

## OPTIMALISASI PENJUALAN PRODUK MELALUI ANALISIS DATA DAN PENGELOMPOKAN PRODUK (STUDY KASUS HOPESPACE COFFE DENGAN ALGORITMA K-MEANS)

Dede Asmaya<sup>1</sup>, Bambang Irawan<sup>2</sup>, Agus Bahtiar<sup>3</sup>

<sup>1,2</sup> Teknik Informatika, STMIK IKMI Cirebon

<sup>3</sup> Sistem Informasi, STMIK IKMI Cirebon

Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135

asmayadede68@gmail.com

### ABSTRAK

Industri makanan dan minuman, terutama dalam bisnis kafe, mengalami persaingan yang semakin ketat dan peningkatan jumlah pesaing. Para pemilik usaha di sektor ini perlu terus berinovasi serta memperbaiki strategi mereka guna menjaga dan meningkatkan pangsa pasar mereka, serta meningkatkan efisiensi dalam operasional. Salah satu cara untuk mencapai tujuan ini adalah dengan melakukan analisis data penjualan secara detail dan menerapkan strategi pemasaran yang efisien, diharapkan dapat mendatangkan pengetahuan yang lebih mendalam terkait dengan tren penjualan harian di *HopeSpace Coffe*. Selain itu, melalui analisis yang dilakukan dengan bantuan *algoritma K-Means*, akan terungkap kelompok produk yang memiliki karakteristik serupa berdasarkan perilaku pembelian konsumen. Hal ini akan memberikan sudut pandang baru terkait strategi pemasaran yang lebih efisien dan memungkinkan pemilik bisnis untuk merencanakan langkah-langkah yang lebih cermat dalam mengelola operasional bisnis mereka. dari hasil nilai *Davies Bouldin Index* (DBI) yang optimal dari penerapan *algoritma K-Means* terhadap data penjualan dan pengelompokan produk kopi di *Hopespace Coffe* menghasilkan nilai DBI 0.035 dengan K4. Adapun Karakteristik dari *cluster* 0 yang terdiri dari 24 Item dengan karakteristik dari tahun 2021-2022 Dengan jenis kopi Robusta memiliki jumlah Cup yang terjual dari rata-rata 224 sampai 246, *cluster* 1 terdiri dari 24 Item dengan karakteristik dari tahun 2021-2022 Dengan jenis kopi Arabica memiliki jumlah Cup yang terjual dari rata-rata 413 sampai 491, *cluster* 2 terdiri dari 24 item dengan karakteristik dari tahun 2021-2022 dengan jenis kopi Kamojang memiliki jumlah Cup yang terjual dari rata-rata 89 sampai 134, *Cluter* 3 terdiri dari 24 item dengan karakteristik dari tahun 2021-2022 dengan jenis kopi Puntang memiliki jumlah Cup yang terjual dari rata-rata 34 sampai 116, dimana setiap kelompok memiliki karakteristik yang berbeda secara signifikan dari kelompok lainnya.

**Kata Kunci** : Penjualan Kafe , *algoritma K-Means* , Data Mining

### 1. PENDAHULUAN

Industri makanan dan minuman, terutama dalam bisnis kafe, mengalami persaingan yang semakin ketat dan peningkatan jumlah pesaing. Dalam situasi ini, para pemilik usaha di sektor ini perlu terus berinovasi serta memperbaiki strategi mereka guna menjaga dan meningkatkan pangsa pasar mereka, serta meningkatkan efisiensi dalam operasional. Salah satu cara untuk mencapai tujuan ini adalah dengan melakukan analisis data penjualan secara detail dan menerapkan strategi pemasaran yang efisien.

*HopeSpace Coffe*, sebagai bagian dari industri kafe, juga dihadapkan pada tantangan ini. Menyadari pola penjualan sehari-hari dan kebiasaan pembelian pelanggan merupakan faktor penting dalam pengelolaan kafe yang efisien. Selain itu, mengelompokkan produk kopi pembeli dapat mempermudah pengenalan tren konsumen dan perancangan strategi pemasaran yang lebih tepat dan efektif.

Menurut Jasenta Mikha Barus, Muhammad Zunaidi , Fifi Sonata. Dalam penelitiannya yang berjudul Implementasi Data Mining Untuk Klasterisasi Penjualan Makanan Dan Minuman Menggunakan Metode *K-Means* (Studi Kasus: *Bloum*

*Coffee*), Data Mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. Pada penelitian ini analisa data mining dilakukan dengan menggunakan metode *K-Means*. Dengan menggunakan metode ini, data-data yang telah didapatkan dapat dikelompokkan kedalam beberapa cluster berdasarkan kemiripandari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster dan yang memiliki karakteristik yang berbeda dikelompokkan dalam cluster yang lain yang memiliki karakteristik yang sama.

Menurut Anggi Octa Fadilah, Baenil Huda, Agustia Hananto, Tukino dalam penelitiannya yang berjudul Strategi Promosi untuk Meningkatkan Penjualan Kedai Kopi Desimal Menggunakan *Algoritma K-Medoids Clustering Anggi* , *K-medoids* merupakan sebuah algoritma clustering yang mirip dengan k-means, tetapi menggunakan medoid sebagai representasi dari setiap cluster bukan mean (rata-rata). Medoid adalah objek dalam cluster yang paling dekat ke rata-rata semua objek dalam cluster tersebut. Algoritma ini berfokus pada memberikan jarak

terpendek antara data yang terkait dalam setiap cluster [1]

Penelitian ini bertujuan untuk menginvestigasi pemanfaatan algoritma *K-Means* dalam mengkategorikan produk yang ada di *HopeSpace Caffe* berdasarkan pola pembelian yang serupa yang diperlihatkan oleh para pelanggan. Keberhasilan algoritma ini telah terbukti di berbagai konteks, terutama dalam menganalisis asosiasi dan mengelompokkan data. Melalui implementasi algoritma ini, diharapkan penelitian ini dapat memberikan wawasan yang lebih mendalam mengenai cara pengelompokan produk di *HopeSpace Caffe* dapat dilakukan berdasarkan kesamaan pola pembelian.

Hasil penelitian ini diharapkan akan memberikan manfaat ganda. Pertama, pemilik *HopeSpace Caffe* diharapkan dapat meningkatkan pemahaman mereka terkait pola penjualan harian dan kebiasaan belanja pelanggan. Hal ini diharapkan akan mendukung perencanaan persediaan, manajemen stok, dan perumusan strategi pemasaran yang lebih efektif. Kedua, dengan mengelompokkan produk berdasarkan perilaku pembelian pelanggan, *HopeSpace Caffe* berpotensi untuk merancang strategi pemasaran yang lebih terfokus dan meningkatkan pengalaman pelanggan secara keseluruhan.

Dalam era pertumbuhan pesat industri makanan dan minuman, penelitian ini akan memberikan petunjuk berharga bagi pemilik bisnis sejenis yang ingin meningkatkan efisiensi operasional mereka dan mencapai kesuksesan di tengah persaingan yang semakin sengit.

## 2. TINJAUAN PUSTAKA

### 2.1. Literature Review

Penelitian yang berjudul Penerapan Data Mining Untuk Clustering Menu Favorit Menggunakan Algoritma *K-Means* (Studi Kasus Kedai Expo) yang ditulis Setri Dwi Prasetiani. [2] Maraknya café juga dibarengi dengan tema dan tujuan tertentu. Misalnya seperti, dengan konsep live music, varian menu yang beragam mulai dari tradisional sampai moderen, harga yang terjangkau, dan juga fasilitas lainnya (wifi, kamar mandi ataupun musholla). Hal tersebut semakin membuktikan bahwa minat masyarakat yang tinggi terhadap keberadaan café karena semakin rame maka menunjukkan bahwa café tersebut sangat rekomen untuk dikunjungi kembali atau bahkan dijadikan tempat favorit.

Penelitian yang berjudul Strategi Promosi untuk Meningkatkan Penjualan Kedai Kopi Desimal Menggunakan Algoritma *K-Medoids Clustering* yang ditulis Anggi Octa Fadilah, Baenil Huda, Agustia Hananto, Tukino. [1] tentang cara berpromosi harus didasarkan pada data yang ada, data tersebut dapat menjadi aspek penting dalam menentukan pola penjualan dengan mengamati perilaku pembelian konsumen. Jika data dianalisis secara efektif, maka

dapat terungkap produk mana yang paling populer dan mana yang tidak laku, dengan begitu memungkinkan perbaikan item menu yang kurang baik untuk meningkatkan strategi promosi.

Penelitian yang berjudul Implementasi Data Mining Untuk Klasterisasi Penjualan Makanan Dan Minuman Menggunakan Metode *K-Means* (Studi Kasus: Bloum Coffee) yang ditulis oleh Jasenta Mikha Barus, Muhammad Zunaidi, Fifin Sonata. [3] Dengan menggunakan metode ini, data-data yang telah didapatkan dapat dikelompokkan kedalam beberapa cluster berdasarkan kemiripan dari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster dan yang memiliki karakteristik yang berbeda dikelompokkan dalam cluster yang lain yang memiliki karakteristik yang sama.

Penelitian yang berjudul Analisis Algoritma Apriori Dalam Menentukan Pola Pemesanan Konsumen Pada Ucokopi yang ditulis oleh Dewinta Marthadinata Sinaga, Wiranto Hernandez Sirait, Agus Perdana Windarto. [4].

Penelitian yang berjudul Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode *K-Means* yang ditulis oleh Fintri Indriyani, Ani Irfiani [5] Data mining merupakan penggalian data yang tersembunyi dari database. Proses Clustering merupakan suatu proses pengelompokan berdasarkan atas prinsip kesamaan kelas serta mengurangi kesamaan antar kelas. Berbagai algoritma dalam clustering telah dikembangkan untuk menghasilkan kinerja yang baik.

Penelitian oleh Hani Prastiwi, Jeny Pricilia, Errissya Raswir. [6] yang berjudul Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode *K-Means Clustering*. Metode *K-Means* sudah menjadi salah satu teknik data mining yang digunakan untuk merancang strategi persediaan barang yang efektif menggunakan data transaksi penjualan. Tujuan dari penelitian ini adalah untuk menerapkan algoritma *K-Means* dan data transaksi penjualan dari MM Glory sebagai contoh tipikal.

Penelitian oleh dengan judul Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma *K-Means* Dan *K-Medoids* dengan penulis Reza Gustrianda, Dadang Iskandar Mulyana. [7]

Penelitian oleh Sabrina Aulia Rahmah. [8] dengan judul penelitiannya Klasterisasi Pola penjualan Pestisida Menggunakan Metode *K-Means Clustering* (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja) Dalam penelitiannya Data Mining adalah suatu proses pencarian atau penggalian suatu informasi yang dilakukan di dalam sebuah data yang besar. Clustering (pengelompokan) merupakan salah satu metode yang dilakukan di dalam proses Data Mining.

Penelitian oleh Indah Nuryani. [9] Dedi Darwis dengan judul penelitiannya Analisis Clustering Pada

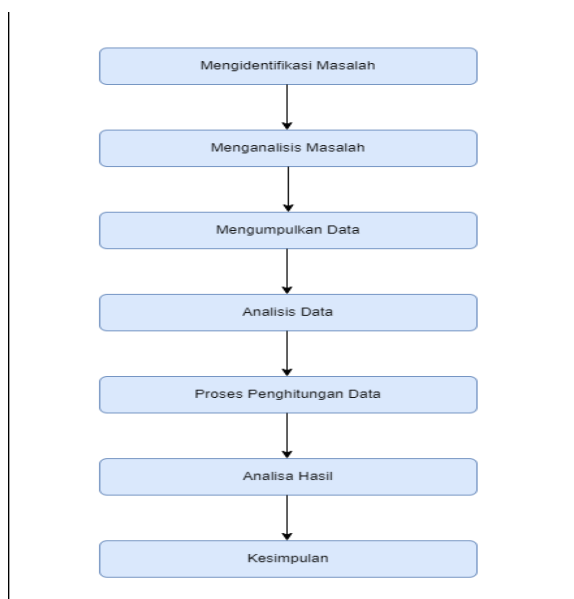
Pengguna Brand Hp Menggunakan Metode K-Means, dalam penelitiannya Pada pembahasan penelitian ini memilih metode clustering karena dapat mengelompokkan beberapa data/obyek ke dalam suatu cluster oleh karena itu dalam masing-masing cluster berisi data yang mirip. Dalam clustering menghasilkan obyek yang mirip berjarak dekat dalam satu cluster dan menghasilkan jarak antar cluster sejauh mungkin.

Penelitian yang dilakukan oleh Suhandio Handoko, Fauziah, Endah Tri Esti Handayani dengan judul penelitiannya Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering, Data mining yang biasa disebut sebagai knowledge discovery in database (KDD) merupakan kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola hubungan dalam himpunan data yang berukuran besar. Output dari data mining ini dapat digunakan untuk pengambilan keputusan di masa depan. Salah satu teknik yang dikenal dalam data mining yaitu clustering.[10]

### 3. METODE PENELITIAN

#### 3.1. Metode Penelitian

Metode penelitian yang digunakan adalah metode deskriptif analitis dengan pendekatan kuantitatif. Artinya dalam penelitian yang dilakukan penekanannya adalah pada analisis data numerik (angka), yang tujuannya untuk memperoleh gambaran yang jelas tentang keadaan dengan menyajikan, mengumpulkan dan menganalisis informasi yang diperoleh, sehingga menjadi informasi baru. informasi. cocok untuk analisis. Tahapan penelitian dapat dilihat pada gambar 3.1 dibawah ini.



Gambar 1. Tahapan Metode Penelitian

#### 3.2. Sumber Data

Sumber data yang digunakan penulis ialah data primer, dalam data primer ialah diperoleh secara langsung oleh peneliti melalui observasi, wawancara pribadi, atau eksperimen. Penelitian ini mengambil dataset secara langsung ke hopespace caffe yang ber alamat Jl. Soekarno Hatta, Bojongsari, Kec. Indramayu, Kabupaten Indramayu, Jawa Barat 45214 atau link maps <https://maps.app.goo.gl/KWx1HMQJGCDdC26e7>, dataset penjualan produk kopi di hopespace caffe memiliki 9 atribut dan 96 record.

#### 3.3. Populasi dan Sampel

Populasi dalam penelitian ini adalah data penjualan kopi di hopespace caffe di indramayu jawa barat yang dataset diambil langsung ke kafe tersebut, data tersebut diambil dari tahun 2021 sampai tahun 2022 dengan jumlah populasi pada penelitian ini berjumlah 96 data. Sampel dikelompokkan dalam 9 atribut yaitu tahun, bulan, jenis kopi, harga perbungkus, isi perbungkus, produksi, jumlah cup, harga unit.

#### 3.4. Teknik Pengumpulan Data

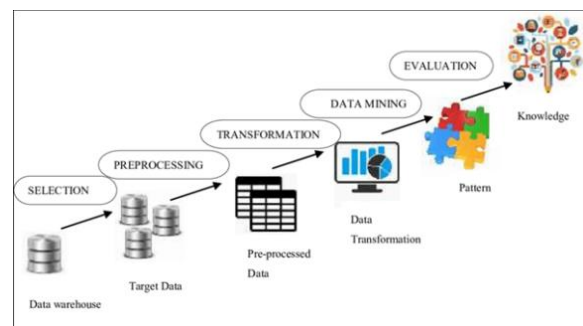
##### 3.4.1. Observasi

Observasi adalah mengemukakan observasi merupakan sebuah pengamatan secara langsung terhadap suatu objek yang ada di lingkungan baik itu yang sedang berlangsung atau masih dalam tahap yang meliputi berbagai aktivitas perhatian terhadap suatu kajian objek yang menggunakan penginderaan. Dan merupakan dari suatu tindakan yang dilakukan secara sengaja atau sadar dan juga sesuai urutan.[11]

##### 3.4.2. Wawancara

Wawancara merupakan interaksi tanya jawab antara dua pihak yang setidaknya satu diantara mereka memiliki tujuan serius yang telah ditetapkan, pelatihan wawancara pada dasarnya harus melibatkan simulasi intensif. Kegiatan pelatihan wawancara kerja ini dilaksanakan secara interaktif melalui seminar dan praktik secara langsung.[12].

#### 3.5. Teknik Analisis Data



Gambar 2. KDD (Knowledge Discovery in Database)

Dalam penelitian ini, metode yang digunakan adalah *Knowledge Discovery in Database (KDD)*. Langkah-langkah dan proses yang ada dalam *Knowledge Discovery in Database* diuraikan sebagai berikut:

a. *Selection*

Tahap pertama adalah tahap seleksi, dimana pemilihan data merupakan proses menganalisis data yang relevan dari database, karena tidak semua data dibutuhkan dalam proses data mining. Pemilihannya dilakukan dengan data penjualan kopi Hopespace tahun 2021-2022 yang berjumlah 96 data.

b. *Preprocessing*

Langkah selanjutnya adalah praprocessing, di mana data yang dikumpulkan disaring untuk mengetahui kesalahan atau kelalaiannya. Proses ini mencakup tindakan seperti penanganan data kosong, duplikat, atau hilang untuk memastikan bahwa data sesuai dengan format yang diinginkan.

c. *Transformation*

Langkah ini dilakukan untuk menafsirkan atau memodelkan data secara lebih efektif, dan sering kali melibatkan transformasi, agregasi variabel, normalisasi, atau pengelompokan.

d. *Data mining*

Tahapan ini merupakan tahap utama dari proses data mining. Banyaknya prediksi harus mengidentifikasi hubungan dan tren antar parameter dalam data.

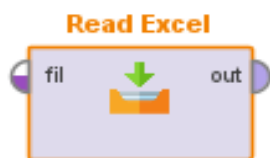
e. *Evaluation*

Interpretasi melibatkan analisis dan pemahaman hasil model regresi linier. Dalam tahap ini penulis mencoba menafsirkan arti algoritma k-means, menentukan nilai davies bouldin. Evaluasi menilai kualitas dan efisiensi algoritma k-means yang dibangun. Tujuannya adalah untuk memastikan bahwa algoritma tersebut dapat diandalkan dan memberikan hasil yang relevan.

4. HASIL DAN PEMBAHASAN

4.1. Data Selection

Tahapan yang pertama yaitu data *selection*. Pada tahap ini memasukkan operator *Read Excel* yang berfungsi untuk membaca data *excel* Penjualan kopi. Selanjutnya yaitu *import* data. Seperti pada Gambar.



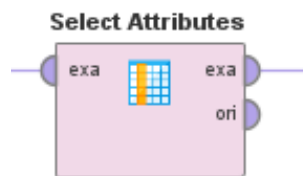
Gambar 3. Operator Read Excel

Parameter pada operator *Read Excel* menggunakan parameter *default*. Data yang berhasil di *import* menggunakan *read excel* akan muncul *preview* data.berikut hasil gambar 4 data *Read Excel*.

2021	Januari	Arabica	7200000	50 kg	500
2021	Januari	Robusta	3000000	30 kg	300
2021	Januari	Kamojang	1600000	15kg	150
2021	Januari	Puntang	1400000	15 kg	150
2021	Februari	Arabica	7200000	50 kg	500
2021	Februari	Robusta	3000000	30 kg	300
2021	Februari	Kamojang	1600000	15kg	150
2021	Februari	Puntang	1400000	15 kg	150
2021	Maret	Arabica	7200000	50 kg	500
2021	Maret	Robusta	3000000	30 kg	300
2021	Maret	Kamojang	1600000	15kg	150
2021	Maret	Puntang	1400000	15 kg	150

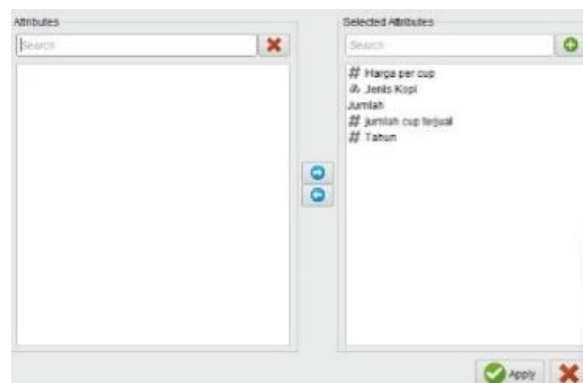
Gambar 4. Data Read Excel

Pada langkah selanjutnya, *operator Select Attributes* dimasukkan, untuk memilih atribut mana yang dipilih . Berikut Gambar 5 *select attributes*.



Gambar 5. Operator Select Attributes

Parameter pada operator *Select Attributes* hanya mengganti *attribute filter type a subset*. Hasil *preview* pada *Select Attributes*. Lalu pilih atribut yang ingin digunakan Berikut dapat dilihat pada gambar 6 *Sellect Attributes*.



Gambar 6. Select Attributes

Dari total 9 atribut yaitu tahun, bulan jenis kopi, harga/bks, isi/bks,produksi cup, jumlah cup terjual, harga per cup, dan jumlah.dipilih menjadi 6 atribut yaitu tahun, bulan, harga per cup, jenis kopi, jumlah, jumlah cup terjual. Pemilihan atribut ini didasarkan pada relevansi informasi terkait dengan analisis yang ingin dilakukan.. Kombinasi dari keenam atribut ini dapat memberikan wawasan yang mendalam tentang penjualan produk kopi dalam suatu konteks tertentu. yang tertera pada tabel 1 dibawah ini.

Tabel 1. Hasil Select Attributes

Atribut Yang Tidak Digunakan	Atribut Yang Digunakan
Harga/bks	Bulan
Isi/bks	harga per cup
Produksi cup	jenis kopi
	Total pendapatan
	jumlah cup terjual
	tahun

Model proses pada *rapidminer* dilangkah *Selection* tampak pada gambar 7 terkait model proses langkah *selection* di *rapidminer*.

Gambar 7. Model proses langkah Selection di Rapidminer

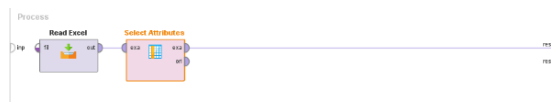
4.2. Preprocessing

Pada tahap *Pre-Processing*, langkah-langkah dilakukan untuk membersihkan dan menyiapkan data sebelum dilakukan analisis atau pemodelan. Tujuan utamanya adalah memastikan bahwa data yang digunakan memiliki kualitas yang baik, akurat, dan sesuai dengan kebutuhan analisis yang akan dilakukan. Proses ini mencakup langkah-langkah seperti menangani data yang kosong, menghapus data duplikat, serta mengatasi nilai yang hilang agar sesuai dengan format yang diinginkan. Data yang telah dibersihkan menjadi lebih mudah untuk dianalisis dan dapat menghasilkan hasil yang lebih akurat. Salah satu operasi yang umum dilakukan dalam proses ini adalah menggunakan operator "*Replace Missing Values*" yang bertujuan untuk menggantikan atau mengisi nilai yang hilang dalam dataset dengan nilai yang valid atau dapat diproses lebih lanjut. Dengan menggunakan operator ini, kita dapat menentukan metode pengisian nilai yang hilang, seperti pengisian dengan nilai rata-rata, median, modus, atau menggunakan teknik lain yang sesuai dengan karakteristik data. Dalam konteks ini, *Replace Missing Values* digunakan untuk mengatasi masalah data yang tidak lengkap, memastikan bahwa data yang digunakan untuk analisis atau pemodelan tidak memiliki nilai yang hilang, sehingga memungkinkan analisis yang lebih tepat dan akurat. Gambar 8 adalah operator *replace missing value*.



Gambar 8. Operator Missing Value

Setelah dilakukan pembersihan data oleh operator *Replace Missing Values* memiliki hasil seperti pada gambar 4.6 yang menghasilkan tidak adanya perubahan data dan masih tetap sama berjumlah 96 data dari tahun 2021 sampai tahun 2022, dikarenakan data yang digunakan merupakan data lengkap tidak ada data yang kosong, ganda ataupun *missing* seperti pada gambar 9 *Missing Values*.



Gambar 9. Missing Values

4.3. Transformation

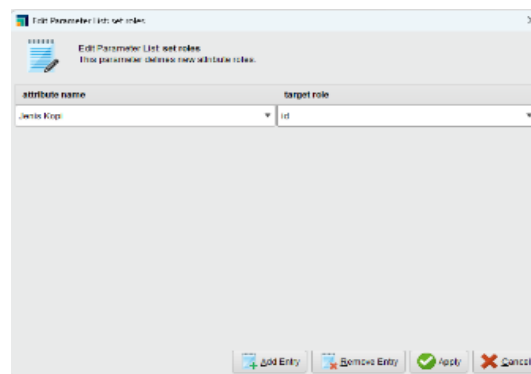
Pada tahap berikutnya, dilakukan proses transformasi data dengan menggunakan parameter "*Set Role*" dan "*Normalize*". Transformasi ini bertujuan untuk menetapkan peran atribut dan mengubah skala nilai variabel. Dengan menetapkan peran atribut, data menjadi lebih terstruktur dan siap untuk dianalisis lebih lanjut. Normalisasi, yang mencakup pengubahan skala nilai variabel ke rentang antara 0 dan 1, dilakukan untuk memastikan bahwa semua variabel memberikan dampak yang seimbang pada analisis.

Operator *set role* dapat ditampilkan pada gambar 10 dibawah ini



Gambar 10. Operator Set Role

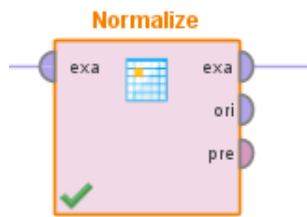
Proses "*Set Rol*" merupakan tahap penting dalam persiapan data sebelum menjalani analisis lebih lanjut. Pada tahap ini, peneliti memanfaatkan *set rol* dengan tujuan menetapkan atribut Jenis Kopi sebagai ID. Dengan menetapkan atribut Jenis Kopi sebagai ID, data menjadi lebih terstruktur dan dapat diakses dengan lebih mudah dalam tahap analisis berikutnya. Proses *Set Role* ditampilkan pada gambar 11



Gambar 11. Tampilan Set Role

#### 4.4. Normalize

Tahap selanjutnya yaitu dengan menggunakan operator Normalize. operator Normalize ditampilkan pada gambar 12



Gambar 12. Operator Normalize

Dengan menggunakan operator normalize, tahap transformasi data dapat dijalankan dengan lebih terstruktur, sehingga data siap untuk dianalisis lebih mendalam. Normalisasi, sebagai metode transformasi data dilakukan untuk mengadaptasi skala nilai variabel, tujuannya untuk menjamin bahwa semua variabel memiliki pengaruh yang seimbang pada proses analisis. Data normalisasi ditampilkan pada gambar 4.11.

Row No.	Tahun	Jenis Kopi	cluster	Bulan	Harga per cup	jumlah cup L.	Total Penda...
1	2021	Arabica	cluster_0	Januari	1.177	1.711	1.731
2	2021	Robusta	cluster_2	Januari	-0.168	0.459	0.337
3	2021	Kamojang	cluster_1	Januari	0.504	-0.720	-0.693
4	2021	Puntang	cluster_3	Januari	-1.513	-0.760	-0.553
5	2021	Arabica	cluster_0	Februari	1.177	1.724	1.744
6	2021	Robusta	cluster_2	Februari	-0.168	0.373	0.260
7	2021	Kamojang	cluster_1	Februari	0.504	-0.793	-0.761
8	2021	Puntang	cluster_3	Februari	-1.513	-0.932	-0.966
9	2021	Arabica	cluster_0	Maret	1.177	1.459	1.484
10	2021	Robusta	cluster_2	Maret	-0.168	-0.045	-0.117
11	2021	Kamojang	cluster_1	Maret	0.504	-0.879	-0.843
12	2021	Puntang	cluster_3	Maret	-1.513	-1.303	-1.270
13	2021	Arabica	cluster_0	April	1.177	1.207	1.236
14	2021	robusta	cluster_2	April	-0.168	0.101	0.015

ExampleSet (96 examples, 4 special attributes, 3 regular attributes)

Gambar 13. Hasil Normalisasi Data

#### 4.5. Data Mining

Tahap selanjutnya memasukkan operator Clustering yang berfungsi untuk menentukan Cluster atau pengelompokan data. Penelitian ini menggunakan algoritma K-Means dan menggunakan 4 Cluster max perulangan 10. Seperti pada gambar 4.12 operator clustering menggunakan k-means.

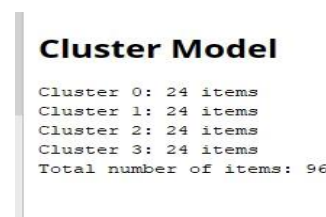


Gambar 14. Operator Clustering

Parameter pada operator Clustering K-Means dimana diketahui  $K = 4$ ,  $max\ run = 10$  kali putaran,  $tipe\ measure = Numerical\ Measure$  yang  $Euclidean\ Distance$  dan  $max\ optimal\ steps = 100$ .

Hasil operator Clustering K-Means diperoleh 4 Cluster yaitu Cluster 0 sebanyak 24 items, Cluster 1

sebanyak 24 items, Cluster 2 sebanyak 24 items, Cluster 3 sebanyak 24 items, dengan total items sebanyak 96 items. Seperti pada gambar 4.13 Cluster Model.



Gambar 15. Cluster Model

Row No.	Tahun	Jenis Kopi	cluster	Bulan	Harga per cup	jumlah cup L.	Total Penda...
1	2021	Arabica	cluster_1	Januari	1.177	1.711	1.731
2	2021	Robusta	cluster_2	Januari	-0.168	0.459	0.337
3	2021	Kamojang	cluster_3	Januari	0.504	-0.720	-0.693
4	2021	Puntang	cluster_4	Januari	-1.513	-0.760	-0.553
5	2021	Arabica	cluster_1	Februari	1.177	1.724	1.744
6	2021	Robusta	cluster_2	Februari	-0.168	0.373	0.260
7	2021	Kamojang	cluster_3	Februari	0.504	-0.793	-0.761
8	2021	Puntang	cluster_4	Februari	-1.513	-0.932	-0.966
9	2021	Arabica	cluster_1	Maret	1.177	1.459	1.484
10	2021	Robusta	cluster_0	Maret	-0.168	-0.045	-0.117
11	2021	Kamojang	cluster_3	Maret	0.504	-0.879	-0.843
12	2021	Puntang	cluster_4	Maret	-1.513	-1.303	-1.270
13	2021	Arabica	cluster_1	April	1.177	1.207	1.236
14	2021	robusta	cluster_0	April	-0.168	0.101	0.015

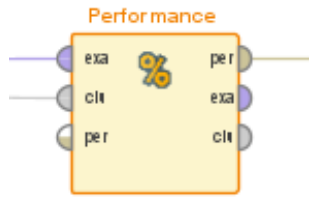
ExampleSet (96 examples, 4 special attributes, 3 regular attributes)

Gambar 16. Hasil Operator Clustering menggunakan K-Means

Setelah proses pengelompokan data menggunakan algoritma clustering seperti K-Means, langkah selanjutnya adalah melakukan validasi terhadap cluster yang dihasilkan. Salah satu metode validasi yang umum digunakan adalah dengan menggunakan performa jarak cluster (Cluster distance performance) untuk mengevaluasi seberapa baik cluster yang dihasilkan. Pada tahap ini, evaluasi dilakukan menggunakan model performance untuk menghitung nilai Davies Bouldin Index (DBI). DBI adalah salah satu metrik evaluasi yang digunakan untuk mengukur kualitas klaster berdasarkan jarak antara Cluster, yang lebih baik ketika nilainya mendekati 0.

Tujuan dari penggunaan operator ini adalah untuk mengidentifikasi nilai k yang paling optimal dalam klasterisasi. Fungsi dari operator ini adalah melakukan evaluasi DBI menggunakan nilai k yang ditentukan secara acak atau diuji secara berulang, dengan tujuan untuk menentukan nilai k yang paling optimal. Dengan melakukan perhitungan DBI pada berbagai nilai k, operator tersebut membantu dalam mengevaluasi performa cluster yang dihasilkan oleh model K-Means. Hasil dari evaluasi ini memberikan pemahaman tentang seberapa baik pembagian Cluster yang dihasilkan oleh algoritme K-Means dengan nilai k tertentu. Nilai DBI yang mendekati 0 menunjukkan bahwa pembagian cluster memiliki jarak antar Cluster yang optimal, menunjukkan bahwa nilai k tersebut mungkin merupakan pilihan yang lebih baik untuk mewakili struktur data yang ada seperti pada gambar 16 operator performance.





Gambar 17. Operator Performance

Hasil dari nilai pada *Davies Bouldin Index* yang di dapat dari penelitian ini dan yang hasilnya paling mendekati 0 yaitu 0.035 , menggunakan nilai k = 4 dengan maksimal perulangan 10 kali seperti gambar 17 hasil operator *performance*.



Gambar 18. Hasil Operator Performance

Untuk mendapatkan nilai k yang paling optimal, dilakukan 10 kali percobaan. Nilai k yang paling optimal adalah nilai k dengan hasil DBI mendekati 0. Dari 4 kali percobaan, nilai k yang paling optimal terdapat pada k = 4 dengan hasil DBI nya adalah 0.035. Hasil percobaan yang telah dilakukan dapat dilihat pada tabel 4.2. hasil DBI Model proses pada *rapidminer* semua langkah tampak pada gambar 18 Model proses semua langkah di *Rapidminer*.



Gambar 19. Model proses semua langkah di Rapidminer

**4.6. Hasil nilai *Davies Bouldin Index*(DBi) yang optimal dari penerapan algoritma *K-Means* terhadap Data Penjualan dan Pengelompokan Produk Kopi di *Hopespace Coffe* Dengan Penerapan Algoritma *K-Means***

Tabel 2. Hasil Nilai DBI

No.	Cluster	Davies Bouldin Indeks
1	K2	0.249
2	K3	0.302
3	K4	0.035
4	K5	0.119
5	K6	0.161
6	K7	0.241
7	K8	0.148
8	K9	0.235
9	K10	0.295

**4.7. Pembahasan**

Dari hasil analis data penjualan dan pengelompokan produk kopi di *Hopespace Coffe* dengan *algoritma k-means*, *algoritma K-Means* dapat membantu mengidentifikasi tren konsumen terhadap produk kopi di *HopeSpace Coffee* dengan cara Segmentasi konsumen, *Algoritma k-means* bisa digunakan untuk menentukan *clustering* produk kopi dengan langkah langkah pemilihan fitur, pemrosesan data, penentuan jumlah cluster, inialisasi pusat *cluster*, penetapan, dan evaluasi cluster. dan hasil analisis penjualan dan pengelompokan produk kopi dapat memberikan wawasan strategis untuk meningkatkan kinerja bisnis dan profitabilitas *HopeSpace Coffee* adalah dengan cara pemahaman konsumen, penyesuaian harga, strategi pemasaran yang lebih tepat dan efisiensi.

Dilihat dari hasil nilai *Davies Bouldin Index* (DBI) yang optimal dari penerapan *algoritma K-Means* terhadap data penjualan dan pengelompokan produk kopi di *Hopespace Coffe* menghasilkan nilai DBI 0.035 dengan K4. Adapun Karakteristik dari *cluster* 0 yang terdiri dari 24 Item dengan karakteristik dari tahun 2021-2022 Dengan jenis kopi Robusta memiliki jumlah Cup yang terjual dari rata-rata 224 sampai 246, *cluster* 1 terdiri dari 24 Item dengan karakteristik dari tahun 2021-2022 Dengan jenis kopi Arabica memiliki jumlah Cup yang terjual dari rata-rata 413 sampai 491, *cluster* 2 terdiri dari 24 item dengan karakteristik dari tahun 2021-2022 dengan jenis kopi Kamojang memiliki jumlah Cup yang terjual dari rata-rata 89 sampai 134, *Cluter* 3 terdiri dari 24 item dengan karakteristik dari tahun 2021-2022 dengan jenis kopi Puntang memiliki jumlah Cup yang terjual dari rata-rata 34 sampai 116, dimana setiap kelompok memiliki karakteristik yang berbeda secara signifikan dari kelompok lainnya.

Ini berarti *algoritma K-Means* berhasil mengidentifikasi pola pembelian yang konsisten dan berbeda diantara kelompok produk kopi yang ada di *Hopespace Coffe*. dengan demikian hasil ini memberikan keyakinan bahwa segmentasi produk kopi yang dilakukan berdasarkan *Algoritma K-Means* memiliki validasi yang tinggi dan dapat diandalkan untuk digunakan dalam pengambilan Keputusan strategis terkait pemasaran pengelola stok dan pengembangan produk di *Hopespace Coffee*.

**5. KESIMPULAN**

Dari hasil penelitian dengan pengelompokan tingkat penjualan jenis kopi di *hopespace cafee* yang telah dilakukan oleh penulis , dapat disimpulkan bahwa : Penelitian ini melakukan 10 kali percobaan dengan k yang berbeda. Dan diketahui dari 10 kali percobaan, nilai k yang paling optimal terdapat pada k = 4 dengan hasil DBI nya adalah 0. 035. *Cluster* yang dihasilkan memperoleh sebanyak 10 *Cluster* Setiap klaster memiliki karakteristiknya sendiri dalam tingkat penjualan yang berbeda-beda tergantung jenis kopi yang dijual. *Cluster* tersebut menunjukkan

sejumlah pola yang paling banyak diminati oleh pengunjung *hopespace cafee* terhadap menu yang dijual. Dari hasil *cluster* yang dilakukan penulis terdapat beberapa jenis kopi yang paling diminati oleh pengunjung diantaranya jenis kopi Arabica yang menjadi *best seller* di *hopespace cafee*.

Saran dari penelitian ini adalah Analisis Lebih Mendalam terhadap *Cluster* yang Ditemukan: Lakukan analisis lebih lanjut terhadap karakteristik masing-masing *Cluster*. Perhatikan tidak hanya tingkat penjualan, tetapi juga faktor lain seperti meningkatkan penjualan kopi dan berbagai menu lainnya di *hopespace cafee*. Eksplorasi Lebih Lanjut terhadap Variabel Lain: penambahan variasi di kopi untuk mengeksplorasi rasa yang mungkin memengaruhi tingkat penjualan. Misalnya, penambahan jenis kopi, menambahkan varian rasa yang bervariasi Hal ini dapat membantu dalam peningkatan penjualan di *hopespace cafee*.

#### DAFTAR PUSTAKA

- [1] A. O. Fadilah, B. Huda, A. Hananto, dan Tukino, "Strategi Promosi untuk Meningkatkan Penjualan Kedai Kopi Desimal Menggunakan Algoritma K-Medoids Clustering," *JURIKOM (Jurnal Ris. Komputer)*, vol. 10, no. 1, hal. 2407–389, 2023, doi: 10.30865/jurikom.v10i1.5561.
- [2] S. D. Prasetiani dan N. Rochmawati, "Penerapan Data Mining Untuk Clustering Menu Favorit Menggunakan Algoritma K-Means (Studi Kasus Kedai Expo)," *J. Informatics Comput. Sci.*, vol. 3, no. 03, hal. 278–286, 2022, doi: 10.26740/jinacs.v3n03.p278-286.
- [3] J. Mikha Barus, M. Zunaidi, dan F. Sonata, "Implementasi Data Mining Untuk Klasterisasi Penjualan Makanan Dan Minuman Menggunakan Metode K-Means (Studi Kasus: Bloum Coffee) STMIK Triguna Dharma \*\* Program Studi Sistem Informasi, STMIK Triguna Dharma \*\*\* Program Studi Manajemen Informatika, STMIK Tr," *J. CyberTech*, vol. 3, no. 6, hal. 1195–1207, 2020, [Daring]. Tersedia pada: <https://ojs.trigunadharma.ac.id/>
- [4] D. M. Sinaga, W. H. Sirait, dan A. P. Windarto, "Analisis Algoritma Apriori Dalam Menentukan Pola Pemesanan Konsumen Pada Ucokopi," *J. Informatics Manag. Inf. Technol.*, vol. 1, no. 2, hal. 68–73, 2021, doi: 10.47065/jimat.v1i2.105.
- [5] F. Indriyani dan E. Irfiani, "Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means," *JUITA J. Inform.*, vol. 7, no. 2, hal. 109, 2019, doi: 10.30595/juita.v7i2.5529.
- [6] H. Prastiwi, Jeny Pricilia, dan Errissya Rasywir, "Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering," *J. Inform. Dan Rekayasa Komputer (JAKAKOM)*, vol. 2, no. 1, hal. 141–148, 2022, doi: 10.33998/jakakom.2022.2.1.34.
- [7] R. Gustrinda dan D. I. Mulyana, "Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma K-Means Dan K-Medoids," *J. Media Inform. Budidarma*, vol. 6, no. 1, hal. 27, 2022, doi: 10.30865/mib.v6i1.3294.
- [8] S. Aulia, "Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means Clustering (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja)," *Djtechno J. Teknol. Inf.*, vol. 1, no. 1, hal. 1–5, 2021, doi: 10.46576/djtechno.v1i1.964.
- [9] I. Nuryani dan D. Darwis, "Analisis Clustering Pada Pengguna Brand Hp Menggunakan Metode K-Means," *Pros. Semin. Nas. Ilmu Komput.*, vol. 1, no. 1, hal. 2021, 2021.
- [10] S. Handoko, F. Fauziah, dan ..., "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering," ... *Ilm. Teknol. dan ...*, 2020.
- [11] H. Hasanah, "TEKNIK-TEKNIK OBSERVASI (Sebuah Alternatif Metode Pengumpulan Data Kualitatif Ilmu-ilmu Sosial)," *At-Taqaddum*, vol. 8, no. 1, hal. 21, 2017, doi: 10.21580/at.v8i1.1163.
- [12] S. Suprayogi, D. Puspita, E. A. D. Putra, dan M. R. Mulia, "Pelatihan Wawancara Kerja Bagi Anggota Karang Taruna Satya Wira Bhakti Lampung Timur," *Community Dev. J. J. Pengabd. Masy.*, vol. 3, no. 1, hal. 356–363, 2022, doi: 10.31004/cdj.v3i1.4494.