

## PENERAPAN METODE *K-MEANS CLUSTERING* PADA PENJUALAN BARANG DI *SPORTS STATION*

Nadia Putri Gantara<sup>1)</sup>, Irfan Ali<sup>2)</sup>

Komputerisasi Akuntansi<sup>1)</sup>, Rekayasa Perangkat Lunak<sup>2)</sup>, STMIK IKMI Cirebon  
Jl. Perjuangan No. 10 B, Majasem Kec. Kesambi Kota Cirebon 45131, Jawa Barat, Indonesia  
E-mail: <sup>1)</sup>[nadiaputrigantara@gmail.com](mailto:nadiaputrigantara@gmail.com), <sup>2)</sup>[irfanaali0.0@gmail.com](mailto:irfanaali0.0@gmail.com)

### ABSTRAK

Penjualan adalah kegiatan menjual barang dan jasa, apabila manajemen penjualan pada perusahaan kurang baik maka akan mempengaruhi keuntungan. Sehingga, membuat perusahaan tidak mencapai tujuannya. Jenis usaha yang ada di Indonesia sangat beragam, salah satunya *Sports Station* merupakan perusahaan retail yang menjual perlengkapan olahraga. *Sports Station* sering mengalami permasalahan dengan ketidakakuratan dan tidak terstrukturanya data penjualan. Sehingga, berdampak dalam kesulitan mengelompokkan produk. Maka, diperlukan sistem yang dapat menentukan pola atau trend dalam penjualan. Algoritma *K-Means* digunakan untuk mengelompokkan produk berdasarkan pola penjualan yang serupa, yaitu dengan membagi data menjadi dua klaster yang dikategorikan sebagai laris dan kurang laris. Tahapan yang diterapkan yaitu *retrieve*, untuk mengambil *dataset*, kemudian menggunakan *K-Means Clustering* untuk memodelkan *dataset* dan *Cluster distance performance* untuk mengevaluasi hasil pengelompokan. Validasi hasil klasterisasi dapat dilakukan menggunakan *Davies Bouldin Index (DBI)*. Algoritma ini menghasilkan pengelompokan menjadi 2 yaitu *cluster 0* dengan nilai 995 sebanyak 121 produk dengan kategori laris dan *cluster 1* dengan nilai 327 sebanyak 2.279 produk dengan kategori kurang laris. Serta hasil DBI yang paling mendekati 0 adalah K 2 menghasilkan nilai 0,10.

**Kata kunci:** *data mining, algoritma, means, clustering*

### ABSTRACT

*Sales are the activities of selling goods and services, if the sales management in the company is not good, it will affect profits. Thus, making the company not achieve its goals. The types of businesses in Indonesia are very diverse, one of which is Sports Station is a retail company that sells sports equipment. Sports Stations often experience problems with inaccuracies and unstructured sales data. Thus, it has an impact on the difficulty of grouping products. Therefore, a system is needed that can determine patterns or trends in sales. With the K-Means algorithm it is used to group products based on similar sales patterns. By dividing the data into two clusters, which are categorized as best-selling and under-selling. The stage applied is Retrieve, to retrieve the dataset, then use K-Means Clustering to model the dataset and Cluster distance performance to evaluate the clustering results. Validation of clustering results can be done using the Davies Bouldin Index (DBI). This algorithm produces a grouping into 2, namely cluster 0 with a value of 995 with 121 products with best-selling categories and cluster 1 with a value of 327 with a value of 2,279 products with less selling categories and the DBI result that is closest to 0 is K 2 resulting in a value of 0.10.*

**Keywords :** *data mining, algorithm, means, clustering*

## 1. PENDAHULUAN

Penjualan atau *sales* merupakan sebuah kegiatan menjual barang atau jasa. Didalam proses penjualan terdapat penyedia barang dan jasa atau biasa disebut penjual yang memberikan produk kepada pembeli dengan harga tertentu [1]. Salah satu aspek penting bagi sebuah perusahaan adalah penjualan. Jika penjualan tidak dikelola dengan baik, hal ini dapat merugikan perusahaan dengan berpotensi mempengaruhi laba dan mengurangi pendapatan perusahaan, bahkan dapat mengakibatkan tidak tercapainya target penjualan serta membuat pelanggan tidak merasa puas. Penjualan juga merupakan kegiatan transaksi yang dilakukan oleh dua belah pihak dengan menggunakan alat pembayaran yang sah [2]. Berdasarkan penelitian dari Firmansyah *et.al* (2020) yang berjudul Prediksi Pola Penjualan Tiket Kapal PT Pelni Cabang Makasar Menggunakan Metode Algoritma Apriori menyimpulkan bahwa dengan adanya kegiatan penjualan maka data penjualan akan semakin bertambah. Tidak hanya berfungsi sebagai arsip dan laporan perusahaan saja data penjualan juga dapat diolah menjadi informasi yang berguna sebagai strategi pemasaran pada perusahaan [3].

Setiap perusahaan perdagangan memiliki tujuan untuk memaksimalkan perkembangan usahanya agar tidak ketinggalan dalam persaingan bisnis yang sangat dinamis. Oleh karena itu, beberapa perusahaan menerapkan strategi tertentu untuk memastikan kelancaran bisnis mereka, salah satunya adalah dengan mengembangkan produk berkualitas tinggi untuk meningkatkan pendapatan perusahaan [4]. *Sports Station* adalah salah satu perusahaan ritel di Indonesia yang fokus pada penjualan perlengkapan olahraga. Dalam beberapa tahun terakhir, *Sports Station* menjadi salah satu usaha yang berkembang pesat, hal ini disebabkan oleh semakin banyaknya masyarakat yang menyadari pentingnya berolahraga untuk menjaga kesehatan tubuh, serta adanya perkembangan teknologi yang memudahkan akses masyarakat dalam membeli produk-produk olahraga. Setiap perusahaan memiliki sistem yang berbeda dalam mengendalikan

usahanya. Namun, *Sports Station* mengalami masalah dalam menentukan performa penjualan yang baik atau buruk karena data penjualan yang tidak terstruktur dan tidak akurat, sehingga sulit untuk mengambil keputusan bisnis yang tepat. Selain itu, untuk meningkatkan penjualan produk dan pendapatan bisnis, *Sports Station* juga perlu mempertimbangkan faktor menentukan produk yang sesuai dengan target pasar. Oleh karena itu, diperlukan suatu sistem yang dapat membantu menentukan pola atau tren dalam penjualan agar bisnis dapat berjalan lebih efektif dan efisien.

Salah satu sistem yang dapat digunakan dalam pengelompokan penjualan produk pada toko menggunakan metode *data mining*. Digunakan untuk menentukan pola atau tren dalam penjualan, identifikasi produk yang laris terjual dan kurang laris serta banyak faktor-faktor lain yang mempengaruhi penjualan. *Data mining* merupakan proses semi otomatis dengan teknik statistik, matematika, kecerdasan buatan serta *mechine learning*. Digunakan untuk memperoleh dan mendapati informasi yang berguna terkait dari basis data yang besar [5].

*Data mining* juga dapat membantu perusahaan membuat keputusan bisnis. Metode *K-Means* adalah salah satu metode *clustering non-hierarki* yang sering digunakan karena kemampuannya dalam mengelompokan data dengan cepat dan efisien dalam bentuk satu atau lebih *cluster*. Metode *K-Means* akan mengelompokan data kedalam suatu kelompok yang dimana data berkarakteristik sama akan dimasukan ke satu kelompok yang sama dan data yang memiliki karakteristik berbeda akan dikelompokan dengan kelompok lainnya [6].

Sitohang *et.al* (2021) dalam penelitiannya berjudul *Data mining* untuk Mengelompokkan Data Penjualan *Cake* dengan Menggunakan Algoritma *K-Means* menyimpulkan bahwa *Jovie Bakery* masih menggunakan proses manual dalam mencatat atau mengumpulkan data penjualan *cake*, sehingga sering terjadi kesalahan dan belum melakukan pengelompokan terhadap *Cake* yang terjual dan tidak terjual. Kondisi ini dapat mengakibatkan kurangnya persediaan *Cake* yang laku dan akumulasi *cake* yang tidak laku.

Untuk menyelesaikan masalah tersebut, dibutuhkan sebuah sistem yang dapat menganalisis masalah dengan mengelompokkan data menggunakan *data mining*, dengan metode *clustering*. Dalam metode *clustering*, pada iterasi pertama, pusat *centroid* awal ditentukan dengan mengacak data yang telah ada. Kemudian, dengan menghitung jarak dari data ke pusat *cluster*, perhitungan jarak dari data pertama hingga data kelima dilakukan. Setelah itu, hasil dihitung dengan menghitung jarak masing-masing data ke pusat *centroid* baru. Selanjutnya, nilai WCV (*Within Cluster Validation*) dihitung dengan memangkatkan dan menjumlahkan jarak terdekat cluster, dan nilai BCV (*Between Cluster Validation*) dihitung dengan menjumlahkan hasil dari jarak setiap *centroid*. Setelah itu, nilai proporsional dihitung dengan membandingkan nilai BCV dan WCV. Pada tahap selanjutnya, *centroid* baru dihitung dengan menghitung rata-rata nilai pada setiap *cluster*. Setelah beberapa tahap, hasil dari penelitian ini dapat membantu para pegawai dalam membuat laporan tentang pengelompokan *cake* yang terjual dan tidak terjual [7].

Penelitian yang dilakukan oleh Fajar Ageng Bramasta dan Risa Helilintar dalam Jurnal Seminar Nasional Inovasi Teknologi tahun 2021 dengan judul "Penerapan Data Mining Untuk Menentukan Strategi Penjualan Toko Sepatu" mengatakan bahwa kasus yang sering terjadi pada toko sepatu adalah dalam penjualan sepatu, pemilik toko kesulitan dalam memastikan produk sepatu apa saja yang sering terjual dan tidak diminati oleh konsumen sehingga data penjualan pada toko tidak dimanfaatkan dengan maksimal. Tujuan dari penelitian ini agar sistem bisa memastikan data untuk mengelompokkan jenis sepatu apa saja yang terjual dengan menggunakan metode *k-means clustering*. *K-means* merupakan salah satu proses tata cara pengelompokan non-hirarki yang cara kerjanya dengan membagi data menjadi beberapa cluster. Cluster tersebut dibagi menjadi 3, yaitu Terlaris, Sedang dan Terendah. data yang dipakai dalam penelitian ini adalah data penjualan pada toko sepatu Ritelindo. hasil dari data yang sudah di-*cluster* menghasilkan 4 item yang tergolong dalam

kelompok penjualan Terlaris, 12 item yang tergolong dalam kelompok Sedang dan 9 item yang tergolong dalam kelompok Terendah [8].

Berdasarkan permasalahan yang sudah diuraikan di atas, penelitian ini menggunakan *data mining* yang merupakan proses untuk mengekstrak informasi yang berguna untuk menentukan pola, hubungan atau tren dalam data yang dapat digunakan untuk mengelompokkan data penjualan yang dapat menentukan produk yang sesuai dengan target pasar dan membuat keputusan bisnis atau meningkatkan penjualan. Serta menggunakan algoritma *K-Means Clustering* yang merupakan sebuah metode pengelompokan data non-hirarki yang berusaha memisahkan data yang ada ke dalam bentuk dua atau lebih kelompok [9]. Serta sebagai metode pengelompokan yang digunakan untuk mengelompokkan data menjadi *K cluster* yang berbeda.

## 2. METODE PENELITIAN

Sumber data dalam penelitian merujuk pada tempat atau lokasi dimana dapat memperoleh informasi atau data yang di perlukan. Sumber data meliputi data penjualan dengan waktu pengambilan data pada bulan Januari sampai Desember 2022. Data penelitian yang dengan sumber data primer diperoleh secara langsung dari tangan pertama dengan menggunakan metode observasi dan metode wawancara terhadap pegawai toko. Dan pada sumber data sekunder diperoleh data yang merujuk pada informasi yang telah ada sebelumnya dari sumber yang sudah ada dengan menggunakan metode dokumentasi.

Metode mengacu pada pendekatan atau strategi yang diterapkan untuk menunjukkan bagaimana teknik atau cara tersebut dapat digunakan, sehingga dapat diperlihatkan penggunaannya melalui:

1. Metode Observasi

Metode observasi dilakukan dengan proses pengamatan yang dilakukan di *Sports Station* dengan prosedur kerja yang ada di toko tersebut.

2. Metode Literatur

Untuk penelitian, informasi yang relevan dengan permasalahan akan dicari,

dikumpulkan, dianalisis, dan dievaluasi dari berbagai sumber, seperti jurnal, *e-book*, literatur, internet, dan sumber lain yang terkait dengan penelitian. Hal ini dilakukan untuk memperoleh teori, metode, dan konsep yang relevan dengan tujuan penelitian.

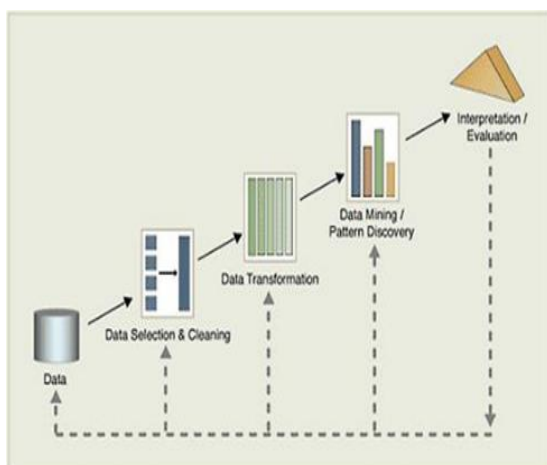
### 3. Metode Wawancara

Metode wawancara digunakan untuk mendapatkan informasi dari pegawai toko yang bertanggung jawab atas data penjualan di *Sports Station*. Karena data penjualan tersebut sangat banyak dan bersifat rahasia, hanya sebagian dari data tersebut yang diberikan untuk penelitian.

### 4. Metode Dokumentasi

Metode dokumentasi dilakukan dengan cara mengumpulkan data yang diperoleh dari dokumen-dokumen yang ada atau tersimpan. Sumber dokumen tersebut dapat berupa dokumen tertulis, rekaman audio atau video, atau bahkan data elektronik. Kemudian dokumen tersebut dipelajari untuk memperoleh data serta informasi yang berkaitan dengan penelitian ini.

Tahapan perancangan mencakup langkah-langkah perancangan dari awal sampai akhir, sebagai acuan penelitian dengan menggunakan proses *Knowledge Discovery In Databases* (KDD) secara garis besar dapat dijadikan sebagai berikut:



**Gambar 1.** Tahapan Perancangan KDD

### 1. Data Selection

*Data Selection* adalah proses pemilihan data dari sekumpulan data operasional yang perlu dilakukan sebelum melakukan tahap KDD (*Knowledge Discovery in Databases*). Data yang telah terpilih dan akan digunakan dalam proses data mining kemudian disimpan dalam suatu berkas yang terpisah dari basis data operasional [10].

### 2. Pre-processing /Cleaning

Pada tahap *Pre-processing*, dilakukan pemeriksaan terhadap nilai yang hilang pada data atau yang disebut dengan *missing value*, agar data yang digunakan menjadi valid. Hal ini dilakukan dengan membersihkan data yang tidak konsisten, termasuk perbaikan data yang tidak sesuai dengan format yang telah ditentukan, serta mengisi nilai yang hilang dengan rata-rata dari data yang tersedia [11].

### 3. Transformation

Data transformasi pada tahap ini digunakan untuk melakukan inialisasi terhadap data yang memiliki nilai normal menjadi nilai numerik [6]. Pada proses transformasi data yang terpilih dalam bentuk *mining procedure*, merupakan proses yang kreatif dan bergantung kepada jenis atau pola yang akan dicari dalam basis data [12].

Dari penjelasan di atas dapat ditarik kesimpulan dalam tahap tata transformasi, dilakukan inialisasi terhadap data yang memiliki nilai normal menjadi nilai numerik. Proses ini bersifat kreatif dan bergantung pada jenis atau pola yang ingin dicari dalam basis data yang terpilih. Proses transformasi data sangat penting dalam KDD karena dapat mempengaruhi kualitas dan keakuratan hasil analisis yang akan dilakukan. Oleh karena itu, proses ini harus dilakukan dengan hati-hati dan

berdasarkan pemahaman yang mendalam mengenai karakteristik data yang diproses.

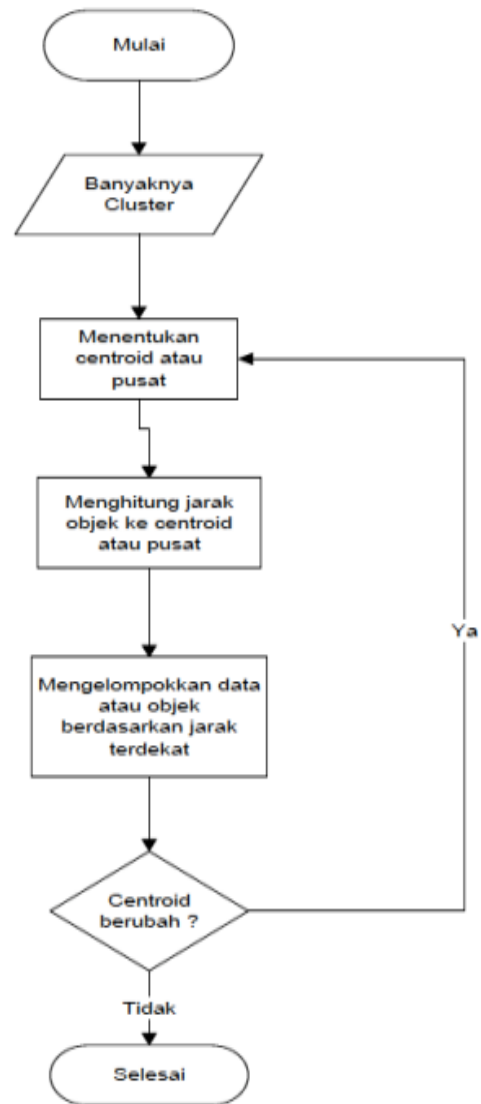
#### 4. *Data mining*

Pada tahap ini, dilakukan proses pemilihan algoritma atau teknik data mining yang akan digunakan untuk mencari pola atau informasi yang relevan dalam data yang telah dipilih. Pemilihan algoritma ini harus didasarkan pada tujuan yang ingin dicapai dan juga harus sesuai dengan keseluruhan proses KDD[13]. *Data mining* adalah salah satu operasi dasar dalam proses KDD yang meliputi berbagai tahapan, termasuk proses pembersihan data. Proses pembersihan data ini mencakup penghapusan data duplikat, pengecekan data yang tidak konsisten atau tidak relevan, serta perbaikan kesalahan pada data [12].

#### 5. *Interpretation/Evaluation*

Tahap evaluasi digunakan untuk mengevaluasi hasil dari tahap data mining dan memastikan bahwa hasil tersebut dapat menjawab tujuan yang telah ditetapkan. Evaluasi melibatkan profilisasi pada setiap *cluster* yang terbentuk dan analisis lebih lanjut untuk menghubungkannya dengan atribut yang relevan. *Data mining* memperlihatkan pola informasi dengan cara yang mudah dipahami oleh pihak yang berkepentingan [14].

Tahapan perancangan yang digunakan sebagai gambaran umum untuk mengetahui cara kerja dan proses dari metode *K-Means Clustering* dengan menggunakan *flowchart*. Berikut merupakan gambaran dari *flowchart* menggunakan metode *K-Means Clustering*.



**Gambar 2.** *Flowchart K-Means Clustering*

### 3. HASIL DAN DISKUSI

Penelitian ini menggunakan tools *RapidMiner* studio versi 10.01. dengan menggunakan data yang diambil pada data penjualan *Sports Station*, data yang digunakan dalam penelitian ini sebanyak 2.400 sampel data. Penelitian ini dilakukan dengan penerapan algoritma *K-Means Clustering*.

#### 1. *Data Selection*

Berikut ini merupakan data penjualan yang diambil dengan menggunakan teknik data primer yang berarti data tersebut diambil langsung dari penjualan, sebagaimana terlihat pada Tabel 1 berikut ini:

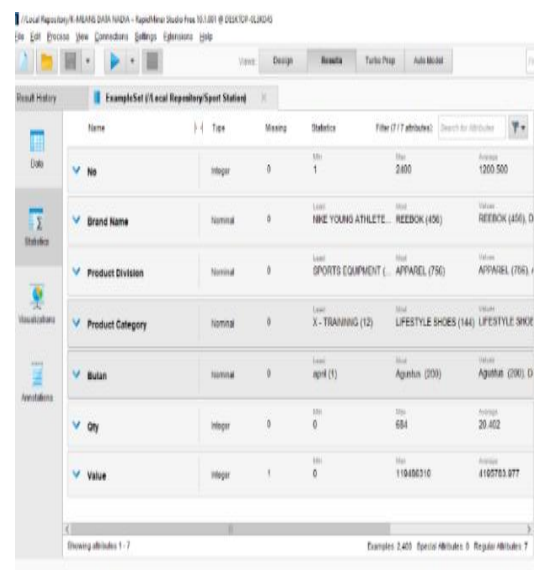
**Tabel 1.** Data Penjualan

<b>Id</b>	<b>Nama Brand</b>	<b>Devisi produk</b>	<b>Kategori produk</b>	<b>Bulan</b>	<b>Qty</b>	<b>Value</b>
1	Airwalk	Accessories	Hat / Cap	Januari	32	2476200
2	Airwalk	Accessories	Lifestyle	Januari	1	99000
3	Airwalk	Accessories	Socks	Januari	1	99000
4	Airwalk	Accessories	Sports Accessories	Januari	32	2476200
5	Airwalk	Apparel	Jacket	Januari	2	642600
6	Airwalk	Apparel	Pants	Januari	2	502600
7	Airwalk	Apparel	Short	Januari	2	418600
8	Airwalk	Apparel	Sports Lifestyle Apparel	Januari	9	2054570
9	Airwalk	Apparel	Tshirt Full Sleeve	Januari	0	0
10	Airwalk	Apparel	Tshirt Short Sleeve	Januari	3	490770
11	Airwalk	Bags	Backpack	Januari	0	0
12	Airwalk	Bags	Shoulder Bag	Januari	4	437800
...	...	...	...	...	...	...
2399	U-Lace	Accessories	Sports General Accessories	Desember	0	0
2400	U-Lace	Footwear	Lifestyle Shoes	Desember	0	0

Berdasarkan Tabel 1. menunjukkan atribut-atribut yang akan digunakan dalam algoritma *K-Means Clustering*. *Data Selection* merupakan bagian yang sangat penting, karena hasil keputusan dan kesimpulan yang dihasilkan sangat bergantung pada kualitas data yang digunakan.

2. *Pre-processing /Cleaning*

Saat mengolah sebuah kumpulan data, seringkali terdapat nilai yang tidak tersedia atau hilang (*missing value*) yang dapat mempengaruhi hasil pemodelan. Oleh karena itu, diperlukan tindakan untuk menangani nilai-nilai tersebut. Salah satu cara yang dapat dilakukan adalah dengan menggunakan operator *retrieve missing value* yang bertujuan untuk mengganti nilai-nilai yang hilang dalam data. Contohnya dapat dilihat pada gambar di bawah ini:



**Gambar 3.** Missing Value

Langkah berikutnya melibatkan penggunaan operator "*nominal to numerical*" yang bertujuan untuk mengkonversi data yang diperlukan oleh beberapa metode pemrosesan data menjadi data numerik. Operator ini hanya dapat menerima data numerik, sehingga data nominal perlu dikonversi menjadi data numerik sebelum diproses dalam pemodelan. Data nominal yang sudah dikonversi akan menjadi fitur baru



atau "dummy variable" dengan nilai 0 dan 1.

Dalam proses selanjutnya, dilakukan perubahan pada tipe filter atribut dengan menggunakan *subset*, dan dipilih atribut tertentu dari *data set* yang digunakan dalam penelitian. Penggunaan operator ini bertujuan untuk membatasi jumlah data dan memfokuskan pada data yang relevan serta memudahkan dalam proses penelitian.

### 3. Transformation

Pada tahap ini, keseluruhan data operasional akan diolah untuk menghasilkan kelompok atribut yang akan digunakan dalam proses transformasi *data mining*. Kelompok atribut tersebut terdiri dari nama merek, divisi produk, kategori produk, bulan, kuantitas, dan nilai. Transformasi data dilakukan untuk mengubah tipe data ke dalam format yang sesuai dengan kebutuhan dalam proses *data mining*. Jika ada data yang memiliki nilai nominal, maka akan dilakukan inisialisasi untuk mengubahnya menjadi nilai numerik sehingga dapat diproses lebih lanjut.

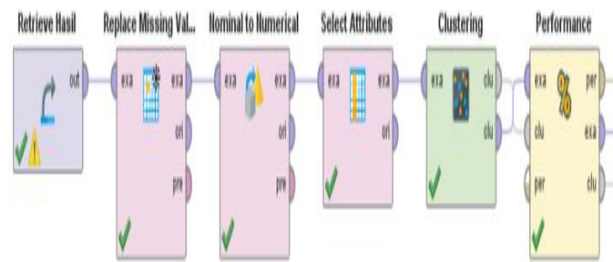
No	Brand Name	Product Div...	Product Ca...	Bulan	Qty
Integer	polynomial	polynomial	polynomial	polynomial	Integer
1	AIRWALK	ACCESSORIES	HAT / CAP	Januari	32
2	AIRWALK	ACCESSORIES	LIFESTYLE	Januari	1
3	AIRWALK	ACCESSORIES	SOCKS	Januari	1
4	AIRWALK	ACCESSORIES	SPORTS ACCES...	Januari	32
5	AIRWALK	APPAREL	JACKET	Januari	2
6	AIRWALK	APPAREL	PANTS	Januari	2
7	AIRWALK	APPAREL	SHORT	Januari	2
8	AIRWALK	APPAREL	SPORTS LIFEST...	Januari	9
9	AIRWALK	APPAREL	TSHIRT FULL SL...	Januari	0
10	AIRWALK	APPAREL	TSHIRT SHORT ...	Januari	3
11	AIRWALK	BAGS	BACKPACK	Januari	0
12	AIRWALK	BAGS	SHOULDER BAG	Januari	1

Gambar 4. Transformasi Data

### 4. Data mining

Model *data mining* adalah hasil akhir dari proses *data mining* yang

menghasilkan informasi berharga dan bermanfaat dari kumpulan data yang besar. Model ini dapat digunakan untuk model prediksi (*prediction modelling*), analisis kelompok (*cluster analysis*), analisis asosiasi (*association analysis*) dan deteksi anomaly (*anomaly detection*) [14]. Model *data mining* terdiri dari beberapa tahapan, seperti pemilihan data, *preprocessing data*, pemilihan algoritma, pembangunan model, evaluasi model, dan penggunaan model.



Gambar 5. Model Data mining

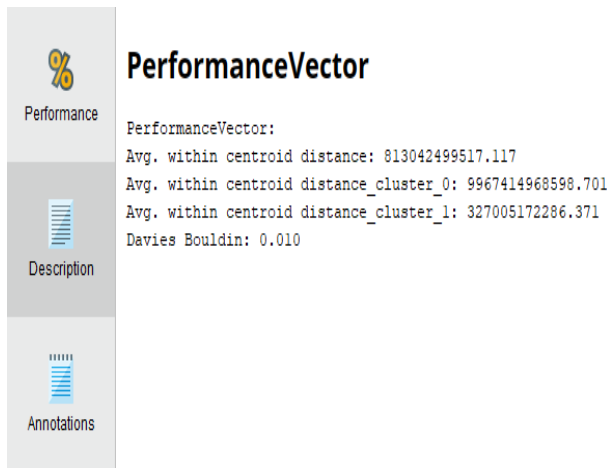
### 5. Hasil Evaluation

Dengan menggunakan operator *cluster distance performance*, data yang telah ada dapat dibentuk menjadi beberapa *cluster* dengan membaginya atau menentukannya. Untuk menentukan jumlah *cluster* K yang paling optimum, dapat digunakan operator *performance* untuk menghitung *Devies Boulding Index* (DBI) yang merupakan salah satu uji validitas internal pengelompokan berdasarkan jumlah kedekatan data dengan *centroid* dari *cluster* yang diikuti dan antar dua *cluster*, diukur dengan kedekatan dua *centroid cluster* [15].

Nilai K yang paling optimal dapat dicari dengan mencari nilai DBI yang mendekati 0. Berdasarkan hasil operator *performance*, nilai DBI yang paling mendekati 0 adalah pada K 2. Oleh karena itu, jumlah *cluster* yang dibuat adalah 2, yaitu laris dan tidak laris. Selain itu, terdapat 4 atribut yang digunakan dalam pengelompokan data, yaitu nama *brand*, bulan, *qty*, dan *value*.

(a) Hasil *Performance Vector*

Berdasarkan hasil dari penerapan algoritma *k-mean clustering* maka hasil yang didapatkan pada *performance* yaitu seperti gambar berikut:



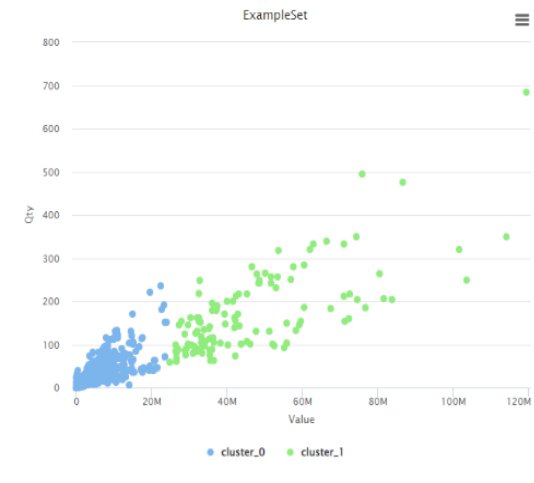
**Gambar 6.** Hasil *Performance Vector*

Berdasarkan Gambar 6 yang menunjukkan data *performance*, dapat dijelaskan bahwa data tersebut telah dikelompokkan menjadi dua kelompok, yaitu kelompok produk yang laris terjual dan kurang laris terjual. Kelompok produk yang laris terjual dapat ditemukan pada *cluster* 0 dengan nilai sebesar 995, sedangkan kelompok produk yang kurang laris terjual dapat ditemukan pada *cluster* 1 dengan nilai sebesar 327. Parameter untuk mengukur kinerja algoritma *K-Means* dilakukan dengan menghitung nilai DBI, dimana nilai DBI yang mendekati 0 yaitu pada  $K=2$  dengan nilai 0,10.

1. Hasil *Visualizations*

Pada proses *data mining* di *RapidMiner*, hasil visualisasi merujuk pada representasi hasil dalam bentuk visual seperti grafik, plot, dan tabel. Dalam penelitian ini, *scatter/bubble* digunakan untuk tujuan mempermudah aksesibilitas dan pemahaman terhadap data. Visualisasi data

dilakukan dengan cara menyajikan data dalam bentuk grafik *scatter/bubble* seperti yang ditunjukkan pada Gambar 7 di bawah ini.



**Gambar 7.** Hasil Visualisasi

Berdasarkan Gambar 7 dapat dilihat bahwa visualisasi *scatter/bubble* terdiri dari sekumpulan titik atau "*bubble*" yang merepresentasikan satu atau lebih variabel dalam suatu dataset. Sumbu x dan y digunakan untuk membandingkan variabel-variabel tertentu, sedangkan ukuran setiap titik menunjukkan nilai variabel ketiga.

6. Diskusi

Berdasarkan hasil dari *performance* di atas menjelaskan pada data *performance* terdapat dua kategori yaitu kategori laris terjual sebagai *cluster* 0 yang bernilai sebesar 995 dan berjumlah sebanyak 121 produk. Sedangkan pada kategori kurang laris sebagai *cluster* 1 yang berilai sebesar 327 dan berjumlah sebanyak 2.279 produk. Dapat dilihat pada tabel di bawah ini.

- a. *Cluster* 0 dikategorikan sebagai laris
- b. *Cluster* 1 dikategorikasn sebagai kurang lari

**Tabel 2.** Hasil dari *Cluster* 0

Id	<i>Cluster</i>	<i>Brand Name</i>	<i>Product Division</i>	<i>Product Category</i>	Bulan	<i>Qty</i>	<i>Value</i>	Katego ri
15	<i>cluster</i> 0	Airwalk	Footwear	Casual Boots	Januari	33 3	6286815 0	Laris



17	cluster 0	Airwalk	Footwear	Lifestyle Shoes	Januari	32 0	6200575 0	Laris
72	cluster 0	Converse	Footwear	Casual Boots	Januari	10 1	4389477 0	Laris
73	cluster 0	Converse	Footwear	Lifestyle Shoes	Januari	10 1	4389477 0	Laris
10 7	cluster 0	Diadora	Footwear	Running	Januari	21 3	4245183 0	Laris
10 9	cluster 0	Diadora	Footwear	Sports Shoes	Januari	24 2	4862300 0	Laris
13 7	cluster 0	New Balance	Footwear	Lifestyle Shoes	Januari	14 3	4334537 0	Laris
13 8	cluster 0	New Balance	Footwear	Sports Shoes	Januari	32 0	1015981 30	Laris
17 5	cluster 0	Reebok	Footwear	Running	Januari	89	2702073 0	Laris
17 7	cluster 0	Reebok	Footwear	Sports Shoes	Januari	11 6	3627573 0	Laris
19 1	cluster 0	Skechers	Footwear	Lifestyle Shoes	Januari	10 4	3388212 0	Laris
19 2	cluster 0	Skechers	Footwear	Sports Shoes	Januari	13 2	5825098 0	Laris
21 5	cluster 0	Airwalk	Footwear	Casual Boots	Februa ri	17 9	3608306 0	Laris
21 7	cluster 0	Airwalk	Footwear	Lifestyle Shoes	Februa ri	18 9	3730982 0	Laris
30 7	cluster 0	Diadora	Footwear	Running	Februa ri	15 4	2784621 0	Laris
30 9	cluster 0	Diadora	Footwear	Sports Shoes	Februa ri	16 2	3030131 0	Laris
33 7	cluster 0	New Balance	Footwear	Lifestyle Shoes	Februa ri	10 6	3659527 0	Laris
33 8	cluster 0	New Balance	Footwear	Sports Shoes	Februa ri	18 6	6055280 0	Laris
37 5	cluster 0	Reebok	Footwear	Running	Februa ri	68 4	1194863 10	Laris
37 7	cluster 0	Reebok	Footwear	Sports Shoes	Februa ri	10 0	3050660 0	Laris
39 1	cluster 0	Skechers	Footwear	Lifestyle Shoes	Februa ri	15 5	4233233 0	Laris
39 2	cluster 0	Skechers	Footwear	Sports Shoes	Februa ri	20 7	8174068 0	Laris
...	...	...	...	...	...	...	...	...

Tabel 3. Hasil dari Cluster 1

Id	Cluster	Brand Name	Product Division	Product Category	Bulan	Qty	Value	Kategori
----	---------	------------	------------------	------------------	-------	-----	-------	----------

1	<i>cluster</i> 1	Airwalk	Accessories	Hat / Cap	Januar i	32	247620 0	Kurang Laris
2	<i>cluster</i> 1	Airwalk	Accessories	Lifestyle	Januar i	1	99000	Kurang Laris
3	<i>cluster</i> 1	Airwalk	Accessories	Socks	Januar i	1	99000	Kurang Laris
4	<i>cluster</i> 1	Airwalk	Accessories	Sports Accessories	Januar i	32	247620 0	Kurang Laris
5	<i>cluster</i> 1	Airwalk	Apparel	Jacket	Januar i	2	642600	Kurang Laris
6	<i>cluster</i> 1	Airwalk	Apparel	Pants	Januar i	2	502600	Kurang Laris
7	<i>cluster</i> 1	Airwalk	Apparel	Short	Januar i	2	418600	Kurang Laris
8	<i>cluster</i> 1	Airwalk	Apparel	Sports Lifestyle Apparel	Januar i	9	205457 0	Kurang Laris
9	<i>cluster</i> 1	Airwalk	Apparel	Tshirt Full Sleeve	Januar i	0	0	Kurang Laris
1 0	<i>cluster</i> 1	Airwalk	Apparel	Tshirt Short Sleeve	Januar i	3	490770	Kurang Laris
1 1	<i>cluster</i> 1	Airwalk	Bags	Backpack	Januar i	0	0	Kurang Laris
1 2	<i>cluster</i> 1	Airwalk	Bags	Shoulder Bag	Januar i	4	437800	Kurang Laris
1 3	<i>cluster</i> 1	Airwalk	Bags	Sports Bags	Januar i	40	477640 0	Kurang Laris
1 4	<i>cluster</i> 1	Airwalk	Bags	Waist Bag	Januar i	4	636000	Kurang Laris
1 6	<i>cluster</i> 1	Airwalk	Footwear	Lifestyle	Januar i	0	0	Kurang Laris
1 8	<i>cluster</i> 1	Airwalk	Footwear	Sandal	Januar i	19	284020 0	Kurang Laris
1 9	<i>cluster</i> 1	Airwalk	Footwear	Skate	Januar i	0	0	Kurang Laris
2 0	<i>cluster</i> 1	Airwalk	Footwear	Walking	Januar i	0	0	Kurang Laris
2 1	<i>cluster</i> 1	Astec	Accessories	Accessories	Januar i	1	99000	Kurang Laris
2 2	<i>cluster</i> 1	Astec	Accessories	Accessories Collection	Januar i	0	0	Kurang Laris
2 3	<i>cluster</i> 1	Astec	Accessories	Bottle	Januar i	0	0	Kurang Laris
2 4	<i>cluster</i> 1	Astec	Accessories	Health Equipment	Januar i	2	19000	Kurang Laris
...	...	...	...	...	...	...	...	...

#### 4. KESIMPULAN DAN SARAN

Berdasarkan hasil dari penelitian dapat disimpulkan bahwa dengan memanfaatkan data penjualan pada *Sport Station* yang telah dilakukan dari bulan Januari sampai Desember 2022 melalui algoritma *K-Means* dengan metode *Clustering* dapat membantu tingkat penjualan berdasarkan data yang telah diolah menggunakan operator-operator yang digunakan. Hasil pengelompokan terbagi menjadi 2 (dua) kelompok yaitu *cluster 0* dengan nilai 995 dan *cluster 1* dengan nilai 327 yang menunjukkan hasil dari pembentukan cluster. Pada *cluster 0* memiliki jumlah produk sebanyak 121 produk dengan kategori laris dan *cluster 1* memiliki jumlah sebanyak 2.279 produk dengan kategori kurang laris. Serta pada hasil *Devies Boulding Index* (DBI) dengan nilai yang paling mendekati 0 adalah K 2 menghasilkan nilai 0,10.

#### 5. DAFTAR PUSTAKA

- [1] M. Marsono, D. Saripurna, and M. Zunaidi, "Analisis Data Mining Pada Strategi Penjualan Produk PT Aquasolve Sanaria Dengan Menggunakan Metode K-Means Clustering," *J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD)*, vol. 4, no. 1, p. 127, 2021.
- [2] N. D. Rusida and Z. M. Noer, "Perancangan perangkat lunak bantu sistem penjualan berbasis aplikasi dekstop pada cafe instamie pangandaran," vol. 1, no. 1, pp. 341–350, 2018.
- [3] A. Harist N, I. R. Munthe, and A. P. Juledi, "Implementas[1] N. D. Rusida and Z. M. Noer, 'Perancangan perangkat lunak bantu sistem penjualan berbasis aplikasi dekstop pada cafe instamie pangandaran,' vol. 1, no. 1, pp. 341–350, 2018.i Data Mining Algoritma Apriori untuk Meningkatkan Penjualan," *J. Tek. Inform. UNIKA St. Thomas*, vol. 06, pp. 188–197, 2021.
- [4] Y. Dharma Putra, M. Sudarma, and I. B. A. Swamardika, "Clustering History Data Penjualan Menggunakan Algoritma K-Means," *Maj. Ilm. Teknol. Elektro*, vol. 20, no. 2, p. 195, 2021.
- [5] R. Supardi and I. Kanedi, "Implementasi Metode Algoritma K-Means Clustering pada Toko Eidelweis," *J. Teknol. Inf.*, vol. 4, no. 2, pp. 270–277, 2020.
- [6] A. Nugraha, N. Odi, and G. Dwilestari, "Penerapan Data Mining Merode K-Means Clustering untuk Analisa penjualan pada Toko Yana Sport," *J. Mhs. Tek. Inform.*, 2022.
- [7] D. Saripurna and W. Ristamaya, "Data Mining Untuk Pengelompokan Data Penjualan Cake Dengan Menggunakan Algoritma K-means Clustering Pada Jofie Bakery.," *J. Cyber Tech*, no. x, 2021.
- [8] F. A. Bramasta and R. Halilintar, "Penerapan Data Mining Untuk Menentukan Strategi Penjualan Toko Sepatu," *Pros. SEMNAS INOTEK ...*, pp. 236–241, 2021, [Online]. Available: <https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/1135%0Ahttps://proceeding.unpkediri.ac.id/index.php/inotek/article/download/1135/736>
- [9] A. Mardiana and Budiman, "Implementasi Metode K-Means untuk Clustering Produk yang Kurang Diminati Berdasarkan Data Penjualan".
- [10] R. R. Putra and C. Wadisman, "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means Implementation," *J. Inf. Teknol. Comput. Sci.*, vol. 63, no. 2, pp. 1–3, 2018.
- [11] M. Mulyadien and U. Enri, "Algoritma K-Means Untuk Pengelompokan Bantuan Langsung Tunai (BLT)," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 12, pp. 198–210, 2022.
- [12] N. Qorimah, K. Kunci, P. BPR Bank Cirebon, and N. Bayes, "Implementasi Metode Naïve Bayes Untuk Menentukan Minat Tabungan Pada Perumda BPR Bank Cirebon," vol. 2, no. 1, pp. 15–20, 2022.
- [13] S. Nurajizah, U. Bina, S. Informatika, R. Barat, and J. Barat, "Analisa Transaksi Penjualan Obat menggunakan Algoritma Apriori," vol. d, pp. 35–44, 2019.

- [14] K. Fatmawati and P. A. Windarto, “Data Mining : Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (DBD) Berdasarkan Provinsi,” *J. Comput. Eng. Syst. Sci.*, vol. 3, no. 2, pp. 173–178, 2018.
- [15] B. M. Islami, C. Sukmayadi, and T. N. Padilah, “Clustering Fasilitas Kesehatan Berdasarkan Kecamatan Di Karawang Dengan Algoritma K-Means,” *Clust. Fasilitas Kesehat. Berdasarkan Kec. Di Karawang Dengan Algorith. K-Means*, vol. 8, no. 25279777, 2021.