vodka Blush	5	4	1	1
27	Pie Shell - 5	5	4	1	1
28	Pecan Raisin - Tarts	5	4	1	1
29	Beef - Bones, Marrow	16	4	12	2
30	Cake - Lemon Chiffon	16	4	12	2
31	Oxtail - Cut	16	4	12	2
32	Raspberries - Fresh	16	4	12	2
33	Beets - Pickled	16	4	12	2
34	Bread Base - Italian	16	4	12	2
35	Tea - Lemon Scented	16	4	12	2
36	Tea - Herbal Sweet Dreams	16	4	12	2
37	Beef - Tenderloin Tails	16	4	12	2
38	French Kiss Vanilla	16	4	12	2
39	Beef - Baby, Liver	16	4	12	2
40	Soup - Campbells Asian Noodle	16	4	12	
41	Hand Towel	16	4	12	

Pada tabel 6 diatas, merupakan hasil cluster metode DBSCAN tanpa menggunakan elbow.Cluster yang terbentuk terdapat 3 cluster dengan jumlah data noise 959.Dari 3 cluster tersebut terdapat masing-masing perbedaan yang bisa dilihat.Pada Cluster berisikan data dengan produk jumlah penjualan paling sedikit.Cluster 0 berisikan data dengan produk dengan jumlah penjualan atau persediaan yang sedang atau sedikit lebih banyak dari cluster 1. Kemudian cluster 2 yang berisikan data dengan produk yang memiliki jumlah penjualan ataupun persediaan yang paling banyak.

Jadi, berdasarkan pengujian yang telah dilakukan yaitu menggunakan metode elbow dan DBSCAN, diperoleh hasil 144 cluster dan data noise 0. Lalu pengujian metode DBSCAN tanpa elbow dengan eps 0,1 dan minPts 13, memperoleh hasil jumlah cluster 3 dan 959 jumlah data noise.

#### 4. KESIMPULAN

Berdasarkan hasil pengujian yang telah dilakukan pada data set dengan atribut kuantitas awal, kuantitas terjual dan kuantitas akhir menggunakan python, dapat disimpulkan bahwa metode DBSCAN dapat digunakan untuk data

dengan jenis kerapatan yang tinggi, dan metode elbow digunakan untuk menentukan nilai eps dan minPts yang optimal. Adapun pengujian pada metode DBSCAN dilakukan dengan menggunakan metode elbow dan tanpa elbow untuk memperoleh hasil yang lebih baik. Metode DBSCAN dengan menggunakan metode elbow didapat eps 0,1 dan minPts 1 dan memperoleh hasil 144 cluster dan 0 data noise. 144 cluster dikelompokkan berdasarkan jumlah persediaan yang serupa atau mirip dengan cluster 2 yang merupakan produk dengan penjualan dan persediaan yang tinggi. Lalu pengujian metode DBSCAN yang dilakukan tanpa elbow dengan menggunakan eps 0,1 dan minPts 13 diperoleh hasil 3 cluster dan 959 data noise. Adapun pada 3 cluster yang terbentuk terdapat perbedaan pada masing-masing cluster. Cluster 1 merupakan produk dengan jumlah penjualan atau persediaan yang paling sedikit, cluster 0 berisikan produk dengan penjualan dan persediaan yang sedang atau sedikit lebih banyak dari cluster 0, dan cluster 2 dengan produk yang memiliki jumlah penjualan dan persediaan yang paling banyak atau tinggi. Dari kedua hasil pengujian dapat dilihat bahwa pengujian menggunakan elbow terdapat kelebihan masing-masing. Pengujian menggunakan elbow dapat memberikan gambaran visual titik clustering maupun menentukan jumlah cluster dan eps yang optimal. Sementara tanpa menggunakan pengujian elbow dapat lebih sederhana dan cepat karena dapat langsung memilih konfigurasi DBSCAN tanpa melakukan analisis tambahan.

## REFERENCES

- [1] Ika Anikah, Agus Surip, Nela Puji Rahayu, Muhammad Harun Al- Musa, and Edi Tohidi, “Pengelompokan Data Barang Dengan Menggunakan Metode K-Means Untuk Menentukan Stok Persediaan Barang,” *KOPERTIP J. Ilm. Manaj. Inform. dan Komput.*, vol. 4, no. 2, pp. 58–64, 2022, doi: 10.32485/kopertip.v4i2.120.
- [2] Syardiansah, M. Fuad, and P. Sari, “Analisis Pengendalian Persediaan Produksi Pada Cv. Fanara Abadi,” *J. Ilm. Manaj.*, vol. 8, no. 2, pp. 80–91, 2020.
- [3] J. Heizer and B. Render, *Operations management*. New Jersey: Pearson Education, 2004.
- [4] W. Wilda Rina Hasibuan, “Sosialisasi Sistem Informasi Persediaan Barang Pada Pt. Immunotec Profarmasia,” *J. Pengabd. Bareleng*, vol. 4, no. 1, pp. 20–27, 2022, doi: 10.33884/jpb.v4i1.4674.
- [5] M. H. Fakhriza and K. Umam, “Analisis Produk Terlaris Menggunakan Metode K-Means Clustering Pada “Pt.Sukanda Djaya,” *JIKA (Jurnal Inform.*, vol. 5, no. 1, p. 8, 2021, doi: 10.31000/jika.v5i1.3236.
- [6] D. . Larose, *Discovering knowledge in data : an introduction to data mining*. New Jersey John Wiley & Sons, 2005.
- [7] F. Indriyani and E. Irfiani, “Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means,” *JUITA J. Inform.*, vol. 7, no. 2, p. 109, 2019, doi: 10.30595/juita.v7i2.5529.
- [8] L. Muflikhah, D. E. Ratnawati, and R. R. M. Putri, *Data Mining*. Malang: UB Press, 2018.
- [9] R. Rustam, S. Rahmatullah, S. Supriyanto, and S. Wahyuni, “Penerapan Data Mining Untuk Prediksi Penjualan Produk Triplek Pada Pt Puncak Menara Hijau Mas,” *J. Inf. dan Komput.*, vol. 8, no. 2, pp. 75–86, 2020, doi: 10.35959/jik.v8i2.186.
- [10] R. Sarno, shoffi I. Sabilla, Malikhah, D. P. Purbawa, and S. H. Ardani, *Machine Learning Deep Learning Konsep dan pemrograman python*. Penerbit ANDI, 2019.
- [11] R. R. Muhima et al., *Kupas Tuntas Algoritma Clustering : Konsep,Perhitungan Manual,dan Program*. Penerbit ANDI, 2022.
- [12] M. Ramadhani and D. Fitriana, “Implementation of data mining analysis to determine the tuna fishing zone using DBSCAN algorithm,” *Int. J. Mach. Learn. Comput.*, vol. 9, no. 5, pp. 706–711, 2019, doi: 10.18178/ijmlc.2019.9.5.862.
- [13] D. Ramdhan, G. Dwilestari, R. D. Dana, A. Ajiz, and K. Kaslani, “Clustering Data Persediaan Barang Dengan Menggunakan Metode K-Means,” *MEANS (Media Inf. Anal. dan Sist.*, vol. 7, no. 1, pp. 1–9, 2022, doi: 10.54367/means.v7i1.1826.
- [14] D. Armiady, “Analisis Metode DBSCAN (Density-Based Spatial Clustering of Application with Noise) dalam Mendeteksi Data Outlier,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 6, p. 2158, 2022, doi: 10.30865/jurikom.v9i6.5080.
- [15] R. Y. Sari, H. Oktavianto, and H. W. Sulisty, “ALGORITMA K-MEANS DENGAN METODE ELBOW UNTUK MENGELOMPOKKAN KABUPATEN/KOTA DI JAWA TENGAH BERDASARKAN KOMPONEN PEMBENTUK INDEKS PEMBANGUNAN MANUSIA,” pp. 1–9, 2022.
- [16] R. Swastika, S. Mukodimah, F. Susanto, M. Muslihudin, and S. Ipnuwati, *Implementasi Data Mining (Clustering, Association, Prediction, Estimation, Classification)*. Penerbit Adab, 2023.
- [17] A. B. H. Kiat, Y. Azhar, and V. Rahmayanti, “PENERAPAN METODE K-MEANS DENGAN METODE ELBOW UNTUK SEGMENTASI PELANGGAN MENGGUNAKAN MODEL RFM(Recency, Frequency, & Monetary),” *J. Repos.*, vol. 2, no. 7, p. 945, 2020, doi: 10.22219/repositor.v2i7.973.
- [18] A. Kristianto, “Analisa Performa K-Means dan DBSCAN dalam Clustering Minat Penggunaan Transportasi Umum,” *Elkom J. Elektron. dan Komput.*, vol. 14, no. 2, pp. 368–372, 2021, doi: 10.51903/elkom.v14i2.551.
- [19] A. Kristianto, “Implementasi DBSCAN dalam Clustering Data Minat Mahasiswa Setelah Pandemi Covid19,” *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 2, no. 2, pp. 426–431, 2022, doi: 10.24002/konstelasi.v2i2.5638.
- [20] Z. V. Vi. Khan, D. Alamsyah, and W. Widhiarso, “Klasterisasi Topik Skripsi Informatika dengan Metode DBSCAN,” *J. Algoritma*, vol. 3, no. 1, pp. 82–90, 2022, doi: 10.35957/algoritme.v3i1.3337.