

ANALISIS PENGELOMPOKAN PENJUALAN MENGGUNAKAN METODE K-MEANS PADA BISNIS CENTER SMK WAHIDIN

Muhamad Aji Dian Permana ¹, Martanto ², Umi Hayati ³

^{1,3} Program Studi Teknik Informatika, STMIK IKMI Cirebon

² Program Studi Manajemen Informatika, STMIK IKMI Cirebon

Jalan Perjuangan No.10B, Karyamulya Cirebon, Indonesia

ajidianpermana2@gmail.com, martantomusijo@gmail.com, umi.haya41@gmail.com

ABSTRAK

Bisnis Center SMK Wahidin adalah entitas bisnis yang beroperasi dalam konteks pendidikan, yang menyediakan berbagai produk dan layanan kepada komunitas sekolah dan orang tua siswa. Bisnis Center ini menjual berbagai barang, termasuk seragam sekolah, peralatan pendidikan, dan souvenir. Metode K-Means adalah salah satu teknik clustering yang populer dalam analisis data, yang dapat membantu mengidentifikasi pola-pola tersembunyi dalam data penjualan Bisnis Center SMK Wahidin. Dalam penelitian ini, kami mengumpulkan data penjualan dari 1 Juli 2023 sampai 31 Oktober 2023, yang mencakup informasi tentang jenis produk yang dijual, harga produk serta jumlah penjualan. Data ini kemudian digunakan sebagai dasar untuk analisis pengelompokan penjualan dengan metode K-Means. Penelitian ini akan merinci bagaimana hasil analisis ini dapat digunakan sebagai panduan bagi Bisnis Center di sekolah-sekolah lain yang memiliki unit bisnis serupa. Analisis ini dapat menjadi model bagi upaya pengelompokan penjualan dan manajemen operasional yang lebih efektif dalam lingkungan sekolah. Sebagai langkah selanjutnya, penelitian ini dapat diperluas dengan mempertimbangkan teknik clustering lainnya, serta pengumpulan data yang lebih luas dan waktu observasi yang lebih panjang untuk analisis yang lebih mendalam. penelitian ini memberikan landasan bagi pengembangan strategi bisnis yang lebih cerdas dan efisien di Bisnis Center SMK Wahidin serta unit bisnis serupa dalam lingkungan pendidikan.

Kata kunci : *Bisnis Center SMK Wahidin, Analisis Penjualan, Metode K-Means, Strategi Pemasaran, SMK Wahidin.*

1. PENDAHULUAN

Bisnis Center SMK Wahidin merupakan sebuah entitas yang memiliki peran penting dalam memfasilitasi aktivitas penjualan di lingkungan sekolah. Sebagai bagian dari upaya pengembangan keterampilan siswa, Bisnis Center ini telah aktif dalam menjalankan berbagai kegiatan pemasaran dan penjualan. Dalam era persaingan bisnis yang semakin ketat, perlu adanya analisis yang mendalam terkait pola penjualan yang ada. Analisis ini bertujuan untuk mengoptimalkan strategi pemasaran, meningkatkan efisiensi operasional, dan pada akhirnya, meningkatkan pendapatan sekolah. Oleh karena itu, penelitian ini bertujuan untuk menerapkan metode pengelompokan data, khususnya metode *K-Means*, untuk mengidentifikasi pola penjualan yang ada di Bisnis Center SMK Wahidin.

Dalam melakukan analisis pengelompokan penjualan, terdapat beberapa permasalahan yang perlu diidentifikasi. Pertama, Bisnis Center SMK Wahidin memiliki beragam produk yang dijual, termasuk snack, minuman, dan barang-barang sekolah. Pola penjualan yang beragam ini mungkin memerlukan pendekatan yang berbeda dalam strategi pemasaran. Kedua, penjualan Bisnis Center ini terdistribusi pada berbagai waktu dan hari yang berbeda, sehingga penting untuk mengidentifikasi pola waktu terbaik untuk masing-masing produk. Ketiga, kebijakan harga juga beragam, dan perlu dikaji lebih lanjut agar dapat menentukan strategi harga yang paling efektif. Keempat, terdapat variasi dalam preferensi pelanggan yang mungkin

perlu diperhatikan dalam pengelompokan penjualan. Kelima, analisis terkait performa individu dalam tim penjualan perlu diperhitungkan untuk mendukung pengembangan keterampilan siswa. Keseluruhan, pemahaman mendalam tentang karakteristik penjualan Bisnis Center ini menjadi dasar penting untuk pengambilan keputusan yang lebih baik dan terinformasi di masa depan. Manajemen stok barang yang tidak tepat dapat menyebabkan ketidakseimbangan antara penawaran dan permintaan produk, yang berpotensi mengakibatkan kerugian finansial dan kekecewaan pelanggan. Oleh karena itu, analisis pengelompokan penjualan dengan metode *K-Means* diharapkan dapat memberikan solusi yang bermanfaat bagi Bisnis Center SMK Wahidin.

Dari hasil penelitian yang dilakukan oleh [1], Menguraikan *text mining* pada barang elektronik yang dijual di bisnis berbasis web Shopee dengan menggunakan strategi perhitungan *K-Means* dengan memanfaatkan bahasa pemrograman *Python*. Ulasan teks, angka penjualan, dan peringkat bintang digunakan sebagai data. Informasi pemeriksaan yang didapat dari investigasi pesan audit item dengan menggunakan *wordcloud* untuk item ponsel yang dievaluasi rendah menunjukkan bahwa informasi survei Pasar Shopee Indonesia menunjukkan bahwa ponsel yang dievaluasi rendah dan yang dievaluasi tinggi pada umumnya memiliki desain kata yang serupa, khususnya kata-kata yang sering muncul adalah kata-kata yang tidak memihak dan bagus, sedangkan kata-kata dengan makna dasar yang

disesalkan cenderung tidak terlalu banyak. Sedangkan kata-kata yang sering muncul adalah produk, konsisten, cepat, kirim, dan hebat dengan nilai tingkat yang berbeda sebesar 92%. Sedangkan konsekuensi dari proses wordcloud biaya menengah-signifikan menunjukkan bahwa kata-kata yang sering muncul adalah kata-kata (kirim, cepat, dan hebat) dengan nilai ketepatan sebesar 94%. Selain itu, jika dilihat dari grafik hasil pengelompokan informasi k-means, terlihat bahwa proyeksi pemasaran dari 0 hingga 1000 sampai pada nilai rating bintang yang paling tinggi dan penawaran dengan nilai rating bintang yang paling rendah berada di kisaran 1500 dan 2000 atau lebih. Selanjutnya menurut penelitian yang dilakukan oleh [2], menyelesaikan seluruh tahap pengelompokan data. Di toko obat M23, dalam beberapa kasus terjadi kekurangan atau kelebihan obat. Untuk mengatasi masalah ini, strategi penggalian informasi diterapkan dengan memecah penggunaan obat untuk menghasilkan data yang dapat digunakan untuk pengendalian persediaan dan perencanaan obat. Teknik yang digunakan dalam penelitian ini adalah strategi *K-Means*. Strategi pengelompokan *K-Means* bermaksud untuk mengelompokkan informasi yang memiliki atribut yang sama ke dalam kelompok yang sama dan informasi yang memiliki kualitas yang berbeda dikumpulkan ke dalam kelompok yang berbeda. menggunakan metode *Davies Bouldin Index* (DBI) untuk menentukan jumlah *cluster* yang ideal. Hasil dari penelitian ini mengkonfirmasi bahwa jumlah kelompok terbaik adalah 2 kelompok, pengumpulan informasi obat terdiri dari kelompok 1 dengan penggunaan obat rendah yang terdiri dari 144 jenis obat dan kelompok 2 dengan penggunaan obat tinggi yang terdiri dari 6 jenis obat.

Penelitian ini bertujuan untuk mengidentifikasi pola penjualan di Bisnis Center SMK Wahidin dengan menggunakan metode *K-Means*, sehingga dapat membantu sekolah dalam merancang strategi pemasaran yang lebih efektif. Dengan menganalisis data penjualan, penelitian ini juga bertujuan untuk meningkatkan pemahaman terhadap preferensi pelanggan, waktu penjualan terbaik, dan strategi harga yang optimal. Penelitian ini memiliki signifikansi dalam mendukung perkembangan Bisnis Center, serta dapat menjadi acuan bagi sekolah-sekolah lain yang memiliki bisnis serupa.

Dalam melaksanakan penelitian ini, kami akan menggunakan metode pendekatan berbasis data, yaitu metode *K-Means*. Metode *K-Means* adalah sebuah teknik pengelompokan data yang akan memungkinkan kami untuk mengelompokkan penjualan Bisnis Center SMK Wahidin ke dalam kelompok-kelompok yang memiliki karakteristik serupa. Data penjualan yang akan kami gunakan melibatkan variabel seperti jenis produk, waktu penjualan, harga, dan preferensi pelanggan. Proses analisis akan dimulai dengan pra-pemrosesan data, termasuk normalisasi data dan pemilihan jumlah kluster yang optimal. Selanjutnya, kami akan menerapkan algoritma *K-Means* untuk

mengelompokkan data penjualan ke dalam kelompok-kelompok yang relevan. Hasil analisis ini akan membantu kami dalam merumuskan rekomendasi strategi pemasaran yang lebih efektif dan mendukung pengembangan Bisnis Center SMK Wahidin secara keseluruhan.

Implikasi hasil penelitian dari analisis pengelompokan penjualan di Bisnis Center SMK Wahidin menggunakan metode *K-Means* sangat relevan dan berdampak luas. Pertama, penelitian ini akan memberikan wawasan yang mendalam tentang preferensi pelanggan dan pola penjualan Bisnis Center, yang dapat digunakan untuk merancang strategi pemasaran yang lebih efektif. Hasil analisis ini dapat memungkinkan Bisnis Center untuk menyusun penawaran produk yang lebih sesuai dengan kebutuhan pelanggan. Kedua, hasil penelitian ini memiliki implikasi finansial yang signifikan. Dengan pemahaman yang lebih baik tentang waktu penjualan yang optimal dan strategi harga yang tepat, Bisnis Center dapat meningkatkan pendapatannya. Hal ini akan berkontribusi positif terhadap dana sekolah, memungkinkan pengembangan fasilitas dan program pendidikan yang lebih baik.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Hasil literature review yang telah dilakukan pada jurnal-jurnal penelitian terkait topik Algoritma K-Means Clustering dapat dijabarkan sebagai berikut:

Paper (Dharma Putra et al., 2021) membahas mengenai penggunaan algoritma *K-Means Clustering* dalam analisis Memanfaatkan Algoritma *K-Means* untuk Mengelompokkan Riwayat Data Penjualan Tujuan dari penelitian ini adalah untuk mengetahui tingkat penjualan dari waktu ke waktu. Organisasi X adalah sebuah organisasi yang bergerak di bidang inovasi data dan media komunikasi. Berapa banyak informasi mengenai produk dari stok yang ada di seluruh dunia sehingga perlu dikumpulkan. Informasi *history* transaksi merupakan informasi yang dapat dimanfaatkan untuk melihat perkembangan perkembangan barang dagangan dari yang paling baik hingga yang kurang memikat. Hal ini diperlukan sebagai acuan bagi organisasi dalam meninjau ulang barang dagangan untuk membantu transaksi dan menjaga *loyalitas* konsumen. Eksplorasi ini menggunakan strategi pengelompokan yang merupakan teknik yang dapat digunakan untuk melihat tingkat transaksi yang telah dilakukan berdasarkan tandan yang terbentuk. Algoritma *K-Means* merupakan metode yang cocok untuk mengolah data *historis* penjualan karena mudah diimplementasikan, berjalan dengan cepat, dan mudah dimodifikasi dibandingkan dengan metode pengelompokan lainnya.

Paper (Anikah et al., 2020) membahas mengenai penggunaan algoritma *K-Means Clustering* dalam analisis Memanfaatkan teknik *K-Means* untuk mengelompokkan data barang untuk menentukan persediaan stok barang. Titik fokus dari pemeriksaan

ini adalah untuk menentukan persediaan barang dagangan tersebut. Masalah yang dilihat oleh PT SAS Gathering berhubungan dengan aksesibilitas produk yang merupakan komponen vital, sehingga diperlukan administrasi yang baik selama waktu yang dihabiskan untuk berurusan dengan aksesibilitas stok barang, untuk menjauhkan diri dari pengumpulan barang dagangan dan produk yang kurang menarik bagi pembeli dengan memanfaatkan penggalian informasi. Metode *K-Means* digunakan untuk pengelompokan. Tahapan eksplorasi yang dilakukan adalah identifikasi masalah, studi penulisan, pengumpulan informasi, penanganan informasi, penggalian informasi dan penilaian. Motivasi dibalik penelitian ini adalah untuk membantu PT SAS Group dalam menentukan beban stok produk. Konsekuensi dari penelitian ini adalah masing-masing kelompok, lebih spesifiknya cluster_0 sebanyak 1282 atau 98.6%, cluster_1 sebanyak 18 atau 1.4%. Hal ini menunjukkan bahwa stok barang di cluster_0 memiliki rata-rata volume penjualan yang signifikan yaitu sebesar 2368 barang. Hal ini menunjukkan bahwa pengumpulan barang dagangan di cluster_0 perlu menyimpan banyak produk karena merupakan barang yang paling banyak diminta.[4]

Paper (Rahayu et al., 2020) membahas mengenai penggunaan algoritma *K-Means Clustering* dalam analisis Pemanfaatan Information Mining Strategi *K-Means Bunching* untuk Pemeriksaan Transaksi Rotan. Informasi yang sangat banyak akan sulit untuk diselidiki. Pemeriksaan diharapkan dapat melihat desain penawaran barang sehingga dapat memberikan penawaran barang yang diharapkan, memisahkan data dari informasi yang cakupannya sangat besar harus dapat dilakukan dengan prosedur information mining, strategi yang digunakan adalah pengelompokan (*bunching*) yang dapat digunakan sebagai alasan untuk dinamis dalam mengantisipasi, Mengingat hasil percobaan dipartisi menjadi 3 (Tiga) Kelompok Untuk lebih spesifik Tidak dalam kerangka berpikir, populer dan populer. Hasil tes kelompok adalah konsekuensi dari kelompok 0 mendapatkan nilai 4636996106454.166. Hasil kelompok 1 mendapatkan nilai 1088684561455.892. Hasil kelompok 2 mendapatkan nilai 4624429728751.111. Terdapat 6 data di kelompok Tidak Berlaku, 114 di kelompok Cukup Berlaku, dan 9 di kelompok Diminati.[5]

Paper (Riadi, 2023) membahas mengenai penggunaan algoritma *K-Means Clustering* dalam analisis Pemanfaatan Information Mining yang Melibatkan Perhitungan *K-Means* untuk Pemeriksaan Transaksi Parfum. Sebuah toko bernama IM Parfum Rantaupratap menjual berbagai macam parfum bermerek IM. Meskipun memberikan keputusan yang berbeda, tidak semua parfum terjual dengan cepat, ada yang populer dan ada juga yang kurang memikat. Informasi mengenai penawaran, pembelian, dan pengeluaran di toko tidak terkoordinasi, sehingga informasi tersebut hanya mampu menjadi catatan sejarah tanpa memiliki opsi untuk digunakan demi kemajuan teknik periklanan. Informasi yang telah

dikumpulkan harus digunakan sebagai kerangka kerja yang dinamis untuk mengatasi masalah bisnis. Untuk mencapai hal ini, pencipta merencanakan sebuah aplikasi penambahan informasi dalam ulasan ini dengan harapan dapat memberikan hasil yang menarik dan paling ekstrim dalam memeriksa penawaran wewangian di toko IM Parfum Rantaupratap. Pemanfaatan *Information Mining* dengan Perhitungan *K-Means* ditunjukkan untuk memberikan pemeriksaan yang terbaik dan menjadi jawaban dalam mengembangkan bisnis parfum. *Rapidminer* mampu membuat tiga cluster dengan pemodelan clustering dengan algoritma *K-Means* dan membagi jumlah cluster menjadi tiga. Dari 25 item produk yang diamati, cluster 1 berisi 9 produk, cluster 2 berisi 3 produk, dan cluster 3 berisi 13 produk.[6]

Paper (Siregar et al., 2023) membahas mengenai penggunaan algoritma *K-Means Clustering* dalam analisis Pemanfaatan Data Mining pada Transaksi Rumah Makan Kasih Ibu Menggunakan Strategi *K-Means Bunching*. Rumah Makan Kasih Ibu dibuka pada tahun 2013 di Jalan Desa Dusun III, Medan Tebing Tinggi, Sei Sijenggi Permasalahan yang sering terjadi di Rumah Makan Kasih Ibu adalah tidak adanya informasi tentang bagaimana menentukan menu yang umumnya digemari oleh para pelanggan sehingga menyebabkan menipisnya menu secara terus menerus, yang menimbulkan rasa frustrasi dari para pembeli yang ada. Oleh karena itu, Restoran Kasih Ibu perlu mencari cara untuk mengatasi masalah persediaan menu. Salah satu cara untuk melakukannya adalah dengan melihat bagaimana pelanggan membeli sesuatu untuk mengetahui menu mana yang memiliki keterkaitan dan perlu disediakan dalam jumlah yang lebih banyak. Dengan menggunakan Metode *K-Means Clustering*, ilmu data mining dapat diterapkan pada masalah tersebut. Dengan adanya tinjauan ini, Information Mining *K-Means Clustering* dapat memberikan data kepada pemilik Rumah Makan Kasih Ibu sehubungan dengan pengumpulan penawaran menu yang ada, sebagai 10 macam menu yang Paling Disukai, 19 menu yang Cukup Disukai, dan 4 menu yang Kurang Disukai, sehingga dapat digunakan sebagai referensi ketika mengambil keputusan tentang apa yang akan dimasukkan ke dalam menu Rumah Makan Kasih Ibu.[7]

Paper (FITHRI, 2023) membahas mengenai Informasi pertukaran transaksi yang otentik dapat digunakan sebagai data untuk pilihan bisnis. Toko Usaha Mandiri memiliki masalah dalam memuat barang dagangan, ada barang dagangan yang laten yang membuat nyaman pusat distribusi untuk jangka waktu yang cukup lama. Eksplorasi masa lalu membedah informasi untuk membingkai kumpulan informasi menjadi data yang bermanfaat. Eksplorasi ini bermaksud untuk memecah informasi transaksi dengan menerapkan perhitungan *K-Means Grouping* untuk membentuk tandan transaksi. Konsekuensi dari pengelompokan informasi membentuk struktur cluster1, cluster2 dan cluster3 dengan nilai tingkat

62% (11 informasi), 8% (56 informasi) dan 30% (25 informasi) secara terpisah. Davies Bouldin Index digunakan untuk memvalidasi algoritma pengelompokan K-Means, dan hasilnya adalah nilai 0.2. Pengelompokan data penawaran dapat menjadi pengaturan pilihan, bahan masukan untuk administrasi saham dan teknik periklanan.[8]

Paper (Pambudi & Witanti, 2021) membahas mengenai Pengelompokan barang selama ini masih dilakukan secara manual, yaitu berdasarkan pengamatan pemilik toko. Oleh karena itu, penelitian ini bertujuan untuk mendukung proses penentuan stok barang dengan membangun model yang dapat mengelompokkan barang ke dalam kategori penjualan tinggi dan rendah dengan menggunakan k-means clustering agar dapat menjaga persediaan stok barang dan mengeliminasi barang yang kosong. Centroid yang paling tinggi akan menjadi centroid yang paling banyak dicari, sedangkan centroid yang rendah adalah centroid yang kurang diminati. Informasi yang digunakan dalam penelitian ini diambil dari informasi transaksi pada tahun 2017 dan 2018. Situasi pengelompokan menggunakan faktor nama barang, informasi barang yang mendekati, informasi barang yang aktif, dan stok barang. Konsekuensi dari penelitian ini menunjukkan nilai ekshibisi dari kerangka kerja dalam mengelompokkan produk sebesar 83,33%.[9]

Paper (Aprilianur & Hadisaputro, 2022) membahas mengenai Informasi mengenai penawaran, perolehan produk dan biaya yang mengejutkan di Myam Hijab tidak terlalu terorganisir, dengan tujuan bahwa informasi tersebut hanya berkemampuan sebagai dokumen toko dan tidak dapat digunakan untuk mengembangkan teknik showcasing. Oleh karena itu, penting untuk menerapkan penggalan informasi dengan menggunakan strategi K-Means pada toko Myam Hijab. Strategi K-Means dapat diterapkan pada toko Myam Hijab untuk mengetahui penawaran hijab mana yang banyak dicari, populer dan tidak banyak dicari. Data stok hijab yang ada di toko Myam ini dikelompokkan dengan menggunakan metode K-Means. Kemudian pilih 3 kumpulan sembarang sebagai centroid. Setelah informasi dari masing-masing gathering tidak mengalami transformasi, maka akan terlihat bahwa hasil akhirnya adalah 24 item yang laris manis, 59 item yang laris, dan 17 item yang tidak laris. Kemudian, pada titik tersebut, penggunaan teknik K-Means di Rapidminer selesai dengan memasukkan informasi stok barang, khususnya stok awal, stok terjual, dan stok terakhir, yang akan diubah menjadi kumpulan data di Ms. Produk yang banyak diminati, permintaan tinggi, dan permintaan rendah kemudian akan diproduksi oleh Rapidminer.[10]

Paper (Azmi et al., 2023) membahas mengenai pada Apotek M23 saat ini, menyelesaikan seluruh tahap pengelompokan data, Di toko obat M23, dalam beberapa kasus terjadi kekurangan atau kelebihan obat. Untuk mengatasi masalah ini, strategi penggalan

informasi diterapkan dengan memecah penggunaan obat untuk menghasilkan data yang dapat digunakan untuk pengendalian persediaan dan perencanaan obat. Teknik yang digunakan dalam penelitian ini adalah strategi *K-Means*. Strategi pengelompokan *K-Means* bermaksud untuk mengelompokkan informasi yang memiliki atribut yang sama ke dalam kelompok yang sama dan informasi yang memiliki kualitas yang berbeda dikumpulkan ke dalam kelompok yang berbeda. menggunakan metode *Davies Bouldin Index* (DBI) untuk menentukan jumlah *cluster* yang ideal. Hasil dari penelitian ini mengkonfirmasi bahwa jumlah kelompok terbaik adalah 2 kelompok, pengumpulan informasi obat terdiri dari kelompok 1 dengan penggunaan obat rendah yang terdiri dari 144 jenis obat dan kelompok 2 dengan penggunaan obat tinggi yang terdiri dari 6 jenis obat.[11]

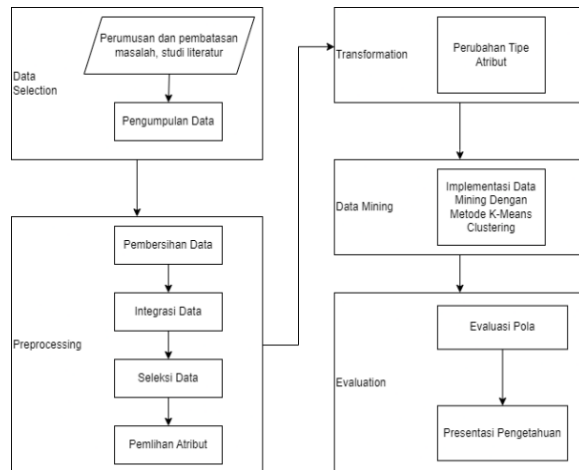
Paper (Rahman & Suroyo, 2021) membahas Informasi yang tergores adalah teks komentar, proyeksi pemasaran, dan skor peringkat bintang. Informasi dari hasil pemeriksaan yang diperoleh pada investigasi pesan komentar item ponsel dengan biaya minimal wordcloud item ponsel menunjukkan informasi komentar pusat komersial Shopee Indonesia dapat berupa bahwa baik di ponsel dengan biaya minimal dan biaya yang signifikan secara umum akan memiliki desain wordcloud yang setara di mana kata-kata yang berlaku yang muncul secara umum akan memihak dan positif, sedangkan kata-kata yang berarti negatif cenderung tidak dominan. Sedangkan kata-kata yang sering muncul adalah barang dagangan, konsisten, cepat, kirim dan hebat dengan nilai tingkat ketepatan sebesar 92%. Sedangkan pada konsekuensi dari proses wordcloud medium significant expense didapatkan kata yang sering muncul adalah kata (kirim, cepat, dan hebat) dengan nilai ketepatan sebesar 94%. Selanjutnya, jika dilihat dari efek grafis dari proses k-implies information bunching menunjukkan bahwa proyeksi pemasaran 0 hingga 1000 mendapatkan nilai star rating yang paling tinggi dan penawaran dengan nilai star rating paling rendah berada di antara 1500 hingga 2000 atau lebih..[12]

3. METODE PENELITIAN

3.1. Jenis Penelitian

Pada penelitian ini menggunakan metode penelitian K-Means Clustering dengan pendekatan kuantitatif yang dimana pada hasilnya akan menampilkan angka statistik dalam bentuk visualisasi grafik. Metode K-Means clustering merupakan suatu bentuk pengelompokkan sebuah nilai pada subjek yang diambil nilai berdekatan kemudian dijadikan sebagai satu kelompok, dan apabila nilai tersebut saling berjauhan maka akan dibuat kelompok baru. Penentuan kelompok tergantung oleh seseorang yang akan disaring menjadi beberapa kelompok, umumnya sebuah angka statistik dinilai akurat dengan membuat menjadi tiga kelompok. Pada metode penelitian ini menggunakan tahapan KDD (Knowledge Discovery Database). KDD merupakan sebuah Langkah praktis

dalam melakukan tahapan data mining untuk mencari, mengidentifikasi pola pada data, dimana pola yang ditemukan bersifat sah dan bermanfaat dan mudah dimengerti. Dalam tahapan KDD terdapat beberapa tahapan diantaranya pembersihan data, integrasi data, pemilihan data, seleksi data, transformasi data. Berikut adalah gambar mengenai tahapan – tahapan dalam KDD (Knowledge Discovery Database):



Gambar 1. Metode Penelitian

Tahapan Aktivitas yang di lakukan penulis diuraikan sebagai berikut :

- 1. Data Selection**
Perumusan dan pembatasan masalah, studi literatur. Merumuskan masalah berdasarkan pada latar belakang, kemudian menentukan Batasan masalah sehingga penelitian lebih terfokus pada satu masalah, setelah itu mencari studi literatur terkait topik yang diteliti.
- 2. Pengumpulan Data**
Melakukan pengumpulan data dari sumber yang telah ditentukan.
- 3. Preprocessing**
Pembersihan Data atau Melakukan pembersihan pada data yang tidak konsisten atau memiliki nilai dan atribut kosong.
- 4. Integrasi Data**
Menggabungkan data dari setiap database ke dalam database baru.
- 5. Seleksi Data**
Menyeleksi data digunakan untuk menganalisis atribut pada data yang sesuai dan relevean untuk dilakukan metode data mining sehingga data akan diseleksi terlebih dahulu.
- 6. Pemilihan Atribut**
Pemilihan atribut dalam data mining skripsi dilakukan untuk menemukan atribut yang muncul dalam satu waktu.
- 7. Transformation**
Perubahan Tipe Atribut Perubahan tipe atribut untuk penyesuaian dengan perkembangan pengetahuan dan kebutuhan penelitian.
- 8. Data Mining**

Implementasi Data Mining dengan metode K-Means Clustering Menerapkan data mining pada data yang akan diolah dan dianalisis menggunakan sistem pengelompokkan yaitu Clustering dan algoritma K-Means serta menggunakan tools RapidMiner.

9. Evaluation

Evaluasi Pola Mengevaluasi pola pada hasil yang telah ditemukan pada penerapan data mining K-mean Clustering tersebut untuk melihat keakuratan dan ketepatan hasil.

10. Presentasi

Pengetahuan Mempresentasikan pengetahuan dengan cara memberikan informasi yang didapat kemudian direpresentasikan menggunakan gambaran visual.

3.2. Sumber Data

Sumber data ini mencakup catatan penjualan historis Bisnis Center, yang mencakup informasi rinci tentang kode barang, nama barang, penjualan, kategori, supplier. Data ini akan mencakup sejarah penjualan produk di Bisnis Center selama periode 1 Juli 2023 sampai 31 Oktober 2023 dengan 552 baris data, yang akan digunakan sebagai dasar untuk analisis pengelompokan menggunakan metode K-Means.

3.3. Populasi dan Sampel

Data populasi mencakup semua data penjualan yang berjumlah 552. Mencakup kode barang, nama barang, penjualan, kategori, supplier. Sample yang akan diambil yaitu semuanya atau berjumlah 552, teknik pengambilan sample menggunakan Teknik sampling jenuh.

3.4. Teknik Analis Data

Setelah data didapatkan tahapan selanjutnya proses pengolahan dan Teknik Analisa data. Pada tahap ini penulis mengelola data menggunakan Teknik analisis *K-means Clustering*. *K-means Clustering* adalah Teknik untuk mengelompokkan data berdasarkan kemiripan dari suatu data. Pada tahap ini penulis melakukan 9 kali percobaan dalam pencarian jumlah K yang optimal berdasarkan nilai DBI. Adapun rumus dalam pencarian DBI sebagai berikut:

$$DBI = \frac{1}{K} \sum_{i=1}^K MAX_{i \neq j} (R_{i,j}) \quad (1)$$

Rumus ini menjelaskan:

- K adalah jumlah *cluster* yang dihasilkan oleh *K-means*
- $R_{i,j}$ adalah rata rata jarak antara semua pasangan titik data dalam cluster i ke pusat cluster j

4. HASIL DAN PEMBAHASAN

4.1. Data Selection

Tahapan pertama dalam proses KDD yaitu data selection. Pada tahapan ini penulis melakukan seleksi data yang akan digunakan dalam proses pengelompokkan. Data yang akan diproses yaitu data penjualan pada Bisnis Center SMK Wahidin dari

tanggal 1 Juli 2023 sampai 31 Oktober 2023. Tahapan seleksi data bertujuan untuk memilih data yang akan digunakan dalam proses pengelompokan melalui aplikasi Rapidminer, tahapan seleksi data ini dilakukan di Microsoft Excel. Data yang digunakan dalam penelitian ini yaitu data penjualan pada Bisnis Center SMK Wahidin dari tanggal 1 Juli 2023 sampai

31 Oktober 2023 dengan dataset total berjumlah 552 dan 6 atribut diantaranya NO, KODE BARANG, NAMA BARANG, PENJUALAN, KATEGORI, SUPPLIER. Dataset yang digunakan dapat dilihat pada tabel 1.

Tabel 1. Data Penjualan

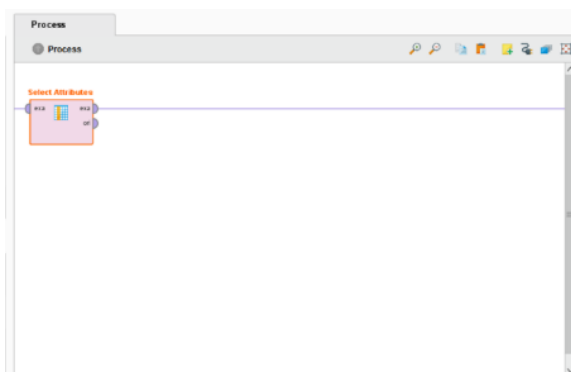
KODE BARANG	NAMA BARANG	PENJUALAN	KATEGORI
8991002122000	ABC CHOCO MALT COFFE 200ML	7	DRINK
8992761139018	ADES 600ML	185	DRINK
8885013132083	AICE CHOCO MALT 49ML	51	DRINK
8995899216132	BONCABE MAKARONI LV 2 27G	67	FOOD
8992775311981	CHOCOLATOS CHOCOLATE	43	FOOD
8992761002039	FANTA STRAWBERRY 390ML	16	DRINK
8994834003950	FRENCH FRIES 2000	56	FOOD
8992775311622	GERY SALUUT COCONUT ROLL 6G	24	FOOD
8998866202893	GOLDA CAPPUCINO 200ML	53	DRINK
-	-	-	-
8993200668915	YOGURT STRAWBERRY 40G	115	DRINK

4.2. Preprocessing

Setelah melakukan tahapan data selection maka tahapan selanjutnya adalah melakukan preprocessing. Tujuan dari preprocessing yaitu untuk penghapusan atribut data yang tidak diperlukan, data yang null atau missing, menghilangkan data yang tidak konsisten, serta menambahkan id kedalam dataset yang akan digunakan. Pada tahapan preprocessing ini penulis menggunakan 2 jenis operator yaitu operator Select Attributes dan operator Set Role. Adapun langkah-langkah yang dilakukan penulis dalam menggunakan 2 operator tersebut yaitu sebagai berikut :

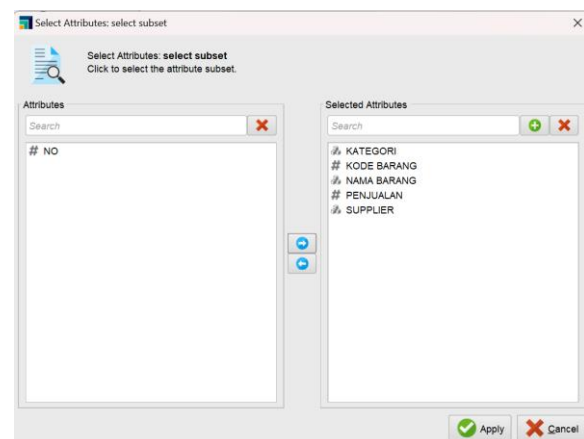
a. Operator Select Attributes

Langkah pertama yang dilakukan penulis dalam tahapan preprocessing yaitu menggunakan operator Select Attributes. Operator Select Attributes adalah operator yang berfungsi untuk memilih subset atribut yang akan digunakan dari ExampleSet dan menghapus atribut lainnya yang tidak digunakan. Tampilan operator Select Attributes dapat dilihat pada gambar 2.



Gambar 2. Tampilan Operator Select Attribute

Pada operator Select Attributes terdapat parameter yang harus disesuaikan terlebih dahulu. Tampilan proses menggunakan parameter select attributes untuk menghilangkan atribut yang tidak digunakan dapat dilihat pada gambar 3.



Gambar 3. Parameter Select Attributes

Atribut data yang tidak digunakan dalam penelitian ini adalah atribut data NO, maka setelah dilakukan proses *Select Attributes* maka dataset yang semula berjumlah 6 atribut sekarang berjumlah 5 atribut yang akan digunakan yaitu KODE BARANG, NAMA BARANG, PENJUALAN, KATEGORI, SUPPLIER. Hasil yang diperoleh dari proses operator ini dapat dilihat pada gambar 4.

Row No.	KODE BAR...	NAMA BAR...	PENJUALAN	KATEGORI	SUPPLIER
1	107	1 PAKET AT...	0	FASHION	HMS
2	8991002122...	ABC CHOCO...	7	DRINK	TOKO MEKA...
3	8991002122...	ABC KOPI S...	3	DRINK	TOKO MEKA...
4	8992761139...	ADES 500ML	185	DRINK	TOKO MEKA...
5	8885013131...	AICE BINGO...	10	FOOD	PT. MY
6	8885013132...	AICE CHOC...	51	DRINK	TOKO MEKA...
7	8885013130...	AICE CHOC...	33	FOOD	PT. MY
8	8885013130...	AICE CHOC...	24	FOOD	PT. MY
9	8885013131...	AICE CHOC...	15	FOOD	PT. MY
10	8885013130...	AICE COFFE...	32	FOOD	PT. MY
11	8885013130...	AICE FREEZ...	63	FOOD	PT. MY
12	8885013131...	AICE FRUZZ...	31	FOOD	PT. MY
13	8885013132...	AICE FRUZZ...	27	FOOD	PT. TANJUN...
14	8885013130...	AICE JERUK...	25	FOOD	PT. TANJUN...
15	8885013130...	AICE MANGO...	1	FOOD	PT. MY
16	8885013131...	AICE MARKI...	81	FOOD	PT. MY

ExampleSet (552 examples, 0 special attributes, 5 regular attributes)

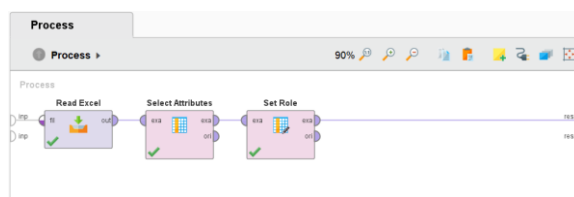
Row No.	KODE BAR...	NAMA BAR...	PENJUALAN	KATEGORI	SUPPLIER
537	8996001350...	WAFELLO C...	1	FOOD	TOKO MEKA...
538	8996001350...	WAFELLO C...	38	FOOD	TOKO MEKA...
539	8996001358...	WAFELLO C...	158	FOOD	TOKO MEKA...
540	8994171101...	WHITE KOFF...	73	FOOD	ALFAMART
541	8998866802...	WPC PEMBE...	28	NON FOOD	ALFAMART
542	8993200666...	YOGURT BL...	3	DRINK	CIMORY
543	8993200668...	YOGURT BL...	115	DRINK	CIMORY
544	8993200668...	YOGURT BR...	62	DRINK	CIMORY
545	8993200668...	YOGURT CA...	1	DRINK	CIMORY
546	8993200661...	YOGURT MA...	3	DRINK	CIMORY
547	8993200667...	YOGURT PE...	19	DRINK	CIMORY
548	8993200668...	YOGURT PU...	1	DRINK	CIMORY
549	8993200666...	YOGURT ST...	5	DRINK	CIMORY
550	8993200668...	YOGURT ST...	115	DRINK	CIMORY
551	48	YOU C1000...	201	DRINK	ALFAMART
552	8997213770...	YOU-YOU S...	91	FOOD	TOKO MEKA...

ExampleSet (552 examples, 0 special attributes, 5 regular attributes)

Gambar 4. ExampleSet Data Setelah Menggunakan Select Attribute

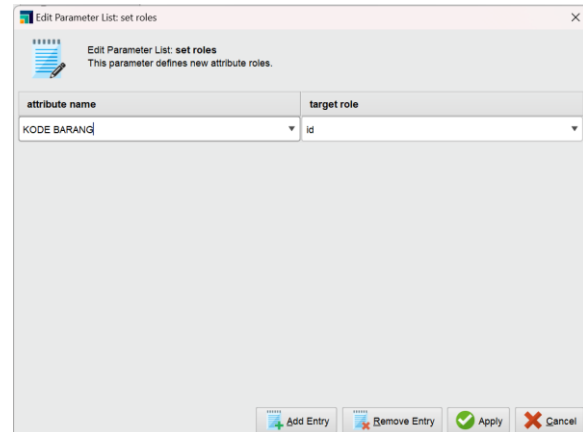
b. Operator Set Role

Langkah yang kedua dalam tahapan preprocessing adalah menggunakan operator Set Role. Operator Set Role digunakan untuk mengubah peran dari satu atau lebih atribut. Tampilan operator Set Role dapat dilihat pada gambar 5.



Gambar 5. Tampilan Operator Set Role

Pada operator Set Role terdapat parameter yang harus disesuaikan terlebih dahulu. Tampilan proses menggunakan parameter Set Role untuk mengubah atribut menjadi id dapat dilihat pada gambar 6.



Gambar 6. Parameter Set Role

Atribut data yang menjadi target role id yaitu atribut KODE BARANG. Hasil yang diperoleh dari proses operator ini dapat dilihat pada gambar 7.

Row No.	KODE BAR...	NAMA BAR...	PENJUALAN	KATEGORI	SUPPLIER
1	107	1 PAKET AT...	0	FASHION	HMS
2	8991002122...	ABC CHOCO...	7	DRINK	TOKO MEKA...
3	8991002122...	ABC KOPI S...	3	DRINK	TOKO MEKA...
4	8992761139...	ADES 500ML	185	DRINK	TOKO MEKA...
5	8885013131...	AICE BINGO...	10	FOOD	PT. MY
6	8885013132...	AICE CHOC...	51	DRINK	TOKO MEKA...
7	8885013130...	AICE CHOC...	33	FOOD	PT. MY
8	8885013130...	AICE CHOC...	24	FOOD	PT. MY
9	8885013131...	AICE CHOC...	15	FOOD	PT. MY
10	8885013130...	AICE COFFE...	32	FOOD	PT. MY
11	8885013130...	AICE FREEZ...	63	FOOD	PT. MY
12	8885013131...	AICE FRUZZ...	31	FOOD	PT. MY
13	8885013132...	AICE FRUZZ...	27	FOOD	PT. TANJUN...
14	8885013130...	AICE JERUK...	25	FOOD	PT. TANJUN...
15	8885013130...	AICE MANGO...	1	FOOD	PT. MY
16	8885013131...	AICE MARKI...	81	FOOD	PT. MY

ExampleSet (552 examples, 1 special attribute, 4 regular attributes)

Row No.	KODE BAR...	NAMA BAR...	PENJUALAN	KATEGORI	SUPPLIER
537	8996001350...	WAFELLO C...	1	FOOD	TOKO MEKA...
538	8996001350...	WAFELLO C...	38	FOOD	TOKO MEKA...
539	8996001358...	WAFELLO C...	158	FOOD	TOKO MEKA...
540	8994171101...	WHITE KOFF...	73	FOOD	ALFAMART
541	8998866802...	WPC PEMBE...	28	NON FOOD	ALFAMART
542	8993200666...	YOGURT BL...	3	DRINK	CIMORY
543	8993200668...	YOGURT BL...	115	DRINK	CIMORY
544	8993200668...	YOGURT BR...	62	DRINK	CIMORY
545	8993200668...	YOGURT CA...	1	DRINK	CIMORY
546	8993200661...	YOGURT MA...	3	DRINK	CIMORY
547	8993200667...	YOGURT PE...	19	DRINK	CIMORY
548	8993200668...	YOGURT PU...	1	DRINK	CIMORY
549	8993200666...	YOGURT ST...	5	DRINK	CIMORY
550	8993200668...	YOGURT ST...	115	DRINK	CIMORY
551	48	YOU C1000...	201	DRINK	ALFAMART
552	8997213770...	YOU-YOU S...	91	FOOD	TOKO MEKA...

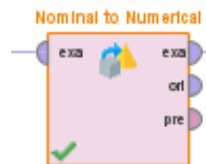
ExampleSet (552 examples, 1 special attribute, 4 regular attributes)

Gambar 7. ExampleSet Statistic Setelah Menggunakan Set Role

4.3. Transformation

Setelah melakukan tahapan preprocessing maka tahapan selanjutnya adalah melakukan transformation. Pada tahapan transformation penulis mengubah tipe atribut data yang semula non-numeric diubah menjadi

tipe data numeric. Atribut data yang diubah yaitu atribut NAMA BARANG, KATEGORI dan SUPPLIER menjadi angka menggunakan operator Nominal to Numerical agar sesuai dengan tipe data yang dibutuhkan algoritma K-Means pada proses clustering. Tampilan operator Nominal to Numerical dapat dilihat pada gambar 8.



Gambar 8. Operator nominal to numerical

Dalam menggunakan operator Nominal to Numerical terdapat parameter yang harus disesuaikan terlebih dahulu agar sesuai dengan kebutuhan pada proses *clustering* penelitian. Tampilan proses menggunakan parameter Nominal to Numerical dapat dilihat pada gambar 9.



Gambar 9. Parameter Nominal To Numerical

Hasil yang diperoleh dari proses transformasi atribut data NAMA BARANG, KATEGORI dan SUPPLIER menggunakan operator *Nominal to Numerical* dapat dilihat pada gambar 10.

Row No.	KODE BAR...	NAMA BAR...	KATEGORI	SUPPLIER	PENJUALAN
1	107	0	0	0	0
2	0581002122...	1	1	1	7
3	0581002122...	2	1	1	3
4	0582761139...	3	1	1	105
5	8885013131...	4	2	2	10
6	8885013132...	5	1	1	51
7	8885013130...	6	2	2	33
8	8885013130...	7	2	2	24
9	8885013131...	8	2	2	16
10	8885013130...	9	2	2	32
11	0585013130...	10	2	2	63
12	0585013131...	11	2	2	31
13	0585013132...	12	2	3	27
14	8885013130...	13	2	3	29
15	8885013130...	14	2	2	1
16	8885013131...	15	2	2	81

ExampleSet (552 examples, 1 special attribute, 4 regular attributes)

Row No.	KODE BAR...	NAMA BAR...	KATEGORI	SUPPLIER	PENJUALAN
537	8996001350...	536	2	1	1
538	8996001350...	537	2	1	38
539	8996001358...	538	2	1	108
540	8994171101...	539	2	8	73
541	899866802...	540	3	8	20
542	8993200656...	541	1	12	3
543	8993200656...	542	1	12	115
544	8993200658...	543	1	12	62
545	8993200658...	544	1	12	1
546	8993200651...	545	1	12	3
547	8993200657...	546	1	12	19
548	8993200658...	547	1	12	1
549	8993200656...	548	1	12	5
550	8993200658...	549	1	12	115
551	48	550	1	8	201
552	8997213770...	551	2	1	91

ExampleSet (552 examples, 1 special attribute, 4 regular attributes)

Gambar 10. ExampleSet Data setelah di Transformasi

4.4. Data Mining

Setelah melakukan tahapan transformation maka tahapan selanjutnya adalah melakukan penerapan pada data mining. Pada tahapan penerapan data mining penulis menggunakan tools rapidminer versi 10.2.000 dengan metode algoritma K-Means. Pada tahapan data mining ini penulis membagi menjadi dua tahapan yaitu tahapan pertama penulis melakukan proses clustering menggunakan algoritma K-Means dengan Operator yang digunakan yaitu operator Clustering (K-Means) dan tahapan kedua penulis melakukan pengujian dan mengevaluasi hasil dari clustering menggunakan operator Cluster Disance Performance dengan metode yang digunakan evaluasi David Bouldin Index (DBI).

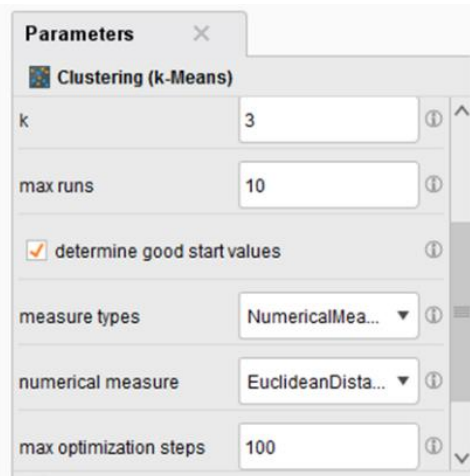
4.4.1. Tahapan Clustering K-Means

Tahapan pertama yaitu penulis melakukan proses clustering K-Means dilakukan 9 kali percobaan dimulai dari $k = 2$ sampai $k = 10$. Pemodelan data mining pada pengelompokkan data pengiriman paket menggunakan algoritma K-Means dapat dilihat pada gambar 11.



Gambar 11. Model Penerapan Algoritma K-Means

Pada proses clustering terdapat parameter yang harus disesuaikan terlebih dahulu. Tampilan parameter clustering dapat dilihat pada gambar 12.



Gambar 12. Tampilan Parameter K-Means Clustering

Untuk penjelasan lebih jelas terhadap parameter yang digunakan yaitu sebagai berikut :

1. K adalah jumlah cluster yang digunakan. Jumlah percobaan cluster yang digunakan oleh penulis yaitu dari (2-10 cluster).
2. Max runs adalah Jumlah maksimum run K-Means (default). Jumlah yang digunakan penulis yaitu 10.
3. Measure types adalah Jenis pengukuran numeric. Jenis pengukuran numerik yang digunakan penulis yaitu Numerical Measures.
4. Numerical Measure adalah perhitungan jarak dari 2 buah titik. Perhitungan jarak yang digunakan penulis yaitu Euclidean Distance.
5. Max optimization steps adalah Jumlah maksimal iterasi (default). Jumlah maksimal iterasi yang digunakan penulis yaitu 100.

4.4.2. Evaluasi David Bouldin Index (DBI)

Tahapan yang kedua yaitu melakukan pengujian hasil clustering dengan menerapkan nilai DBI pada parameter Cluster Distance Performance. Pada pengujian DBI, cluster yang memiliki nilai DBI terkecil atau mendekati 0 dijadikan sebagai cluster yang terbaik. Untuk mencari cluster terbaik penulis melakukan percobaan dari cluster 2 sampai dengan 10 rekapitulasi. Jumlah k cluster yang dihasilkan dari nilai Davies Bouldin Index dapat dilihat dari tabel 2 berikut:

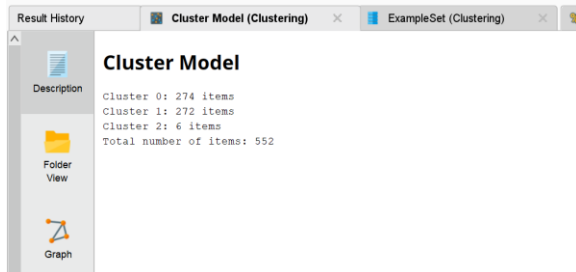
Tabel 2. Hasil Cluster K2-K10

Clustering	Jumlah Anggota Cluster	Hasil Nilai DBI
2	Cluster 0: 271 items	0.182
	Cluster 1: 281 items	
3	Cluster 0: 274 items	0.149
	Cluster 2: 6 items	
4	Cluster 0: 180 items	0.172
	Cluster 1: 179 items	
	Cluster 3: 187 items	
5	Cluster 0: 175 items	0.151
	Cluster 1: 182 items	
	Cluster 2: 172 items	
	Cluster 4: 20 items	
6	Cluster 0: 131 items	0.165
	Cluster 1: 129 items	
	Cluster 3: 20 items	
	Cluster 5: 138 items	
7	Cluster 0: 126 items	0.172
	Cluster 1: 129 items	
	Cluster 2: 127 items	
	Cluster 3: 128 items	
	Cluster 5: 3 items	
8	Cluster 0: 124 items	0.172
	Cluster 1: 125 items	
	Cluster 2: 3 items	
	Cluster 3: 129 items	
	Cluster 4: 125 items	
	Cluster 5: 15 items	
9	Cluster 0: 13 items	0.168
	Cluster 1: 111 items	
	Cluster 2: 107 items	
	Cluster 3: 97 items	
	Cluster 4: 23 items	
	Cluster 5: 3 items	
	Cluster 6: 91 items	
10	Cluster 0: 94 items	0.175
	Cluster 1: 86 items	
	Cluster 2: 78 items	
	Cluster 3: 88 items	
	Cluster 4: 3 items	
	Cluster 5: 83 items	
	Cluster 6: 20 items	
	Cluster 7: 81 items	
	Cluster 8: 15 items	

4.5. Evaluation

Pada tahapan evaluation merupakan tahapan dalam menghasilkan pola-pola guna memperkirakan tujuan yang diharapkan. Berdasarkan tabel 4.2 Menunjukkan bahwa cluster data penjualan pada Bisnis

Center SMK Wahidin menggunakan perhitungan Davies Bouldin Index nilai yang paling mendekati angka 0 dengan percobaan cluster 2 sampai cluster 10 menghasilkan nilai k terbaik pada cluster 3 yaitu 0.597 dengan jumlah anggota Cluster 0: 274 items, Cluster 1: 272 items dan Cluster 2: 6 items.



Gambar 13. Description Cluster Model K-Means

Berikut merupakan jarak antar cendroid cluster dari data penjualan Bisnis Center SMK Wahidin menggunakan algoritma K-Means bisa dilihat pada gambar 13 dan hasil dari deskripsi PerformanceVector dapat dilihat pada gambar 14.

Attribute	cluster_0	cluster_1	cluster_2
NAMA BARANG	414.131	138.840	149
KATEGORI	2.018	2.008	1.333
SUPPLIER	3.887	3.202	2.187
PENJUALAN	37.533	37.423	1021.500

Gambar 14. Tabel Jarak antar centroid

Performance
PerformanceVector:
Avg. within centroid distance: 3174.183
Avg. within centroid distance_cluster_0: 3054.615
Avg. within centroid distance_cluster_1: 2374.777
Avg. within centroid distance_cluster_2: 40000.903
Davies Bouldin: 0.149

Gambar 15. Hasil Deskripsi PerformanceVektor

Berdasarkan hasil pengujian clustering, jumlah cluster yang di dapat dengan nilai DBI terbaik K=3, dengan nilai DBI mendekati nol yaitu 0.149. Hasil pengelompokan dan pengujian menggunakan software RapidMiner menghasilkan 3 cluster, yaitu Cluster 0 (barang yang sedang laku) dengan 274 item, Cluster 1 (kurang laku) dengan 272 item, Cluster 2 (paling laku) dengan 6 item, Ketiga cluster tersebut terbentuk dari 552 baris data, hasil pembentukan cluster model tersebut pada gambar 16.

Cluster Model

Cluster 0: 274 items
Cluster 1: 272 items
Cluster 2: 6 items
Total number of items: 552

Gambar 16. Hasil Cluster

Berikut hasil cluster berdasarkan nama barang penjualan produk terdapat pada Tabel 3.

Tabel 3. Anggota Cluster 0

NOM OR	NAMA BARANG	PENJUALAN
1	PERMEN KISS CHERRY	500
2	TEH PUCUK HARUM 350ML	497
3	PERMEN KISS APPLEPEACH	426
4	PERMEN KISS BARLEY	351
5	PADIMAS PANDAN 9G	348
6	TARONET 3D POTATO JUNGLE CHICKEN 17G	317
7	PERMEN KISS GRAPE	311
8	TARO PUFF ROASTED CORN 20G	300
9	TARO NET POTATO BARBECUE 10G	294
10	TEH PUCUK HARUM 650ML	262
...	...	...
270	TEMPAT MINUM KECIL	0
271	TEMPAT SOFTLENS	0
272	TEMPAT SOFTLENS BESAR	0
273	THE GREAT STAPLES MODEL 10	0
274	TOPI	0

Tabel 4. Anggota Cluster 1

NOM OR	NAMA BARANG	PENJUALAN
1	ES YOGURT SEDANG	418
2	GOODDAY MOCACINNO 20G	280
3	CHIKI TWIST JAGUNG BAKAR 22,5G	273
4	CHOCOLATOS DARK CHOCOLATE 14G	242
5	GARUDA ROSTA BAWANG 26G	241
6	CRISPY CRACKERS 140G	217
7	GERY SALUUT COKLAT TABUR KELAPA 17,5G	209
8	GOOD MOOD BLACKCURRANT 450ML	203
9	GOOD DAY CARAMELLO 20G	199
10	COLLIN S 80g	198
...	...	...
268	LESTARI STYLER	0
269	LIPSTICK MAC	0
270	LIPSTICK REIZEN	0
271	LIPSTICK REIZEN ORGANIC	0
272	LIVING COLOR LOVELY	0

Tabel 5. Anggota Cluster 2

NOM OR	NAMA BARANG	PENJUALAN
1	AIR MINERAL VIOLA 220ML	1423
2	AIR MINERAL VIONA 200ML	1609
3	ANEKA MAKARONI 2000	569
4	ES YOGURT KECIL	679
5	LE MINERAL 600ML	671
6	PRIMA 600ML	1178

Dari hasil tabel data di atas diperoleh penjualan terbanyak pada tanggal 1 Juli 2023 sampai 31 Oktober 2023 yaitu berdasarkan nama barang AIR MINERAL VIONA 200ML dengan total penjualan sebanyak 1609 items, penjualan terbanyak kedua ada pada AIR MINERAL VIOLA 200ML dengan total penjualan 1423 items, penjualan terbanyak ketiga ada pada PRIMA 600ML dengan total penjualan sebanyak 1178 items. Pada hasil evaluasi setelah dilakukannya proses clustering dengan menggunakan cluster 2 sampai dengan 10 rekapitulasi dapat dilihat nilai DBI pada table berikut :

Tabel 6. Hasil Perhitungan Cluster Menggunakan DBI

Hasil DBI	
Cluster	Nilai
2	0.182
3	0.149
4	0.172
5	0.151
6	0.165
7	0.172
8	0.172
9	0.168
10	0.175

Dari hasil nilai DBI di atas dapat disimpulkan bahwa nilai k yang mendekati optimum adalah k=3 dengan hasil nilai DBI 0.149.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat disimpulkan bahwa. Hasil pengelompokan data penjualan pada Bisnis Center SMK Wahidin menggunakan algoritma K-Means menghasilkan 3 cluster, yaitu Cluster 0 (barang yang sedang laku) dengan 274 item, Cluster 1 (kurang laku) dengan 272 item, Cluster 2 (paling laku) dengan 6 item. Nilai k yang mendekati optimum adalah k=3 dengan hasil nilai DBI 0.149. Pada penelitian selanjutnya peneliti menyarankan untuk perbanyak percobaan terhadap Algoritma selain K-Means seperti Asosiasi, Apriori, dll, yang digunakan untuk selain mengelompokkan penjualan agar hasil akurasinya lebih tepat lagi.

DAFTAR PUSTAKA

- [1] A. Rahman and H. Suroyo, "Analisis Data Produk Elektronik di E-Commerce dengan Metode Algoritma K-Means Menggunakan Python," *Journal of Advances in Information* ..., 2021.
- [2] N. Azmi, H. S. Hafsah, Y. Yuyun, and H. Hazriani, "Penerapan Metode K-Means Clustering Dalam Mengelompokkan Data Penjualan Obat pada Apotek M23," *Prosiding SISFOTEK*, 2023.
- [3] P. Rahayu, I. Anikah, D. B. Saputra, and ..., "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Rotan," ... *Scientific Journal of* ..., 2020.
- [4] I. Anikah, A. Surip, N. P. Rahayu, and ..., "Pengelompokan Data Barang Dengan Menggunakan Metode K-Means Untuk Menentukan Stok Persediaan Barang," *KOPERTIP* ..., 2020, [Online]. Available: <http://jurnal.kopertipindonesia.or.id/index.php/kopertip/article/view/120>
- [5] P. Rahayu, I. Anikah, D. B. Saputra, and ..., "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Rotan," ... *Scientific Journal of* ..., 2020, [Online]. Available: <http://jurnal.kopertipindonesia.or.id/index.php/kopertip/article/view/118>
- [6] R. Riadi, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Analisa Penjualan Parfume," *Journal of Informatics, Electrical and Electronics* ..., 2023, [Online]. Available: <https://djournals.com/jieee/article/view/1181>
- [7] H. A. Siregar, A. Azlan, and N. Y. L. Gaol, "Penerapan Data Mining Pada Penjualan Rumah Makan Kasih Ibu Menggunakan Metode K-Means Clustering," *Jurnal Sistem Informasi* ..., 2023, [Online]. Available: <http://ojs.trigunadharma.ac.id/index.php/jsi/article/view/8955>
- [8] F. A. FITHRI, "Analisis Cluster Data Transaksi Penjualan Menggunakan K-Means Clustering Pada Toko Usaha Mandiri," 2023, [Online]. Available: [https://repository.mercubuana.ac.id/73526/%0Ahttps://repository.mercubuana.ac.id/73526/1/FauziaAllamatulFithri-41517310015-COVER.pdf](https://repository.mercubuana.ac.id/73526/%0Ahttps://repository.mercubuana.ac.id/73526/1/Fauzia>AllamatulFithri-41517310015-COVER.pdf)
- [9] W. T. Pambudi and A. Witanti, "Penerapan Algoritma K-Means Clustering Untuk Menganalisis Penjualan Pada Toko Ayu Collection Barbasis Web," *Jurnal Informatika Universitas Pamulang*, 2021, [Online]. Available: <https://www.neliti.com/publications/466160/penerapan-algoritma-k-means-clustering-untuk-menganalisis-penjualan-pada-toko-ayu>
- [10] G. Aprilianur and E. L. Hadisaputro, "Penerapan Data Mining Menggunakan Metode K-Means Clustering Untuk Analisa Penjualan Toko Myam Hijab Penajam," *JUPITER (Jurnal Penelitian Ilmu* ..., 2022, [Online]. Available: <https://jurnal.polsri.ac.id/index.php/jupiter/article/view/4684>
- [11] N. Azmi, H. S. Hafsah, Y. Yuyun, and H. Hazriani, "Penerapan Metode K-Means Clustering Dalam Mengelompokkan Data Penjualan Obat pada Apotek M23," *Prosiding SISFOTEK*, 2023, [Online]. Available: <http://www.seminar.iaii.or.id/index.php/SISFOTEK/article/view/407>
- [12] A. Rahman and H. Suroyo, "Analisis Data Produk Elektronik di E-Commerce dengan

Metode Algoritma K-Means Menggunakan Python,” *Journal of Advances in Information ...*, 2021, [Online]. Available: <https://journal.ittelkom-sby.ac.id/jaiit/article/view/158>