IMPLEMENTASI METODE CLUSTERING PADA DISTRO BINTANG PLUSH STORE SEBAGAI MEDIA UNTUK MENINGKATKAN PROMOSI

S M Fathur Rahman, A Haidar Mirza, Jemakmun, Muhammad Nasir
Teknik Informatika, Universitas Bina Darma
Jl. Jenderal Ahmad Yani, Kecamatan Seberang Ulu I, Kota Palembang, Sumatera Selatan, Indonesia
smfatur@gmail.com

ABSTRAK

Industri fesyen ritel, khususnya sektor distro, mengalami pertumbuhan pesat di Indonesia seiring meningkatnya minat generasi muda terhadap gaya hidup dan ekspresi diri. Distro Bintang Plush Store, sebagai salah satu pelaku dalam industri ini, menghadapi tantangan dalam efektivitas strategi promosinya yang masih bersifat umum dan kurang personal. Penelitian ini bertujuan untuk mengatasi permasalahan tersebut melalui segmentasi pelanggan berbasis data transaksi menggunakan algoritma K-Means Clustering. Metodologi penelitian mencakup tahapan identifikasi masalah, studi literatur, pengumpulan data transaksi, pra-pemrosesan data (pemilihan fitur, pembersihan, dan standardisasi), serta penerapan K-Means menggunakan Python pada platform *Google Colaboratory*. Dua fitur utama, yaitu *Quantity* dan *TotalAmount*, digunakan dalam proses clustering. Hasil analisis menunjukkan bahwa jumlah klaster optimal adalah dua, yaitu Klaster 0 (Pelanggan Reguler) dan Klaster 1 (Pelanggan Prioritas/Volume Tinggi). Evaluasi dengan *Silhouette Score* sebesar 0.50 menandakan kualitas clustering yang cukup baik. Temuan ini memungkinkan Distro Bintang Plush Store untuk menerapkan strategi promosi yang lebih terpersonalisasi, meningkatkan efisiensi pemasaran, serta memperkuat loyalitas pelanggan. Penelitian ini juga memberikan kontribusi praktis dalam pengembangan strategi pemasaran berbasis data pada sektor ritel fesyen lokal.

Kata kunci : *Segmentasi Pelanggan, K-Means Clustering, Data Mining, Promosi, Python, Google Colaboratory.*

1. PENDAHULUAN

Industri fesyen dan ritel di Indonesia menunjukkan dinamika yang sangat tinggi, didorong oleh demografi usia muda yang besar dan preferensi terhadap gaya hidup serta ekspresi diri melalui pakaian. Di tengah tren ini, industri distro telah muncul sebagai segmen yang vital dan berkembang pesat, menawarkan produk-produk fesyen alternatif yang unik dan personal. Persaingan dalam sektor ini semakin ketat, menuntut setiap pelaku usaha untuk tidak hanya berinovasi dalam produk, tetapi juga dalam strategi pemasaran agar dapat bertahan dan terus tumbuh di pasar yang kompetitif.

Salah satu pemain dalam industri ini adalah Distro Bintang Plush Store, yang didirikan pada 17 Oktober 2017. Berawal dari skala rumahan, Distro Bintang Plush Store telah berhasil berkembang menjadi toko fisik, menunjukkan kapasitas adaptasi dan potensi pasar yang kuat. Namun, seiring dengan pertumbuhan dan intensitas persaingan, muncul tantangan signifikan terkait efektivitas strategi promosi. Promosi yang selama ini dijalankan cenderung bersifat massal, kurang personal, dan akibatnya, kurang optimal dalam menarik minat serta membangun loyalitas pelanggan jangka panjang. Idealnya, strategi promosi harus mampu menyentuh preferensi individu pelanggan, menciptakan pengalaman yang lebih relevan dan mengikat.

Kesenjangan antara kondisi promosi saat ini dan kondisi ideal inilah yang menjadi permasalahan inti dalam penelitian ini. Pemasaran yang digunakan Distro Bintang Plush Store saat ini belum mampu mengolah dan memanfaatkan data transaksi pelanggan secara maksimal. Padahal, data ini merupakan "tambang emas" informasi yang mengandung jejak pola pembelian, preferensi produk, dan karakteristik demografi pelanggan. Pemanfaatan data ini secara strategis dapat membuka jalan menuju segmentasi pelanggan yang lebih akurat dan personalisasi promosi yang lebih efektif.

Untuk mengatasi inefisiensi promosi dan memaksimalkan nilai dari data pelanggan, metode clustering, khususnya algoritma K-Means, menawarkan solusi yang prospektif. Dengan menerapkan K-Means, Distro Bintang Plush Store dapat mengelompokkan pelanggan ke dalam segmen-segmen homogen berdasarkan perilaku pembelian mereka. Proses pengelompokan data ini akan dilakukan menggunakan bahasa pemrograman Python dengan memanfaatkan lingkungan komputasi berbasis cloud Google Colaboratory (Google Colab), yang menyediakan kemudahan akses terhadap sumber daya komputasi dan library yang dibutuhkan untuk analisis data dan implementasi algoritma machine learning. Hasil segmentasi ini akan menjadi dasar kuat untuk merancang kampanye promosi yang tidak hanya tepat sasaran, tetapi juga relevan dan personal bagi setiap kelompok pelanggan, sehingga meningkatkan tingkat respons dan konversi.

Pendekatan ini didukung oleh relevansi hasil penelitian terdahulu. Studi terdahulu mengenai "Optimalisasi strategi pemasaran dengan segmentasi pelanggan menggunakan penerapan K-means clustering pada transaksi online retail" menunjukkan peningkatan tingkat respons kampanye pemasaran online ketika promosi disesuaikan dengan segmen

pelanggan[1]. Demikian pula, penelitian yang menegaskan kapabilitas clustering dalam mengidentifikasi pola pembelian dan mendukung rekomendasi produk yang lebih personal, berkontribusi pada kepuasan dan pembelian ulang pelanggan[2].

Mengingat potensi peningkatan signifikan dalam efektivitas promosi dan daya saing, penelitian ini menjadi urgensi tinggi bagi Distro Bintang Plush Store. Dengan mengimplementasikan metode clustering, diharapkan Distro Bintang Plush Store tidak hanya mampu meningkatkan penjualan dan membangun loyalitas pelanggan, tetapi juga mencapai pertumbuhan bisnis yang berkelanjutan. Selain itu, penelitian ini juga diharapkan dapat menjadi referensi berharga bagi pengembangan strategi pemasaran yang lebih cerdas dan adaptif di seluruh industri distro, mendorong inovasi dalam praktik pemasaran yang berbasis data.

2. TINJAUAN PUSTAKA

Berdasarkan penelitian yang dilakukan oleh Suraya dengan judul "Evaluation of Data Clustering Accuracy using K-Means Algorithm", penelitian ini mengevaluasi akurasi clustering data menggunakan algoritma K-Means pada wine dataset. Proses evaluasi dilakukan menggunakan metode KDD, Silhouette, dan Davies Bouldin Index (DBI). Hasil penelitian menunjukkan bahwa normalisasi data mampu meningkatkan evaluasi clustering, dengan hasil optimal ditemukan pada pembagian 3 klaster yang terdiri dari Superior, Intermediate, dan Standard Variety[3].

Penelitian selanjutnya oleh Tabianan yang berjudul "K-Means Clustering Approach for Intelligent Customer Segmentation" bertujuan untuk mengelompokkan pelanggan berdasarkan perilaku belanja guna meningkatkan profitabilitas bisnis. Dengan menggunakan metode K-Means Clustering, penelitian ini berhasil menghasilkan klaster perilaku pelanggan yang mempermudah personalisasi layanan, meningkatkan eksposur produk, serta memberikan dampak positif terhadap keuntungan bisnis[4].

2.1. Segmentasi Pelanggan

Dalam pasar global yang semakin terfragmentasi, memahami konsumen secara mendalam adalah kunci keberhasilan. Mustahil untuk memuaskan semua orang dengan satu pendekatan "satu ukuran cocok untuk semua". Oleh karena itu, segmentasi pelanggan muncul sebagai strategi pemasaran fundamental. Segmentasi pelanggan adalah proses strategis membagi pasar yang luas atau heterogen menjadi kelompok-kelompok konsumen yang lebih kecil dan lebih homogen[5].

Heterogenitas pasar dapat disebabkan oleh perbedaan kebutuhan, preferensi, perilaku, atau karakteristik demografis. Dengan mengelompokkan konsumen yang memiliki atribut serupa, perusahaan dapat mengembangkan pemahaman yang lebih tajam tentang setiap segmen dan merancang bauran pemasaran (produk, harga, tempat, promosi) yang disesuaikan secara spesifik untuk memenuhi kebutuhan unik kelompok tersebut[1].

Penerapan segmentasi pelanggan tidak hanya sekadar praktik pemasaran yang baik; ini adalah keharusan strategis yang membawa berbagai manfaat signifikan bagi perusahaan:

a. Peningkatan Efektivitas Pemasaran: Dengan fokus pada segmen tertentu, pesan promosi dapat dirancang agar sangat relevan dan menarik, menggunakan bahasa dan saluran yang paling cocok untuk segmen tersebut. Ini mengarah pada peningkatan dramatis dalam tingkat respons kampanye dan konversi penjualan.
b. Optimalisasi Penggunaan Sumber Daya: Sumber daya pemasaran, baik finansial maupun waktu, seringkali terbatas. Segmentasi memungkinkan perusahaan untuk mengalokasikan sumber daya ini secara lebih bijak, memfokuskannya pada segmen-segmen yang paling menjanjikan dan memiliki potensi profitabilitas tertinggi, sehingga menghindari pemborosan pada segmen yang kurang responsif.
c. Pengembangan Produk dan Layanan yang Lebih Baik: Dengan pemahaman yang mendalam tentang kebutuhan dan keinginan spesifik setiap segmen, perusahaan dapat mengidentifikasi peluang untuk mengembangkan atau memodifikasi produk dan layanan yang lebih sesuai dengan harapan pasar, bahkan menciptakan solusi inovatif yang belum ada.
d. Peningkatan Kepuasan dan Loyalitas Pelanggan: Ketika pelanggan merasa bahwa suatu merek atau produk "memahami" mereka dan memenuhi kebutuhan unik mereka, pengalaman pelanggan menjadi lebih positif. Ini tidak hanya meningkatkan kepuasan tetapi juga menumbuhkan rasa loyalitas yang kuat, mendorong pembelian berulang dan *word-of-mouth* positif.
e. Penciptaan Keunggulan Kompetitif Berkelanjutan: Kemampuan untuk secara efektif melayani segmen pasar tertentu dengan penawaran yang disesuaikan memberikan perusahaan keunggulan kompetitif yang sulit ditiru oleh pesaing yang mungkin masih menggunakan pendekatan massal. Ini memungkinkan perusahaan untuk menargetkan celah pasar dan membangun posisi yang kuat.

2.2. Data Mining

Data mining adalah proses untuk menemukan pola dalam data berukuran besar guna mengungkap pengetahuan yang sebelumnya tidak diketahui. Proses ini melibatkan penggunaan berbagai teknik untuk mengidentifikasi pola-pola tersembunyi dalam data yang telah dikumpulkan[6]. Data mining, yang juga dikenal sebagai *Knowledge Discovery in Database* (KDD), menggunakan data historis untuk menemukan

keteraturan, pola, atau hubungan dalam set data yang besar, dengan tujuan menggali informasi yang berguna dari database[7].

Hasil dari data mining ini dapat digunakan untuk meningkatkan pengambilan keputusan di masa depan.

Menurut [8], "Data mining adalah proses untuk menemukan korelasi atau pola dari ratusan atau ribuan *field* dalam database yang besar." Sementara itu, dalam penelitian [9], dijelaskan bahwa "Data mining adalah kegiatan mengumpulkan dan mengolah data berukuran besar untuk menemukan pola yang dapat disimpan dalam database."

Berdasarkan definisi di atas, dapat disimpulkan bahwa data mining adalah proses pencarian pola yang diinginkan dalam database besar untuk mendukung pengambilan keputusan di masa mendatang. Pola-pola tersebut diidentifikasi oleh perangkat khusus yang mampu memberikan analisis data yang bernilai dan berwawasan, yang selanjutnya dapat dipelajari secara lebih mendalam, mungkin dengan bantuan sistem pendukung keputusan lainnya.

Data mining berfokus pada pencarian pola-pola tersembunyi dalam kumpulan data berukuran besar yang tersimpan dalam basis data seperti data warehouse atau sumber penyimpanan lainnya. Proses ini melibatkan ekstraksi informasi penting dari volume data yang besar, seperti dalam bidang penelitian medis dan manajemen hubungan pelanggan. Data mining juga mencakup upaya melindungi data pribadi dengan memodifikasi data asli menggunakan algoritma tertentu.

Dengan demikian, data mining adalah serangkaian proses yang menggunakan teknik untuk mengidentifikasi informasi dan pengetahuan dari data yang tersedia. Hasilnya dapat dimanfaatkan untuk pengambilan keputusan di masa yang akan datang.

2.3. Clustering

Clustering adalah metode untuk mengelompokkan data atau objek-objek yang memiliki kesamaan satu sama lain ke dalam kelompok yang sama, yang berbeda dari objek-objek di kelompok lainnya. Secara umum, clustering digunakan untuk menemukan dan mengelompokkan data berdasarkan karakteristik yang mirip di antara data tersebut.

[10] clustering sebagai "proses partisi sekumpulan data menjadi himpunan bagian yang disebut kelas atau cluster." bahwa "*clustering* adalah sekumpulan objek di mana objek-objek dalam satu kelompok memiliki kemiripan satu sama lain dan berbeda dengan kelompok objek lainnya."

2.4. Partitional Clustering

Partitional clustering adalah jenis metode *clustering* yang membagi data menjadi sejumlah kelompok (*cluster*) yang tidak tumpang tindih. Artinya, setiap data point hanya dapat menjadi anggota dari satu *cluster*[11].

Metode ini bertujuan untuk menghasilkan partisi data yang optimal berdasarkan kriteria tertentu, seperti meminimalkan jarak antara data point dengan pusat cluster atau memaksimalkan kesamaan antar data point dalam cluster yang sama[12].

Beberapa contoh algoritma *partitional clustering* yang umum digunakan adalah:

a. K-Means: Algoritma ini bertujuan untuk meminimalkan jumlah kuadrat jarak antara data point dengan pusat cluster (centroid). K-Means cocok untuk data yang berbentuk bulat atau elips.
b. K-Medoids: Mirip dengan K-Means, tetapi menggunakan data point aktual sebagai pusat cluster (medoid). K-Medoids lebih robust terhadap outlier dibandingkan K-Means.
c. CLARA (Clustering Large Applications): Metode ini dirancang untuk menangani dataset yang besar. CLARA memilih sampel data point secara acak dan menerapkan K-Medoids pada sampel tersebut.

Partitional clustering memiliki kelebihan dalam kemudahan implementasi dan efisiensi komputasi. Namun, metode ini juga memiliki beberapa keterbatasan, seperti sensitivitas terhadap inisialisasi awal dan kesulitan dalam menangani data dengan bentuk cluster yang kompleks.

2.5. K-Means

K-Means merupakan teknik pengelompokan data yang didasarkan pada metode *Partitioned Clustering*. Proses kerja pengelompokan ini dilakukan secara bertahap, mirip dengan *hierarchical clustering*. K-Means memiliki keunggulan dalam mengelompokkan data dalam jumlah besar dengan waktu komputasi yang relatif cepat dan efisien. Namun, kelemahan K-Means terletak pada penentuan pusat awal cluster yang dapat mempengaruhi hasil akhir.

[13] mengungkapkan bahwa "K-Means adalah algoritma non-hirarkis yang berasal dari metode pengelompokan data, di mana algoritma ini memulai dengan partisi klaster [14], "K-Means adalah algoritma pengelompokan iteratif yang membagi kumpulan data menjadi sejumlah K cluster yang telah ditentukan sebelumnya."

Ada banyak metode yang dapat digunakan dalam pengelompokan K-Means, salah satunya adalah metode non-hirarki yang membagi data menjadi dua atau lebih kelompok. K-Means merupakan metode analisis yang bertujuan untuk membagi N objek pengamatan ke dalam K kelompok, di mana setiap objek pengamatan akan ditempatkan dalam kelompok dengan rata-rata (*mean*) terdekat. Setiap *cluster* dioptimalkan berdasarkan kriteria partisi seperti perbedaan jarak, sehingga objek-objek dalam satu cluster memiliki kesamaan, sedangkan objek-objek di cluster yang berbeda memiliki perbedaan yang signifikan berdasarkan atribut dataset.

Algoritma K-Means dikenal karena kemudahannya dalam penggunaannya serta kemampuannya mengelola sejumlah besar data dan

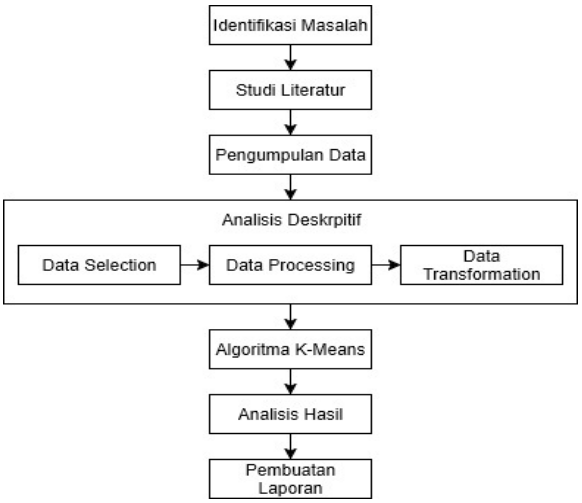
mengidentifikasi *outliers*. Langkah-langkah dalam Algoritma K-Means, seperti yang dijelaskan oleh [15], adalah sebagai berikut:

a. Tentukan nilai K sebagai jumlah *cluster* yang akan dibentuk.
b. Pilih titik pusat awal dari setiap *cluster*.
c. Hitung jarak setiap data input dengan centroid menggunakan rumus jarak Euclidean sampai ditemukan jarak terdekat antara data dengan centroid.
   Persamaan Euclidean Distance adalah sebagai berikut:
   $$D(X,Y)\sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2} \quad (1)$$
d. Klasifikasikan data berdasarkan kedekatannya dengan centroid.
e. Hitung kembali pusat cluster menggunakan anggota cluster yang saat ini ada. Pusat cluster adalah rata-rata dari semua data objek dalam cluster tertentu.

Ulangi perhitungan untuk setiap objek menggunakan pusat cluster yang baru. Jika pusat cluster tidak berubah lagi, maka proses klustering selesai. Jika belum, kembali ke langkah ketiga hingga pusat cluster tidak lagi berubah.

3. METODE PENELITIAN

3.1. Desain Penelitian



Gambar 1. Desain penelitian

Berdasarkan gambar alur penelitian, tahapan-tahapan yang dilakukan dijelaskan sebagai berikut:

a. Identifikasi Masalah
   Langkah awal adalah merumuskan masalah agar penelitian terarah, yaitu terkait penjualan pada Distro Bintang Plush Store.
b. Studi Literatur
   Mengumpulkan teori dari buku, jurnal, dan sumber online untuk memperkuat dasar penelitian.
c. Pengumpulan Data
   Data diperoleh melalui:
   - *Observasi* langsung pada aktivitas penjualan.
   - *Wawancara* dengan pihak terkait untuk mendapatkan *informasi* yang akurat.
d. Analisis Deskriptif
   Meliputi tahapan:
   - *Pemilihan Data* – memilih data sesuai kebutuhan analisis.
   - *Pengolahan Data* – membersihkan dan memilih atribut penting.
   - *Transformasi Data* – menyesuaikan format data untuk analisis.
   - *Clustering K-Means* – menerapkan algoritma K-Means menggunakan Python di Google Colab.
e. Analisis Data
   Menilai hasil clustering berdasarkan *Silhouette Score* sebagai indikator kualitas.
f. Penulisan Laporan
   Menyusun laporan berdasarkan struktur sistematis: pendahuluan, teori, metode, hasil, pembahasan, dan penutup, disertai lampiran hasil penelitian.

3.2. Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang bersifat kuantitatif. Data ini diperoleh dari sumber internal Distro Bintang Plush Store.

a. Jenis Data
   Data yang akan digunakan adalah data transaksi pelanggan, yang mencakup informasi detail mengenai setiap pembelian yang dilakukan oleh pelanggan.
b. Sumber Data
   Data transaksi pelanggan akan diperoleh dari sistem pencatatan penjualan Distro Bintang Plush Store, mencakup periode mulai tanggal pendirian distro (17 Oktober 2017) hingga periode data terbaru yang tersedia dan memungkinkan untuk dianalisis. Data ini mencakup atribut-atribut penting seperti:
   - *CustomerID* – untuk mengidentifikasi setiap pelanggan secara unik.
   - *ProductID* – untuk mengidentifikasi produk yang dibeli dalam setiap transaksi.
   - *Quantity* – menunjukkan jumlah unit produk yang dibeli.
   - *Price* – merupakan harga satuan dari produk.
   - *TransactionDate* – mencatat tanggal transaksi dilakukan, berguna dalam analisis recency.
   - *PaymentMethod* – metode pembayaran yang digunakan oleh pelanggan.
   - *StoreLocation* – lokasi toko tempat transaksi dilakukan.
   - *ProductCategory* – kategori produk yang dibeli, memungkinkan analisis berdasarkan jenis produk.
   - *DiscountApplied* (%) – persentase diskon yang diberikan pada transaksi.
   - *TotalAmount* – total nilai moneter dari setiap transaksi, dihitung dari harga, kuantitas, dan diskon yang diterapkan.

4. HASIL DAN PEMBAHASAN

4.1. Deskripsi Data

Penelitian ini menggunakan data transaksi pelanggan dari sistem internal Distro Bintang Plush Store. Data bersifat kuantitatif dan sekunder, mencakup total 25.056 baris dengan 10 atribut utama, seperti *CustomerID*, *ProductID*, *Quantity*, *Price*, *TransactionDate*, *PaymentMethod*, *StoreLocation*, *ProductCategory*, *DiscountApplied(%)*, dan *TotalAmount*.

Fokus analisis dilakukan pada dua fitur numerik, yaitu:

a. Quantity – menunjukkan jumlah unit produk yang dibeli dalam tiap transaksi.
b. TotalAmount – menunjukkan total nilai transaksi setelah diskon.

Kedua atribut ini dipilih karena dapat mencerminkan perilaku pembelian pelanggan dan digunakan sebagai dasar untuk proses clustering.

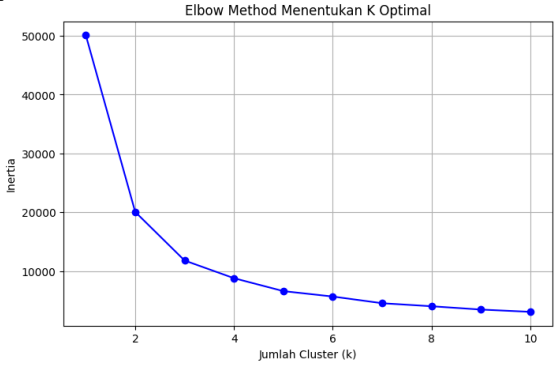
4.2. Pra-pemrosesan Data

Pra-pemrosesan dilakukan untuk memastikan data siap dianalisis oleh algoritma K-Means. Tahapannya mencakup:

a. Pemilihan Fitur: *Quantity* dan *TotalAmount* dipilih karena memiliki variasi tinggi dan relevan untuk segmentasi.
b. Pembersihan Data: Termasuk pengecekan tipe data, penanganan nilai hilang (NaN), penghapusan duplikasi, dan konsistensi data (misalnya menghindari nilai negatif).
c. Transformasi Data: Dilakukan standardisasi menggunakan *StandardScaler* untuk menyamakan skala nilai kedua fitur, karena K-Means sensitif terhadap perbedaan skala.

4.3. Analisis Clustering dengan K-Means

Algoritma K-Means diterapkan untuk mengelompokkan data pelanggan ke dalam klaster yang memiliki karakteristik serupa. Jumlah klaster optimal ditentukan menggunakan Metode Elbow, yang menunjukkan nilai $K = 2$ sebagai titik "siku" penurunan inertia.



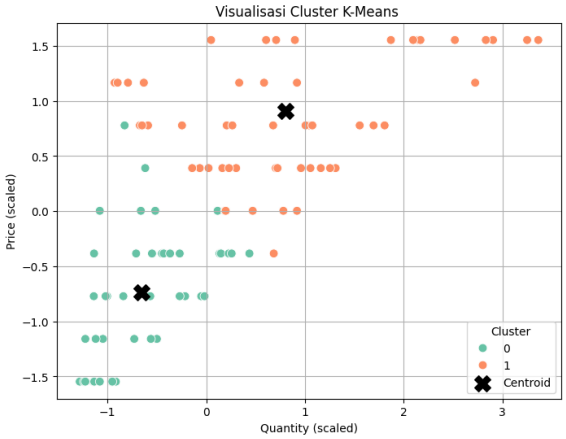
Gambar 2. Grafik elbow metode untuk penentuan k optimal)

Dengan nilai $K = 2$, proses K-Means dilakukan menggunakan Python di Google Colab, menghasilkan dua klaster dengan distribusi sebagai berikut:

Tabel 2. Distribusi jumlah data per klaster

ID Klaster	Jumlah Data
0	13760
1	11296

Untuk memperjelas pemisahan antar klaster, scatter plot digunakan berdasarkan fitur *Quantity* dan *TotalAmount*.



Gambar 3. Hasil k-means clustering dengan k=2

4.4. Interpretasi Klaster

Hasil clustering memperlihatkan dua kelompok pelanggan dengan karakteristik yang berbeda:

a. Klaster 0 (Pelanggan Reguler): Total nilai transaksi sebesar Rp 1.737.419 dan jumlah unit sebanyak 42.318. Mewakili pelanggan dengan transaksi stabil dan rutin namun bernilai sedang.
b. Klaster 1 (Pelanggan Prioritas): Total nilai transaksi mencapai Rp 4.468.083 dan jumlah unit 82.726. Mewakili pelanggan dengan kontribusi tinggi terhadap penjualan. (Lihat: *Tabel 4.3. Rata-rata Fitur per Klaster*)

Visualisasi scatter plot menunjukkan pemisahan dua klaster secara jelas, yang memperkuat interpretasi profil pelanggan.

4.5. Evaluasi Hasil Clustering

Evaluasi dilakukan menggunakan Silhouette Score, yaitu metrik untuk menilai sejauh mana suatu data cocok dengan klasternya sendiri dan berbeda dengan klaster lainnya. Nilai skor yang diperoleh adalah 0.50, menunjukkan pemisahan klaster yang cukup baik dan konsisten.

4.6. Pembahasan

Hasil clustering yang diperoleh melalui implementasi algoritma machine learning pada data pelanggan Distro Bintang Plush Store berhasil mengungkap pola perilaku pembelian yang sebelumnya tersembunyi dalam data mentah. Dengan pendekatan segmentasi berbasis data ini, perusahaan mampu mengelompokkan pelanggan ke dalam beberapa klaster yang memiliki karakteristik berbeda. Hal ini sangat penting dalam dunia bisnis modern,

karena memungkinkan penerapan strategi yang lebih terarah, efisien, dan tepat sasaran.

4.7. Klaster 0: Pelanggan Reguler

Pelanggan dalam klaster ini memiliki frekuensi pembelian yang relatif stabil, namun dengan nilai transaksi yang cenderung sedang atau rendah. Segmentasi ini menunjukkan bahwa mereka adalah pelanggan tetap yang memiliki potensi loyalitas, meskipun kontribusi nilai belanja mereka belum signifikan. Strategi yang dapat diterapkan terhadap klaster ini meliputi:

a. Program loyalitas seperti sistem poin atau kartu member untuk meningkatkan frekuensi pembelian.
b. Promosi reguler, seperti diskon bulanan atau bundling produk, agar mereka merasa dihargai dan terus melakukan pembelian.
c. Upselling dan cross-selling produk yang sesuai dengan histori pembelian mereka untuk meningkatkan nilai transaksi rata-rata per pelanggan.

Dengan pendekatan ini, diharapkan pelanggan reguler dapat ditingkatkan menjadi pelanggan bernilai tinggi di masa depan.

4.8. Klaster 1: Pelanggan Prioritas

Pelanggan dalam klaster ini merupakan segmen paling bernilai, ditandai dengan frekuensi pembelian tinggi dan total belanja yang besar. Kelompok ini berperan signifikan terhadap pendapatan perusahaan dan oleh karena itu perlu dikelola secara khusus agar tidak berpindah ke kompetitor. Strategi pengelolaan pelanggan prioritas ini mencakup:

a. Pemberian layanan premium, seperti program VIP atau fast-track service.
b. Penawaran eksklusif, seperti diskon terbatas hanya untuk mereka atau akses awal terhadap produk baru.
c. Personalisasi promosi, berdasarkan preferensi pembelian atau histori interaksi dengan brand.
d. Penguatan hubungan, misalnya melalui ucapan ulang tahun atau undangan event spesial.

Pendekatan yang berorientasi pada retensi ini bertujuan untuk mempertahankan loyalitas pelanggan bernilai tinggi, serta mendorong mereka menjadi brand advocate yang secara tidak langsung mempromosikan produk ke orang lain.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengimplementasikan algoritma K-Means menggunakan Python di Google Colaboratory untuk melakukan segmentasi pelanggan Distro Bintang Plush Store berdasarkan fitur Quantity dan TotalAmount. Hasil clustering menunjukkan dua klaster utama—Pelanggan Reguler dan Pelanggan Prioritas—yang masing-masing memiliki karakteristik perilaku belanja yang berbeda. Segmentasi ini memungkinkan Distro untuk merancang strategi promosi yang lebih personal dan efisien, dengan pendekatan retensi dan loyalitas yang disesuaikan. Evaluasi menggunakan Silhouette Score sebesar 0.50 menunjukkan pemisahan klaster yang cukup baik. Untuk pengembangan selanjutnya, disarankan agar fitur analisis diperluas mencakup aspek RFM, kategori produk, dan data demografis. Selain itu, membandingkan K-Means dengan algoritma clustering lain seperti GMM atau DBSCAN dapat memberikan perspektif tambahan. Distro juga disarankan untuk segera menerapkan strategi promosi berbasis klaster dan secara rutin memantau efektivitasnya melalui data penjualan setelah implementasi.

DAFTAR PUSTAKA

[1] E. F. L. Awalina And W. I. Rahayu, "Optimalisasi Strategi Pemasaran Dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering Pada Transaksi Online Retail," *Jurnal Teknologi Dan Informasi*, Vol. 13, No. 2, 2023, Doi: 10.34010/Jati.V13i2.10090.

[2] E. Febrianty, L. Awalina, And W. I. Rahayu, "Optimalisasi Strategi Pemasaran Dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering Pada Transaksi Online Retail Optimizing Marketing Strategies With Customer Segmentation Using K-Means Clustering On Online Retail Transactions," *Jurnal Teknologi Dan Informasi (Jati)*, Vol. 13, No. September, 2023.

[3] S. Suraya, M. Sholeh, And U. Lestari, "Evaluation Of Data Clustering Accuracy Using K-Means Algorithm," *International Journal Of Multidisciplinary Approach Research And Science*, Vol. 2, No. 01, 2023, Doi: 10.59653/Ijmars.V2i01.504.

[4] K. Tabianan, S. Velu, And V. Ravi, "K-Means Clustering Approach For Intelligent Customer Segmentation Using Customer Purchase Behavior Data," *Sustainability (Switzerland)*, Vol. 14, No. 12, 2022, Doi: 10.3390/Su14127243.

[5] N. Rohman And A. Wibowo, "Perbandingan Metode K-Medoids Dan Metode K-Means Dalam Analisis Segmentasi Pelanggan Mall," *Sintech (Science And Information Technology) Journal*, Vol. 7, No. 1, Pp. 49–58, 2024.

[6] S. Saputra, "Analisis Kelayakan Penerima Bantuan Covid-19 Menggunakan Metode K–Means Pada Kecamatan Sagulung Kota Batam," Universitas Putera Batam, Batam, 2021.

[7] S. Ependi And M. Akbar, "Implementasi Data Mining Pada Penjualan Produk Dengan Menggunakan Algoritma Apriori," *Bina Darma Conference On Computer Science*, Vol. 3, No. 1, 2021.

[8] F. Yunita, "Penerapan Data Mining Menggunkan Algoritma K-Means Clustring Pada Penerimaan

Mahasiswa Baru," *Sistemasi*, Vol. 7, No. 3, 2018, Doi: 10.32520/Stmsi.V7i3.388.

[9] R. F. N. Alifah And Abd. C. Fauzan, "Implementasi Algoritma K-Means Clustering Berbasis Jarak Manhattan Untuk Klasterisasi Konsentrasi Bidang Mahasiswa," *Ilkomnika: Journal Of Computer Science And Applied Informatics*, Vol. 5, No. 1, 2023, Doi: 10.28926/Ilkomnika.V5i1.542.

[10] K. Annisa, B. S. Ginting, And M. A. Syar, "Penerapan Data Mining Pengelompokan Data Pengguna Air Bersih Berdasarkan Keluhannya Menggunakan Metode Clustering Pada Pdam Langkat," *Jurnal Sistem Informasi Kaputama (Jsik)*, Vol. 6, No. 2, 2022, Doi: 10.59697/Jsik.V6i2.167.

[11] A. H. Lubis And E. Ramayana, "A Review On Appropriateness Of Partitional Clustering Algorithms In Handling Transactional Data," *International Journal Of Research And Review*, Vol. 10, No. 9, 2023, Doi: 10.52403/Ijrr.20230918.

[12] K. Chao, H. Zhao, Z. Xu, And F. Cui, "Robust Hesitant Fuzzy Partitional Clustering Algorithms And Their Applications In Decision Making," *Appl Soft Comput*, Vol. 145, 2023, Doi: 10.1016/J.Asoc.2023.110212.

[13] H. H. W. S. H. S. L. Okta Jaya Harmaja1, "Implementasi Algoritma K-Means Clustering Untuk Pengelompokkan Penyakit Pasien Pada Puskesmas Pulo Brayan," *Sains Dan Teknologi*, Vol. 5, 2023.

[14] A. P. Nanda, D. E. H. Pramono, And S. Hartati, "Menentukan Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademik Menggunakan Metode Algoritma K-Means," *Jurnal Sistem Informasi Dan Telematika*, Vol. 11, No. 1, 2020.

[15] Normah, B. Rifai, S. Vambudi, And R. Maulana, "Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis Smote," *Jurnal Teknik Komputer Amik Bsi*, Vol. 8, No. 2, 2022, Doi: 10.31294/Jtk.V4i2