

thesis adinda irayanti _ 246110100047

ID : 70c5c977241cdb5e994b6fa7931978ad0aa4f6a4



File name : thesis adinda irayanti _ 246110100047.txt

Original file size : 6 MB

Number of words : 6,782

Submitter : UMSIDA Perpustakaan

Submission date : July 2, 2026

Upload type : interface

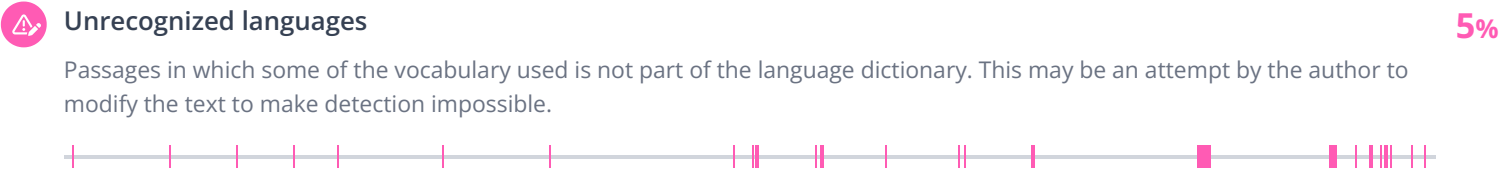
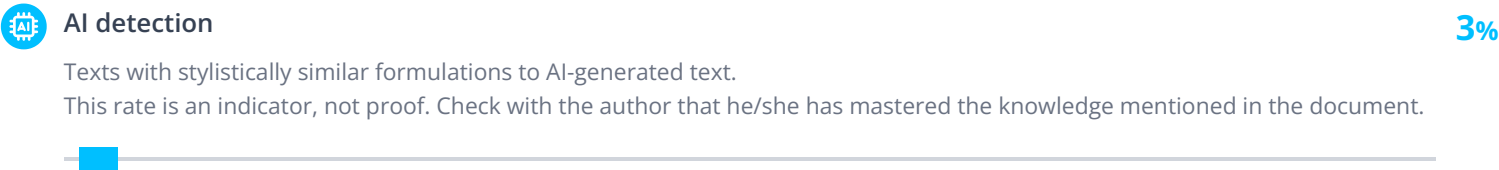
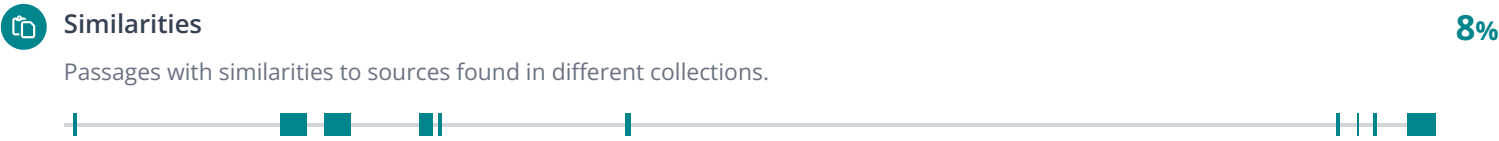
analysis end date : July 2, 2026

 Summary (section 1/3)

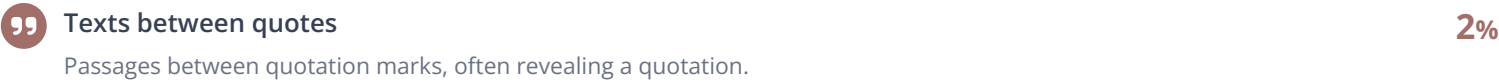
Location of suspect texts in the document :



Included in the suspicious text score :



Not included in the percentage of suspicious texts :



Similarities

Passages with similarities to sources found in different collections.

8%



Main source detected

No.		Description	Similarities	Locations
1		Artikel Ilmiah Siti Mutamimah Alfiah #ede3c1 Comes from my group	6%	
2		archive.umsida.ac.id archive.umsida.ac.id/index.php/archive/preprint/download/7023/5035... ↗	6%	
3		Garuda - Garba Rujukan Digital garuda.kemdiktisaintek.go.id/documents/detail/691537 ↗	<1%	
4		Penerapan K-Means Clustering Pada Ekspor Minyak Kelapa Sawi... prosiding.seminar-id.com/index.php/sainteks/article/view/454 ↗	<1%	
6		IMPLEMENTASI ALGORITMA K-MEANS UNTUK PEMILIHAN... doi.org/10.30813/j-j-alu.v3i2.2157 ↗	<1%	

Source with incidental similarities

No.		Description	Similarities	Locations
5		12132 Artikel-Reza #27bb1e Comes from my group	<1%	
7		Multicultural Education Management Reform with Religious... dx.doi.org/10.21070/ijemd.v1i0.554 ↗	<1%	
8		4 Algoritma K-Means Analisis Cluster: Teori, Algoritma, dan... dedenistiawan.github.io/ClustR/4-km.html ↗	<1%	
9		Analisis Data Mining Menggunakan Algoritma K-Means Clusteri... dx.doi.org/10.22441/incomtech.v8i3.4174 ↗	<1%	
10		prosiding.seminar-id.com prosiding.seminar-id.com/index.php/sainteks/article/download/454/447 ↗	<1%	
11		Mengapa Repeat Order Sangat Vital bagi Bisnis? - BAMS bams.mba/bisnis-industri/mengapa-repeat-order-sangat-vital-bagi-bisn... ↗	<1%	

Customer Segmentation Analysis and Healthy Brand Sales Potential Using K-Means Clustering, RFM, Decision Tree and Swot Analysis at PT. Sukanda Djaya

[Analisis Segmentasi Pelanggan dan Potensi Penjualan Healthy Brand Menggunakan K-Means Clustering, RFM, Decision Tree dan Analisis SWOT pada PT. Sukanda Djaya]

Adinda Irayanti Sunarjo 1), Rita Ambarwati Sukmono 2)*¹⁾

1) Program Studi Magister Manajemen, Universitas Muhammadiyah Sidoarjo, Indonesia

2) Program Studi Magister Manajemen, Universitas Muhammadiyah Sidoarjo, Indonesia

*Email Penulis Korespondensi: ritaambarwati@umsida.ac.id

Abstract : *This study aims to analyze customer segmentation and predict the sales potential of healthy brands at PT Sukanda Djaya during the COVID-19 pandemic period using a data mining approach. The method implemented is K-Means Clustering, integrated with the Recency, Frequency, Monetary (RFM) model and SWOT analysis to support strategic decision-making. The data used in this study are secondary sales data from the food service sector from 2020 to 2024. The results indicate that the clustering process produced very good quality, with silhouette score values ranging from 0.94 to 0.96. Customer segmentation successfully identified customer groups with high and low profit potential. The Diamond brand consistently dominated most clusters, reflecting strong brand performance and high customer loyalty. Meanwhile, other healthy brands such as Oatside and V-Soy still faced challenges in competing with non-healthy brands in several segments. In addition, negative values were found in billing quantity, indicating product return activities and potential issues in the supply chain. The integration of clustering results with SWOT analysis shows that the company is positioned in Quadrant I, which represents a growth strategy. This indicates that the company has strong internal strengths and significant market opportunities. Therefore, the recommended strategies include strengthening healthy brand positioning, product innovation, and distribution optimization to improve the company's competitiveness. This study contributes by integrating analytical and strategic approaches to support data-driven decision-making.*

Keywords: K-Means clustering, RFM, customer segmentation, healthy brand, SWOT, data mining

Abstrak : Penelitian ini memiliki tujuan guna melaksanakan analisa segmentasi pelanggan dan memberikan prediksi potensi penjualan *healthy brand* pada PT. Sukanda Djaya selama periode pandemi COVID-19 dengan memanfaatkan pendekatan *data mining*. Metode yang diimplementasikan yakni K-Means Clustering yang diintegrasikan melalui model Recency, Frequency, Monetary (RFM) serta analisis SWOT guna mendorong pengambilan keputusan strategis. Data yang diimplementasikan merupakan data sekunder penjualan pada sektor *food service* dari tahun 2020 hingga 2024. Hasil kajian mengemukakan bahwa proses clustering menciptakan kualitas yang sangat baik dengan nilai *silhouette score* berkisar antara 0,94 hingga 0,96. Segmentasi pelanggan berhasil mengidentifikasi kelompok pelanggan dengan potensi keuntungan tinggi dan rendah. Brand Diamond secara sistematis mendominasi sebagian besar cluster, yang menggambarkan kekuatan brand dan loyalitas pelanggan yang tinggi. Sementara itu, brand healthy lainnya seperti Oatside dan V-Soy masih mengalami hambatan saat bersaing dengan brand non-healthy pada beberapa segmen. Selain itu, ditemukan adanya nilai negatif pada *billing quantity* yang mengindikasikan adanya aktivitas *return* barang dan potensi permasalahan pada rantai pasok. Integrasi hasil clustering melalui analisa SWOT menggambarkan bahwa perusahaan berada pada kuadran I (*growth strategy*), yang mengindikasikan bahwa perusahaan mempunyai kekuatan internal yang tinggi serta peluang pasar yang besar. Maka dari itu, strategi yang direkomendasikan meliputi penguatan positioning *healthy brand*, inovasi produk, serta optimalisasi distribusi guna meningkatkan daya saing perusahaan. Kajian ini memberi kontribusi melalui integrasi pendekatan analitik serta strategis dalam mendukung pengambilan keputusan berbasis data.

Kata kunci : K-means Clustering, RFM, Segmentasi Pelanggan, healthy brand, SWOT, Data Mining.

I. PENDAHULUAN

Pandemi Covid-19 yang berlangsung sejak tahun 2020 telah mengubah banyak aspek kehidupan. Karakteristik virus corona yang terus bermutasi menyebabkan hidup manusia pun dibayangi ketidakpastian dan harus senantiasa bersiap menghadapi segala keadaan. Namun, ditengah ketidakpastian yang ditimbulkan oleh pandemi, terdapat satu hal yang pasti dalam menghadapi virus covid-19 yakni menjaga kesehatan. Menjaga kesehatan dengan cara berolahraga rutin, mengonsumsi suplemen kesehatan dan juga menjaga pola serta asupan makanan. PT. Diamond Food Indonesia merupakan salah satu perusahaan terkemuka di Indonesia yang bergerak di industri makanan dan minuman. Berdiri sejak beberapa dekade yang lalu, PT Diamond telah membangun reputasi yang kuat dengan memproduksi dan memasarkan produk berkualitas tinggi yang mencakup berbagai kategori. Sejarah perusahaan ini dimulai dari skala kecil, kemudian berkembang pesat hingga mempunyai anak perusahaan yang bernama PT. Sukanda Djaya.

Saat ini PT. Sukanda Djaya telah memiliki lebih dari 2500 sku product, Dalam 2500 sku product tersebut terdapat lima healthy brand yang penjualan nya meningkat saat pandemi covid yakni Biokul, Jungle juice, home brand diamond, outside dan V-soy. Lima Healthy brand tersebut diyakini masyarakat dapat meningkatkan imun dan juga menjaga kesehatan dalam masa pandemi karena dalam lima brand tersebut mengungkapkan komposisi yang menarik di konsumsi saat masa pandemi. Setelah meninjau dari hasil penjualan pada masa pandemi COVID-19 dari tahun 2020 hingga tahun 2023 maka team manajemen PT. Sukanda Djaya perlu melihat langkah kedepan untuk potensi untung dan rugi nya, salah satu nya dengan meneliti healthy brand yang berpotensi untuk diteliti apakah brand tersebut tetap banyak peminat atau kah berpindah ke competitor.

Gambar 1.

Data penjualan pada masa pandemi covid-19

(2020-2023)

Seperti yang terlihat pada Gambar 1 terkait data penjualan global dari 5 brand yang diringkas menjadi 3 category brand pada masa pandemi covid-19 yakni pada tahun 2020 hingga 2023 Berdasarkan pemaparan sebelumnya, perlu adanya pemilihan strategi marketing yang cocok dari 5 brand tersebut, karena tidak semua brand dapat mengimplementasikan strategi pemasaran yang sama. Saat ini PT Sukanda Djaya telah mengelompokkan segment customer berdasarkan customer transaksi tertinggi dan terendah pada segment yang telah ditentukan guna mengetahui potensi keuntungan dan kerugian penjualan berdasarkan brand category serta untuk menentukan strategi pemasaran yang digunakan di PT. Sukanda Djaya berdasarkan *brand strategy*, desain brand dan karakteristik pada brand. Data mining adalah poin penting untuk sebuah proses *Knowledge Discovery in Database* (KDD) yang didalamnya terdapat meliputi asumsi algoritma yang meliputi menelaah data, menciptakan model dan mengidentifikasi pola yang sebelumnya tidak diketahui. KDD memiliki karakteristik otomatis, mampu dimaknai dalam sebuah penerapan proses secara sistematis guna mendorong identifikasi yang tepat dan ideal, bermanfaat dan menemukan pola dari beberapa data yang kompleks [1]. Data mining adalah cara pemecahan masalah melalui proses analisa data yang tersedia di database, melalui kondisi tersebut data tersedia dan otomatis disimpan secara dimana data tersimpan melalui perangkat digital yang selanjutnya pada proses pencariannya diimplementasikan otomatis seperti pada computer.

Pada kajian ini, pendekatan yang diimplementasikan yakni Metode K-Means. Metode K-Means adalah salah satu teknik yang memiliki tujuan fungsi pengelompokan atau *clustering*. *Clustering* merupakan pendekatan analisis data berdasarkan persamaan karakteristik atau pola yang memiliki tingkat kesamaan atau perbedaan tertentu[2]. *Clustering* dikemukakan oleh Xu & Wunsch tahun 2009 memiliki makna dengan membagi objek data seperti entitas, *study case*, unit Analisa ke dalam beberapa kluster, kelompok, atau kategori tertentu yang telah disepakati[3]. Pengelompokan dalam *clustering* adalah proses partisi data ke dalam kelompok berdasarkan kedekatan geometris, di mana setiap kelompok direpresentasikan oleh pusat kluster (*centroid*) [4]. Metode statistik yang bertujuan untuk mengelompokkan objek dalam dataset ke dalam kelompok yang memiliki kesamaan karakteristik tertentu [5].

Copyright © Universitas Muhammadiyah Sidoarjo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





Copyright © Universitas Muhammadiyah Sidoarjo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Sedangkan tujuan proses *clustering* guna meminimalisir terjadi *objective function* yang ditetapkan pada tahap *clustering*, yang pada dasarnya dimanfaatkan guna meminimalkan variasi pada satu cluster maupun cluster yang lain. [6]. Algoritma metode K-Means adalah dimulai dari menentukan jumlah cluster, mengalokasikan data sesuai dengan jumlah cluster yang telah ditentukan [6], menghitung nilai centroid pada tiap-tiap cluster, mengalokasikan masing-masing data ke centroid terdekat lalu Kembali ke step 3, jika terdapat transformasi data antar cluster, atau jika terdapat perubahan nilai centroid di ambang threshold yang disepakati, atau jika terjadi pembaharuan dalam nilai objective function di atas ambang threshold yang disepakati.

Copyright © Universitas Muhammadiyah Sidoarjo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Copyright © Universitas Muhammadiyah Sidoarjo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Beberapa penulis telah melakukan penelitian mengenai K-means clustering sebagai penelitian dalam hal pengelompokan :

Tabel 1. Penelitian Terdahulu

Penulis	Tujuan Penelitian	Hasil Penelitian
Fina Nasari & Nasari Darma [7]	Menerapkan algoritma K- Means Clustering pada data penerimaan mahasiswa baru tahun ajaran 2014/2015(studi kasus : Universitas Potensi Utama).	Hasil K-Means Clustering yang diperoleh ada dua kelompok, pusat cluster dengan Cluster 1 = 1 ; 1.75; 1.5 dan Cluster 2 = 2.95; 1.65; 1.4, Cluster pertama jika asal sekolah adalah SMA atau Sekolah Menengah Pertama maka rata-rata jurusan yang diambil adalah Sistem Informasi dan kedua jika asal Sekolahnya adalah SMK rata-rata jurusan yang diambil adalah Teknik Informatika
Ari Sulistiawaty & Eko Supriyanto [8]	Meningkatkan sistem informasi terintegrasi basis web guna mengatur data penilaian yang bersumber pada laporan hasil belajar peserta didik dan pengklasifikasian peserta didik dalam kelas unggulan pada saat penerapan algoritma K- means Clustering.	Hasil kajian ditemukan bahwa aplikasi pengelolaan data penilaian yang bersumber saat penyajian laporan hasil belajar peserta didik memanfaatkan hasil nilai rapor kurikulum 2013. Penilaian itu menjadi pedoman atribut yang dimanfaatkan pada proses perhitungan guna penentuan peserta didik kelas unggulan. Hasil klarifikasi Kelas unggulan di perlukan sebuah implementasi sistem informasi



		terbentuk 2 cluster pada tiap kelas. Setiap segmentasi data perputaran pada setiap kelas selanjutnya dihasilkan jumlah peserta didik masuk kelas unggulan sebanyak 96 dan peserta didik tidak masuk kelas unggulan sebanyak 96.
Elly Muningsih & Sri kiswati [9]	Kajian ini memiliki tujuan supaya bisa membagi setiap produk yang dijual dalam online shop Ragam Jogja yang terbagi dalam beberapa cluster guna memahami produk yang paling disukai sehingga jumlah stok yang tersedia harus cukup banyak, produk best seller guna dengan kuantitas banyak dan produk dengan sedikit peminat maka stoknya dibuat terbatas.	Hasil kajian atas pengolahan data yang diimplementasikan dengan atribut kode produk, kuantitas transaksi, volume penjualan dan rata-rata penjualan, dihasilkan 3 kelompok produk bestseller berjumlah 3 produk untuk jumlah stok banyak, 11 produk diminati untuk jumlah stok sedang dan 17 produk kurang disukai untuk jumlah stok relative terbatas.
Agus Perdana Windarto [10]	Tujuan dari kajian ini yakni mengimplementasikan k-means dalam mengelompokkan ekspor buah-buahan berlandaskan negara tujuan.	Guna mendapatkan nilai dari produksi ekspor buah-buahan yang berpedoman pada negara tujuan. Data tersebut dianalisa memanfaatkan pendekatan Rapidminer guna memperoleh nilai centroid pada 3 cluster yakni cluster tingkat ekspor tinggi, sedang dan rendah. Centroid data guna cluster tingkat ekspor tinggi 904.276,5, Centroid data untuk cluster tingkat ekspor sedang 265.501 dan Centroid data guna cluster tingkat ekspor rendah 34.280,1. Sehingga diperoleh penilaian berdasarkan indeks ekspor buah-buahan

		dengan 2 negara cluster tingkat ekspor tinggi yakni Pakistan dan India, negara cluster tingkat ekspor sedang yakni Bangladesh, Singapura, dan Negara lainnya dan 6 negara cluster tingkat ekspor rendah yakni Hongkong, Vietnam, Tiongkok, Malaysia, Nepal, dan Iran
Siti Hajar [11]	Tujuan dari penelitian ini untuk clustering ekspor minyak berdasarkan negara tujuan menggunakan k-means.	Sehingga didapat penilaian berpedoman indeks ekspor minyak kelapa sawit dengan 1 negara cluster tingkat ekspor tinggi yaitu India dan 9 yang lain yakni negara cluster tingkat ekspor rendah yakni Pakistan, Tiongkok, Singapura, Malaysia, ,Bangladesh, Sri Lanka, Mesir, Belanda, dan Jerman. Hal ini dapat menjadi pengambilan kebijakan untuk pemerintah, negara yang dijadikan prioritas tertinggi pada kegiatan ekspor minyak kelapa sawit yang berpedoman pada klaster telah diimplementasikan.

Berdasarkan penelitian terdahulu, telah banyak dilakukan penelitian menggunakan K-means clustering dengan tujuan untuk mengelompokkan data sesuai dengan tujuan penelitian. Sesuai dengan latar belakang dan penelitian terdahulu diatas, maka peneliti perlu melakukan *clustering* penjualan *healthy brand* berdasarkan *cluster segments*, *product*, *billing* penjualan pada tahun era pandemi covid-19 yakni tahun 2020-2023 Penelitian ini difokuskan pada sektor pembeli, dimana pembeli dikelompokkan menjadi beberapa kelompok yaitu hotel yang dibagi menjadi hotel bintang 3 kebawah, 4 keatas dan apartement. Restaurant yang dibagi menjadi western, japanese, indian. Bakery yang dibagi international dan lokal. *Café* yang dibagi *café local* kopitiam, local dan internasional. Penelitian ini bertujuan untuk melihat potensi keuntungan dan kerugian dalam penjualan pada era covid-19 yakni dengan meneliti penjualan dari segment customer. selanjutnya di integrasikan dengan *brand strategy* pada brand. Dengan harapan, kedepannya hal ini dapat menjadi pembaruan dalam mengembangkan penjualan strategi.

II. METODE

Penelitian ini menggunakan metode kuantitatif deskriptif menggunakan K-means Clustering dengan menggunakan pendekatan model RFM dan SWOT

A. Model RFM

Model RFM (Recency, Frequency, Monetary) adalah sebuah pendekatan analisa perilaku konsumen yang berguna sebagai identifikasi serta pengklasifikasian konsumen yang berpedoman pada frekuensi transaksi, interval konsumsi, dan nilai moneter pembelian pada sebuah periode waktu tertentu. Model ini membantu dalam pengkajian data konsumen pada skala luas serta mampu mendorong pengambilan keputusan yang lebih tepat di lini pemasaran berbasis data. Lebih detail, ketiga dimensi pada model RFM adalah Recency (R): guna pengukuran seberapa baru proses bisnis terakhir dilaksanakan oleh konsumen, ditentukan sebagai jangka waktu antara pembelian periode akhir dan waktu analisa dilakukan. Nilai R yang lebih tinggi (yaitu interval waktu yang relatif pendek) memberi indikasi konsumen yang lebih aktif. Frequency(F): Memberi gambaran mengenai kuantitas proses bisnis yang dilaksanakan konsumen pada periode tertentu. Nilai F yang semakin tinggi maka semakin besar pula tingkat preferensi konsumen melaksanakan pembelian berulang, yang berpotensi sebagai peluang indikator loyalitas. dan Monetary (M): memberi pedoman bahwa total nilai uang yang dibayarkan konsumen waktu tertentu. Nilai M yang tinggi menggambarkan bahwa konsumen mempunyai nilai ekonomi signifikan bagi sebuah entitas. Ketiga variabel RFM tersebut memiliki tingkat kepentingan yang sama, sehingga dalam proses penghitungan atau pemeringkatan, bobotnya dianggap identik [12]. Namun, pandangan ini ditentang yang berpendapat bahwa kepentingan relatif dari R, F, dan M dapat berbeda-beda tergantung pada karakteristik industri dan perilaku pelanggan [13]. Oleh karena itu, penyesuaian bobot pada masing-masing dimensi sangat disarankan untuk meningkatkan akurasi segmentasi [14].

Selanjutnya, dimensi moneter (M) pada RFM juga dapat dialihkan maupun dikombinasi bersama dimensi lain yang masih linier dengan konteks industri, seperti halnya waktu konsumsi, waktu sebagaimana diimplementasikan pada industri media guna menganalisa lamanya durasi yang digunakan audiens pada konten tertentu. Pendekatan ini menghighlight bahwa fleksibilitas model RFM saat menyelaraskan metrik berbasis perilaku konsumen. Pada pengimplementasiannya, model RFM mampu dimodifikasi pada skala tertentu (misalnya 1 hingga 10), di mana tiap dimensi diberi skor yang berpedoman pada ranking. Langkah ini memungkinkan konversi data transaksional mentah menjadi insight yang terukur dan dapat dibandingkan antar pelanggan. Data yang dibutuhkan umumnya cukup sederhana, yakni mencakup nama pelanggan, tanggal pembelian terakhir, serta nilai pembelian.

B. SWOT

Swot merupakan sebuah teknis analisis strategis yang digunakan untuk mengidentifikasi dan mengevaluasi strenght, weakness, opportunities, dan threats dalam organisasi, bisnis, produk atau program. Penjelasan mengenai strenght merupakan keunggulan organisasi atau proyek yang dapat dijadikan sebagai kekuatan seperti contoh sdm yang berkualitas, teknologi yang unggul dan reputasi brand. Weakness merupakan kelemahan dalam organisasi atau proyek yang dapat menghambat kinerja, contohnya keterbatasan modal dan manajemen yang belum optimal. Opportunities merupakan peluang yang dapat dimanfaatkan untuk mengoptimalkan kinerja. Opportunities termasuk dalam faktor eksternal, contoh yang termasuk opportunities yaitu kemajuan teknologi dan pertumbuhan permintaan. Threats merupakan ancaman yang dianggap merugikan dalam suatu proyek. Threats termasuk dalam faktor eksternal, contoh yang termasuk dalam threats yaitu persaingan yang meningkat dan perubahan regulasi

C. Algoritma K-means

Algoritma K-Means merupakan salah satu algoritma clustering non-hierarki yang paling umum digunakan dalam proses data mining. Algoritma ini digunakan untuk mengklasifikasikan atau mengelompokkan sejumlah besar data ke dalam K kelompok (cluster) berdasarkan kesamaan karakteristik atau atribut yang dimiliki oleh data tersebut. Salah satu ciri khas utama dari algoritma ini adalah bahwa jumlah cluster (K) harus ditentukan sebelum proses clustering dilakukan. Setelah nilai K ditentukan, algoritma K-Means akan mengelompokkan data sedemikian rupa sehingga data yang berada dalam satu cluster memiliki tingkat kemiripan yang tinggi, sedangkan data yang berbeda karakteristik akan dikelompokkan ke dalam cluster yang berbeda. Algoritma K-Means bekerja dengan meminimalkan variasi dalam setiap cluster dan memaksimalkan perbedaan antar cluster [15]. Artinya, data dalam

satu cluster akan memiliki tingkat variasi internal yang rendah, sehingga lebih homogen, sementara perbedaan antar cluster menjadi lebih signifikan. Secara umum, langkah-langkah utama dalam algoritma K-Means dimulai dari Menentukan jumlah cluster K yang diinginkan, Menginisialisasi K centroid secara acak dari data lalu Menghitung jarak antara setiap data ke centroid terdekat menggunakan ukuran jarak tertentu (misalnya, Euclidean distance), Klasifikasi data ke berdasarkan cluster berdasarkan centroid terdekat, Menghitung ulang posisi centroid berdasarkan rata-rata posisi data dalam masing-masing cluster dan mengulangi proses hingga posisi centroid konvergen atau tidak berubah secara signifikan. Keunggulan K-Means terletak pada sederhana dan kecepatan konvergensi, meskipun algoritma ini juga memiliki keterbatasan seperti sensitivitas terhadap pemilihan awal centroid dan nilai K, serta tidak cocok untuk data dengan bentuk cluster yang tidak sferis.

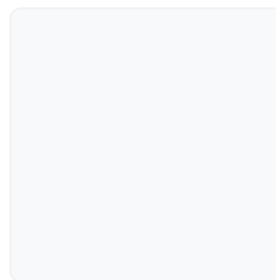
C. Clustering

Clustering adalah suatu metode pengelompokan berdasarkan ukuran kedekatan (kemiripan). Sebuah cluster tidak harus sama akan tetapi pengelompokannya berdasarkan pada kedekatan dari suatu karakteristik sampel yang ada salah satunya dengan menggunakan rumus jarak Euclidean. Metode clustering merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (similarity) antara satu data dengan data yang lainnya. Clustering termasuk salah satu metode data mining yang bersifat tanpa arahan (unsupervised), artinya metode ini diterapkan tanpa adanya latihan (training) dan tanpa ada guru (teacher) serta tidak memerlukan target output [16].

Berdasarkan data yang digunakan dalam klasterisasi untuk menilai kualitas klaster :

a. Sum of Squared Errors (SSE) atau Inertia

Mengukur total jarak kuadrat antara data dengan centroid klasternya dan Nilai SSE yang lebih kecil menunjukkan klaster yang lebih baik.



Digunakan dalam metode **Elbow Method** untuk menentukan jumlah klaster optimal. Grafik SSE vs akan menunjukkan titik "siku" (elbow) yang mewakili jumlah klaster optimal.

b. Silhouette Coefficient

Mengukur seberapa dekat data berada di dalam klasternya dibandingkan dengan klaster lain. Nilai sekitar -1 sampai dengan 1 rating 1 untuk Klaster sangat baik, 0 untuk Data berada di perbatasan dua klaster dan -1 : Klasterisasi buruk (data salah klaster).

c. Dunn Index

Mengetahui ukuran rasio antar jarak minimum antar-klaster dan ukuran maksimum dalam klaster. Nilai yang lebih besar menggambarkan klasterisasi yang lebih tepat dan relevan.

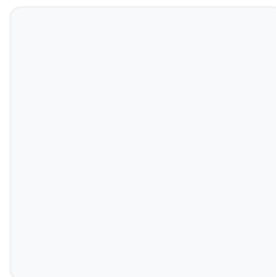
2. Evaluasi Eksternal

Memanfaatkan data dengan label atau ground truth guna menilai akurasi klaster

Rand Index (RI) memberi ukuran kesesuaian antara klaster yang diciptakan dengan label secara real. Nilai RI berkisar antara 0 hingga 1 (1 menggambarkan kesesuaian sempurna). Adjusted Rand Index (ARI) : Versi RI yang disesuaikan guna menghapus tingkat kemungkinan secara acak.

a. Purity

Mengukur persentase data dalam klaster yang benar-benar berasal dari satu kategori.



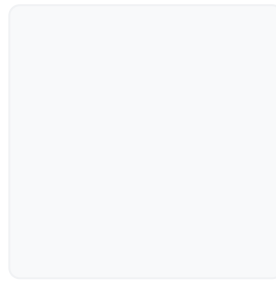
Normalized Mutual Information (NMI) untuk Mengukur kemiripan antara klaster yang dihasilkan dan label sebenarnya berdasarkan entropi informasi.

2. Pemilihan Jumlah Klaster (k) : Beberapa metode digunakan untuk menentukan jumlah klaster yang optimal yakni Elbow Method yang terdiri dari Plot SSE (inertia) terhadap k . lalu pilih k pada titik "siku" di mana penurunan SSE melambat. Silhouette Analysis Menghitung rata-rata nilai Silhouette untuk berbagai nilai k . lalu Pilih dengan nilai rata-rata tertinggi.

3. **Validasi Stabilitas** : Yang terdiri dari 2 yang pertama yakni **Cross-Validation** yaitu Membagi data menjadi beberapa subset untuk memvalidasi hasil klasterisasi pada subset yang berbeda dan **Bootstrapping** yang digunakan pengambilan sampel berulang untuk menilai konsistensi hasil klasterisasi.

C. Data Penelitian

Data dalam penelitian ini diambil dari data sekunder yang berasal dari perusahaan PT. Sukanda Djaya. Data yang digunakan merupakan data yang menggambarkan pembelian yang dilakukan konsumen yang dikelompokkan menjadi beberapa kelompok berdasarkan segment yang ada pada food service, yang kemudian akan di segmentasi ulang berdasarkan tipe dari segmen tersebut. Dalam penelitian ini diambil dari data penjualan pada era covid yakni pada tahun 2020-2022



GAMBAR 2. RESEARCH OVERVIEW DIAGRAM

Gambar diatas merupakan research overview diagram yang akan digunakan dalam penelitian ini, dimulai dari bussines overview yang telah dijelaskan pada diagram penjualan gambar 1 yang menjelaskan mengenai grafik peningkatan penjualan by brand pada masa pandemi covid-19 yakni pada billing tahun 2020 hingga 2023. Dilanjutkan dengan memahami data yang akan diolah dimulai dari seleksi data by channel, brand dan billing tahun, lalu di proses data menggunakan rapid minner untuk mengetahui nilai maksimal dan minimal nya, di proses dengan penyesuaian segmentasi RFM yang dibagi menjadi dua klustering yakni by sales performance dan hasil segment customer menggunakan k-means, keduanya dievaluasi hingga menemukan hasil dan kesimpulan.

III .RESULT AND DISCUSSION

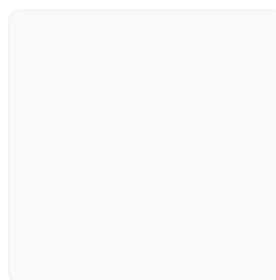
3.1. Potensi Penjualan

Fenomena pandemi COVID-19 dalam penelitian ini bertindak sebagai intervensi eksternal yang secara fundamental mengubah peta preferensi konsumen menuju gaya hidup sehat. Ketidakpastian akibat pandemi mendorong masyarakat untuk lebih memprioritaskan imunitas, yang berdampak langsung pada fluktuasi dan lonjakan penjualan kategori produk sehat PT. Sukanda Djaya (seperti Biokul, Jungle Juice, Diamond, Oatside, dan V-Soy). Oleh karena itu, analisis hasil penjualan periode 2020 hingga 2024 di bawah ini dievaluasi menggunakan pendekatan data mining untuk memetakan bagaimana perubahan perilaku tersebut membentuk segmen pasar baru.

Tabel 2 menjelaskan tentang hasil clusterisasi pada tahun 2020

Tabel 2

Clustering 2020.

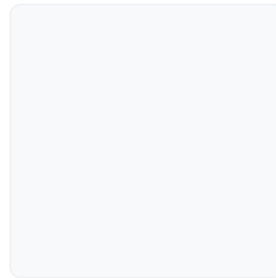


Cluster yang memiliki potensi keuntungan terbesar adalah cluster 1 dan 3. Cluster 1 tidak punya healthy brand yang dominan sama sekali sehingga perlu adanya penguatan pesan kesehatan.

Pada tabel diatas terdapat nilai negatif pada Billing Qty, hal ini bukan error namun dikarenakan pengembalian barang oleh pembeli.Setelah difilterasi dengan lima healthy brand: Biokul, Jungle Juice, Home Brand Diamond, Oatside, V-Soy. Didapat hasil seperti tabel di atas yaitu cluster 0 pada brand diamond, cluster 1 pada brand NFH , cluster 2 pada brand Nutella , cluster 3 pada brand Diamond, dan Cluster 4 pada brand MC Cain. Healthy brand yang

dominan adalah diamond di cluster 0 dan 3 kemudian outside di cluster 4. Pada cluster nomer 3 brand diamond sudah punya posisi yang dominasi sehingga untuk menjaga dominasi, perlu adanya packaging yang lebih menarik atau varian rasa yang lebih banyak. Perlu repositioning terutama pada cluster 4 dimana outside kalah secara penjualan dengan brand MC Cain sehingga perlu adanya perkuatan distribusi pada merk Outside di cluster 4. Action plan yang disarankan adalah mempertahankan dominasi Diamond dengan inovasi design packaging , mempeerkuat pesan kesehatan di cluster 1 dan 2 terutama cluster 1 karena punya potensi untung terbesar dan meningkatkan inovasi Oatside pada cluster 3 agar bisa menyaingi MC Cain dalam dominasi.

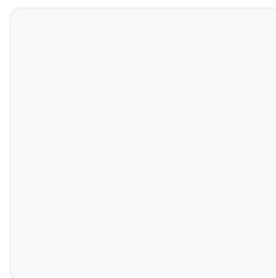
Berikut visualisasi persebaran data berdasarkan clusternya



Gambar 3.

Silhouette score 2020

Clustering pada tahun 2020 memiliki silhouette score = 0,94 yang sangat mendekati 1 yang berarti proses clustering sangat baik dan valid.



Gambar 4

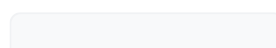
Decision Tree 2020

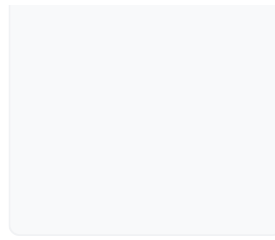
Node utama pada pohon keputusan adalah potensi keuangan karena variabel tersebut menunjukkan kemampuan cluster dalam menghasilkan keuntungan. Cluster dengan potensi keuangan lebih dari 100% dikategorikan sebagai cluster dengan potensi penjualan tinggi. Selanjutnya, nilai billing quantity digunakan sebagai pemisah kedua untuk mengidentifikasi intensitas pembelian pelanggan. Hasil Decision Tree menunjukkan bahwa cluster dengan nilai potensi keuangan tinggi dan jumlah pembelian yang besar memiliki peluang keuntungan yang lebih tinggi. Sebaliknya, cluster dengan nilai potensi keuangan rendah cenderung memiliki kontribusi penjualan yang lebih kecil dan memerlukan strategi pemasaran yang berbeda.

Tabel 3 menunjukkan hasil dari clustering pada tahun 2021 :

Tabel 3

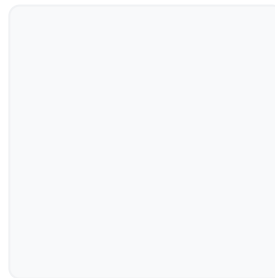
Clustering 2021





Cluster yang memiliki potensi keuntungan terbesar adalah cluster 1 dan 3. Cluster 1 punya healthy brand yang dominan yaitu V-Soy namun masih kalah penjualan dengan MC Cain sehingga perlu adanya penguatan pesan kesehatan di cluster tersebut. Kemudian hampir di seluruh cluster brand Diamond sudah punya posisi dominan sehingga untuk menjaga dominasi, perlu adanya design packaging yang lebih menarik atau varian rasa yang lebih banyak. Setelah difilterasi dengan **lima healthy brand** : Biokul, Jungle Juice, Home Brand Diamond, Oatside, V-Soy. Didapat hasil seperti tabel di atas yakni mengenai brand yang paling diminati, pada cluster 0 terdapat brand diamond, cluster 1 dengan brand MC Cain, cluster 2 dengan brand Diamond, cluster 3 brand USA dan pada cluster 4 dengan brand diamond. Pada tabel diatas menjelaskan bahwa Healthy brand yang dominan di cluster 0,2,3 dan 4 kemudian V- Soy di cluster 1. Menurut hasil tabel diatas ditemukan peluang repositioning terutama pada cluster 1 di mana V-Soy kalah penjualan terhadap MC Cain maka dari itu perlu adanya perkuatan distribusi pada merek V-Soy di cluster 1. Menurut hasil tabel diatas, Action Plan yang disarankan adalah dengan mempertahankan dominasi brand Diamond dengan inovasi design packaging, memperkuat pesan kesehatan di cluster 1 karena mempunyai potensi keuntungan yang besar dan meningkatkan inovasi V-soy pada Cluster 1 agar bisa menyaingi MC Cain dalam dominasi.

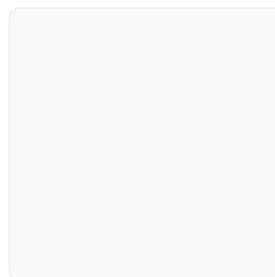
Berikut visualisasi persebaran data berdasarkan clusternya



Gambar 5

silhouette score 2021.

Terdapat nilai negatif pada Billing Qty, hal ini bukan error namun dikarenakan pengembalian barang oleh pembeli.



Gambar 6

Decision Tree 2021

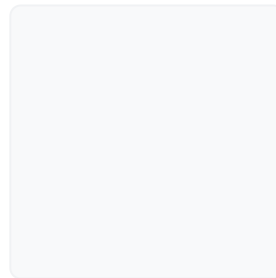
Hasil pohon keputusan menunjukkan bahwa potensi keuangan menjadi variabel utama yang menentukan tingkat

potensi penjualan suatu cluster. Cluster dengan potensi keuangan di atas 100% dikategorikan sebagai cluster dengan potensi penjualan tinggi. Selanjutnya, rata-rata billing quantity digunakan untuk membedakan tingkat kontribusi masing-masing cluster. Cluster 1 memiliki potensi penjualan tertinggi karena memiliki nilai potensi keuangan sebesar 630,78% dan rata-rata billing quantity sebesar 83,694. Sementara itu, cluster 3 termasuk kategori potensi tinggi dengan nilai potensi keuangan sebesar 202,33%. Cluster 4 dikategorikan sebagai potensi sedang, sedangkan cluster 0 dan cluster 2 termasuk kategori potensi rendah karena memiliki nilai keuntungan yang relatif kecil. Hasil ini menunjukkan bahwa Decision Tree mampu menghasilkan aturan klasifikasi yang mudah dipahami dan dapat digunakan sebagai dasar pengambilan keputusan pemasaran.

Tabel 4 menunjukkan hasil clustering pada tahun 2022

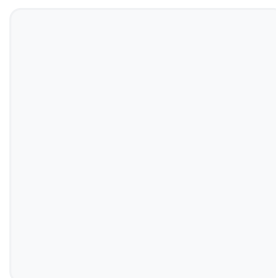
Tabel 4

Clustering 2022



Cluster yang memiliki potensi keuntungan terbesar adalah cluster 1 dan 2. Cluster 1 tidak punya healthy brand yang dominan sama sekali sehingga perlu adanya penguatan pesan kesehatan. Kemudian di cluster 2 perlu adanya penguatan pesan kesehatan dan juga inovasi pada produk Oatside agar bisa menyaingi NN. Terakhir brand Diamond sudah punya posisi dominan sehingga untuk menjaga dominasi, perlu adanya design packaging yang lebih menarik atau varian rasa yang lebih banyak. Setelah difilterasi dengan lima healthy brand: Biokul, Jungle Juice, Home Brand Diamond, Oatside, V-Soy. Didapat hasil seperti tabel di atas yaitu brand yang paling diminati pada cluster 0 adalah diamond, cluster 1 dan 2 pada brand yang sama yaitu NN, pada cluster 3 pada brand diamond dan pada cluster 4 pada brand USA. Healthy Brand yang dominan adalah Diamond di cluster 0, 3, dan 4 kemudian Oatside di cluster 2. Dari hasil tabel diatas ditemukan peluang repositioning terutama pada cluster 2 di mana oatside kalah penjualan terhadap brand NN, maka dari itu perlu ada nya perkuatan distribusi pada merek oatside di cluster. Maka dari itu action plan yang disarankan yaitu mempertahankan dominasi diamond dengan inovasi design packaging, memperkuat pesan kesehatan di cluster 1 dan 2, namun paling yang utama pada cluster 1 karena mempunyai potensi untung yang besar, dan meningkatkan inovasi oatside pada cluster 2 agar bisa bersaing dengan brand NN yang mendominasi.

1. Berikut visualisasi persebaran data berdasarkan clusternya :

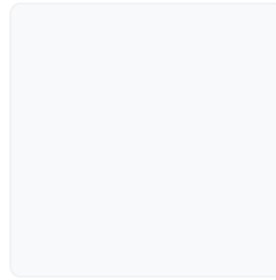


Gambar 7

silhouette score 2022

Clustering pada tahun 2022 memiliki silhouette score = 0,95 yang sangat mendekati 1 yang berarti proses clustering

sangat baik dan valid.



Gambar 8

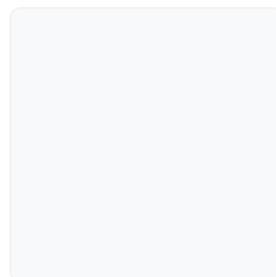
Decision Tree 2022

Hasil pohon keputusan mengemukakan bahwa potensi keuangan adalah faktor pokok guna mengidentifikasi tingkat potensi penjualan sebuah cluster. Cluster yang mempunyai potensi nilai keuangan di atas 50% masuk kategori kelompok peluang keuntungan cukup tinggi. Kemudian, variabel billing quantity dimanfaatkan guna memberi perbedaan tingkat kontribusi pembelian antar cluster. Cluster 2 mempunyai nilai billing quantity tertinggi yang menjadikannya sebagai kategori cluster potensi sangat tinggi. Selain itu, cluster 1 masuk kategori sebagai potensi tinggi karena meskipun menciptakan keuntungan besar, volume pembeliannya relative lebih rendah. Cluster 3 dan cluster 4 mempunyai tingkat potensi yang relatif rendah sehingga entitas seharusnya mengimplementasikan strategi peningkatan promosi, distribusi serta penguatan positioning produk guna memberi kontribusi baik secara langsung maupun tidak langsung terhadap penjualan.

Tabel 5 Menunjukkan hasil dari clustering pada tahun 2023

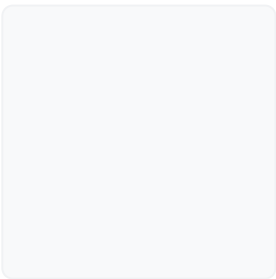
Tabel 5

Clustering 2023



Cluster yang memiliki potensi keuntungan terbesar adalah cluster 1. Cluster 1 didominasi oleh healthy brand Diamond sehingga perlu adanya konsistensi dalam menjaga dominasi. Selain itu, di Diamond juga punya posisi dominan di hampir seluruh cluster selain cluster 4 sehingga untuk menjaga dominasi, perlu adanya design packaging yang lebih menarik atau varian rasa yang lebih banyak. Terakhir, perlu adanya penguatan pesan kesehatan di cluster 4 karena healthy brand yang dominan di cluster tersebut yaitu V-Soy masih kalah penjualan dengan MC Cain. Setelah difilterasi dengan lima healthy brand: Biokul, Jungle Juice, Home Brand Diamond, Oatside, V-Soy. Didapat hasil seperti tabel di atas yaitu Brand yang paling diminati pada cluster 0 dan 1 yaitu brand yang sama yakni diamond, cluster 2 dengan brand LKK, cluster 3 pada brand diamond dan 4 brand MC Cain. Pada tabel diatas dijelaskan healthy brand yang dominan yakni diamond di cluster 0,1,2,dan 3 kemudian brand V-Soy di cluster 4. Dari hasil tersebut ditemukan peluang repositioning terutama pada cluster 4 di mana V-Soy kalah terhadap penjualan MC Cain dan perlu adanya perkuatan distribusi pada merk V-Soy di cluster 4. Action plan yang disarankan dari hasil clustering diatas yakni dengan mempertahankan dominasi Diamond dengan inovasi design packaging di seluruh cluster kecuali cluster 4 dan memperkuat pesan kesehatan di cluster 4 karena V-Soy masih kalah secara penjualan dengan brand MC Cain.

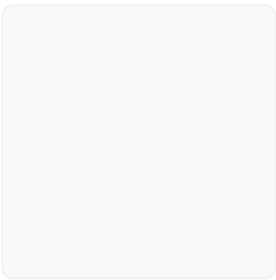
1. Berikut visualisasi persebaran data berdasarkan clusternya



Gambar 9

silhouette score 2023

Clustering pada tahun 2023 memiliki silhouette score = 0,95 yang sangat mendekati 1 yang berarti proses clustering sangat baik dan valid.



Gambar 10

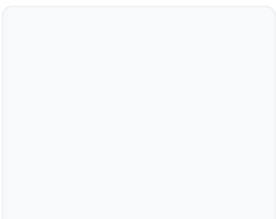
Decision Tree

Hasil pohon keputusan menunjukkan bahwa potensi keuangan menjadi faktor utama dalam menentukan tingkat potensi suatu cluster. Cluster yang memiliki potensi keuangan di atas 50% dikategorikan sebagai kelompok dengan kontribusi keuntungan tinggi. Selanjutnya, billing quantity digunakan untuk membedakan tingkat volume pembelian pada masing-masing cluster.Cluster 1 memiliki potensi penjualan tertinggi dengan nilai potensi keuangan sebesar 83,19% dan rata-rata jumlah pembelian sebesar 57,292, sehingga dikategorikan sebagai potensi sangat tinggi. Cluster 2 termasuk kategori potensi tinggi karena memiliki keuntungan yang cukup besar meskipun jumlah pembeliannya lebih rendah.Sementara itu, Cluster 3 dikategorikan sebagai potensi sedang karena memiliki kontribusi keuntungan yang cukup baik namun masih memerlukan peningkatan strategi pemasaran. Cluster 4 termasuk potensi rendah karena memiliki nilai potensi keuangan yang paling kecil, sehingga memerlukan perhatian khusus untuk meningkatkan kontribusi penjualannya

Tabel 6 menjelaskan tentang hasil clustering pada tahun 2024 :

Tabel 6

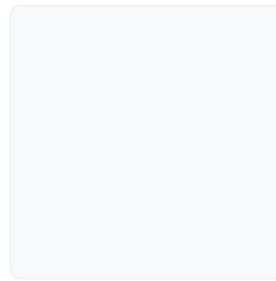
Clustering 2024





Cluster yang memiliki potensi keuntungan terbesar adalah cluster 1 dan 2. Cluster 1 dan 2 tidak punya healthy brand yang dominan sama sekali sehingga perlu adanya penguatan pesan kesehatan. Kemudian, brand Diamond sudah punya posisi dominan di cluster 0 dan 3 sehingga untuk menjaga dominasi, perlu adanya design packaging yang lebih menarik atau varian rasa yang lebih banyak. Selain itu, di cluster 4 perlu adanya penguatan pesan kesehatan dan inovasi pada produk Oatside karena masih kalah dengan MC Cain. Brand Diamond, Oatside, V-Soy. Didapat hasil seperti tabel di atas yaitu brand yang paling diminati pada cluster 0 yakni diamond, cluster 1 pada brand NFH, Cluster 2 pada brand Nutella, cluster 3 dengan brand diamond dan cluster 4 pada Brand MC Cain. Healthy brand yang dominan berdasarkan tabel diatas adalah diamond pada cluster 0 dan 3 kemudian brand Oatside pada Cluster 4. Dari hasil clusterisadi diatas ditemukan adanya peluang repositioning terutama pada cluster 4 dimana oatside kalah penjualan terhadap MC Cain, dan perlu adanya perkuatan distribusi pada merk oatside di cluster 4. Maka dari itu, action plan yang disarankan yakni dengan mempertahankan dominasi diamond dengan inovasi design packaging di cluster 0 dan 3, memperkuat pesan kesehatan di cluster 1 dan 3 karena cluter tersebut mempunyai potensi keuntungan yang besar, dan meningkatkan inovasi oatside pada cluster 4 agar dapat menyaingi MC Cain dalam dominasi.

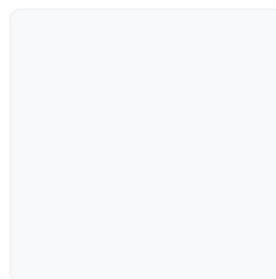
Berikut visualisasi persebaran data berdasarkan clusternya :



Gambar 11

Silhouette score 2024

Dari gambar 7 menjelaskan bahwa Clustering pada tahun 2024 memiliki silhouette score = 0,96 yang sangat mendekati 1 yang berarti proses clustering sangat baik dan valid.



Gambar 12

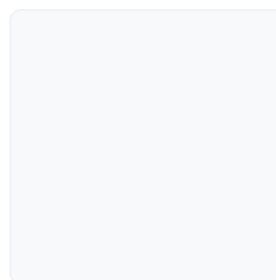
Decision Tree

Hasil pohon keputusan menentukan klasifikasi customer adalah Potensi Keuangan, dengan Billing Quantity Quantity tinggi dikategorikan sebagai Potensi Sangat Tinggi (Cluster 1) karena memiliki kontribusi keuntungan dan volume pembelian terbesar.sebagai faktor pendukung dalam membedakan tingkat kontribusi. Customer dengan Potensi Keuangan >100% dan Billing. Customer dengan potensi keuangan tinggi namun jumlah pembelian lebih rendah masuk ke Cluster 3 (Potensi Tinggi), sedangkan customer dengan kontribusi positif namun lebih rendah

dikategorikan sebagai Cluster 4 (Potensi Sedang). Selain itu, customer yang memiliki Potensi Keuangan $\leq 100\%$ dibedakan menjadi Cluster 2 (Potensi Rendah) dan Cluster 0 (Potensi Sangat Rendah) yang memerlukan strategi peningkatan kontribusi atau evaluasi tingkat lanjutan. Secara garis besar, Cluster 1 jadi segmen paling krusial guna dipertahankan dan terus dilakukan pengembangan, sedangkan Cluster 0 membutuhkan perhatian khusus karena mempunyai risiko kontribusi rendah. Hasil kajian secara keseluruhan mengemukakan bahwa proses clustering memanfaatkan algoritma K-Means Clustering dalam menciptakan nilai silhouette score sebesar 0,94-0,96 pada periode 2020-2024. Nilai ini menunjukkan bahwa kualitas cluster tergolong sangat baik, karena mendekati nilai maksimum yakni 1, yang berarti setiap data memiliki tingkat kesesuaian tinggi dengan cluster nya dan memiliki jarak yang jelas terhadap cluster lain. Secara teoritis, nilai silhouette score yang tinggi mencerminkan keberhasilan algoritma dan meminimalkan *intra cluster distance* dan memaksimalkan *inter cluster distance* [3]. Analisis SWOT digunakan sebagai alat strategis untuk mengidentifikasi kondisi internal dan eksternal yang memengaruhi kinerja healthy brand PT. Sukanda Djaya selama periode 2020–2024. Integrasi SWOT ini berfungsi untuk memperkuat hasil segmentasi pelanggan berbasis K-Means Clustering dan pendekatan RFM, yang sebelumnya telah mengungkap pola dominasi brand, potensi keuntungan tiap klaster, serta posisi kompetitif antar-merek dan selanjutnya dilakukan analisis lingkungan internal-eksternal yang merupakan dasar bagi penyusunan strategi brand yang efektif dan diferensiatif [17], [18]. Selanjutnya, ditambahkan pemahaman mengenai peluang dan ancaman pasar memungkinkan perusahaan menentukan keunggulan bersaing yang berkelanjutan (*sustainable competitive advantage*) [19]. Integrasi model RFM (*Recency, Frequency, Monetary*) memberikan kontribusi yang signifikan dalam mengidentifikasi nilai pelanggan berdasarkan nilai transaksi. Cluster pelanggan dengan nilai billing value dan billing quantity yang tinggi menunjukkan karakteristik sebagai pelanggan bernilai tinggi (*high value-customer*) yang mempunyai dampak kontribusi yang tinggi sehingga dapat digunakan untuk mengidentifikasi pelanggan dengan nilai ekonomi tinggi serta menjadi dasar dalam pengambilan keputusan berdasarkan data [13]. Di sisi lain terdapat cluster pelanggan dengan nilai *billing value* dan *billing quantity* yang rendah yang disebabkan karena proses return, hal ini menunjukkan bahwa tidak semua customer memberikan kontribusi yang positif terhadap perusahaan. Maka diperlukan pendekatan segmentasi pelanggan untuk mengoptimalkan alokasi sumber daya perusahaan dengan menggunakan *customer relationship management* [12]. Fenomena yang terjadi pada penelitian ini yakni brand diamond mendominasi secara konsisten di semua cluster yang dimana menunjukkan bahwa brand diamond memiliki kekuatan yang signifikan dalam hal loyalitas pelanggan, distribusi, dan kualitas produk yang dimana mencerminkan kemampuan yang kuat untuk membangun loyalitas pelanggan serta mempertahankan posisi di pasar yang cukup kompetitif [17], [18].

Tabel 7

Analisis SWOT.



Tabel diatas adalah analisis SWOT yang berpedoman pada hasil clustering tahun 2020–2024, brand Diamond tercatat sebagai healthy brand yang secara berkelanjutan mendominasi sebagian besar klaster pelanggan *food service*. Selain itu, *healthy brand* lain seperti Oatside dan V-Soy mengindikasikan potensi tertentu, namun masih mengalami tekanan secara kompetitif dari merek lain seperti MC Cain, NN, dan USA, utamanya pada klaster berpotensi keuntungan tinggi. Pola clustering dan Analisis swot diatas menunjukan tingkat konsistensi yang tinggi dan perilaku pasar yang relatif stabil meskipun terjadi perubahan kondisi pasar pada masa pandemi covid-19. Konsistensi pola dalam clustering menunjukkan bahwa struktur data memiliki stabilitas yang tinggi, sehingga hasil analisis dapat dipercaya dan digunakan sebagai dasar dalam pengambilan keputusan [20], [21]. Dari hasil penjelasan mengenai pola konsistensi, maka untuk mengintegrasikan strategy brand yang dapat diterapkan dengan perhitungan ifas dan efas lalu dibentuk matriks kuadran swot dibawah ini :

Tabel 8

---	--



					<u>3</u>	<u>0</u>						
					<u>2,5</u>							
					<u>2</u>							
					<u>1,5</u>							
					<u>1</u>							
					<u>0.5</u>							
W						0						S
<u>-3</u>	<u>-2,5</u>	<u>-2</u>	<u>-1,5</u>	<u>-1</u>	<u>-0,5</u>	<u>0,5</u>	<u>1</u>	<u>1,5</u>	<u>2</u>	<u>2,5</u>	<u>3</u>	
					<u>-0,5</u>							
					<u>-1</u>							
					<u>-1,5</u>							

-					-2						-
-					-2,5						-
-	-	-	-	-	-3	T	-	-	-	-	-

Gambar 13

Matriks kuadran SWOT PT.Sukanda Djaya

Berdasarkan Gambar 13, hasil matriks SWOT menunjukkan bahwa perusahaan berada pada Kuadran I (*growth strategy*), yang mengindikasikan bahwa perusahaan memiliki kekuatan internal yang tinggi serta peluang eksternal yang besar. Hasil ini tidak terlepas dari analisis K-Means Clustering berbasis RFM yang menunjukkan adanya segmentasi pelanggan dengan nilai profit tinggi serta tingkat loyalitas yang baik. Posisi ini sangat strategis karena memungkinkan perusahaan untuk memiliki kekuatan internal yang tinggi serta peluang pasar yang besar, posisi ini sangat strategi untuk perusahaan melakukan ekspansi pasar dan inovasi produk sehingga dapat mencapai keunggulan kompetitif yang berkelanjutan [19]. Dengan demikian, hasil clustering menjadi dasar dalam penentuan faktor internal dan eksternal, serta memperkuat posisi strategis perusahaan dalam matriks SWOT. Oleh karena itu, strategi yang dihasilkan berfokus pada pemanfaatan kekuatan perusahaan untuk mengoptimalkan potensi segmen pelanggan bernilai tinggi serta memperluas pasar pada segmen yang belum maksimal.

IV. Kesimpulan

Penelitian ini bertujuan untuk menganalisis segmentasi pelanggan dan potensi penjualan *healthy brand* PT. Sukanda Djaya dengan menggunakan pendekatan K-Means Clustering berbasis model RFM (*Recency, Frequency, Monetary*) serta analisis SWOT sebagai dasar penyusunan strategi bisnis. Berdasarkan hasil penelitian periode 2020–2024, metode K-Means mampu mengelompokkan karakteristik pelanggan berdasarkan pola transaksi, nilai pembelian, dan kontribusi keuntungan sehingga menghasilkan segmentasi customer yang memiliki tingkat potensi penjualan berbeda. Hasil evaluasi menunjukkan nilai silhouette score sebesar 0,94–0,96, yang menunjukkan bahwa hasil clustering memiliki tingkat validitas dan kualitas pengelompokan yang sangat baik. Hasil analisa yang telah dilakukan menggambarkan bahwa Potensi Keuangan adalah faktor pokok untuk merepresentasikan tingkat potensi penjualan pelanggan, sedangkan Billing Quantity menjadi faktor pendukung untuk membedakan tingkat kontribusi pembelian. Berpedoman pada hasil *Decision Tree* , pelanggan dengan nilai potensi keuangan tinggi serta volume pembelian besar masuk kedalam kategori *high value customer* yang memberi kontribusi besar terhadap sebuah entitas. Tahun 2024, Cluster 1 masuk sebagai segmen paling berpotensi dengan nilai prospektif keuangan tertinggi senilai 520,81%, sedangkan Cluster 2 juga mengemukakan peluang besar dengan potensi keuangan senilai 107,58%. Kondisi ini mengemukakan bahwa kedua cluster menjadi sasaran utama pada strategi mempertahankan konsumen dan memberi peningkatan dan pertumbuhan penjualan. Kemudian, hasil kajian mengemukakan bahwa brand Diamond mempunyai posisi yang paling kuat karena secara keberlanjutan mendominasi sebagian besar cluster pada periode kajian, yang memberi indikasi terciptanya kekuatan distribusi, loyalitas pelanggan, dan penerimaan pasar yang sangat baik. Namun demikian, *healthy brand* lain seperti Oatside dan V-Soy masih mengalami tantangan kompetitif di beberapa segmen yang mendorong pentingnya dilakukan strategi repositioning, peningkatan distribusi, dan penguatan komunikasi pada manfaat kesehatan produk. Berpedoman hasil integrasi clustering menggunakan analisa SWOT, PT. Sukanda Djaya berada pada Kuadran I (*growth strategy*) yang menggambarkan bahwa entitas mempunyai kekuatan internal dan peluang eksternal yang luas guna dilakukan pengembangan pasar. Maka dari itu, strategi yang paling disarankan yakni mempertahankan pelanggan dengan nilai ekonomi tinggi, meningkatkan loyalitas pelanggan potensial, memperkuat inovasi produk *healthy brand*, serta dilakukan optimalisasi pemasaran yang berpedoman pada karakteristik masing-masing cluster. Kajian ini memberi kontribusi untuk pengambilan keputusan berbasis data (*data-driven decision making*), dimana hasil segmentasi konsumen bukan sekedar memahami kelompok pelanggan, namun menjadi pedoman guna memprediksi menentukan prioritas pemasaran, potensi penjualan, dan menciptakan strategi pengembangan *healthy brand* PT. Sukanda Djaya secara lebih efektif dan berkelanjutan. Dibandingkan dengan penelitian terdahulu, kajian ini menciptakan kontribusi yang lebih komprehensif dengan integrasi pendekatan analitik (clustering berbasis data mining) dan pendekatan strategis (SWOT). Jika penelitian sebelumnya cenderung hanya berfokus pada pengelompokan data atau analisis strategi secara terpisah, maka penelitian ini menunjukkan bahwa kombinasi keduanya mampu menghasilkan *insight* yang lebih mendalam dan implementatif. Hal ini menjadi nilai tambah utama dalam mendukung pengambilan keputusan yang lebih akurat dan berbasis data.



REFERENSI

- [1] O. Maimon dan L. Rokach, *Data mining and knowledge discovery handbook* , 2 ed. Springer, 2010.
- [2] D. T. Larose, *Discovering knowledge in data: An introduction to data mining* . Wiley, 2005.
- [3] R. Xu dan D. Wunsch, *Clustering* . Wiley-IEEE Press, 2009.
- [4] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* , 1967, hal. 281–297.
- [5] L. Kaufman dan P. J. Rousseeuw, *Finding groups in data: An introduction to cluster analysis* . Wiley, 1990.
- [6] A. Agusta, *Data mining dan warehousing* . Informatika, 2007.
- [7] F. Nasari dan S. Darma, "Penerapan K-Means clustering pada data penerimaan mahasiswa baru (Studi kasus: Universitas Potensi Utama)," *Semnasteknomedia Online* , vol. 3, no. 1, hal. 73, 2013.
- [8] A. Sulistiyawati dan E. Supriyanto, "Implementasi algoritma dan K-Means clustering dalam penentuan siswa kelas unggulan," *J. Tekno Kompak* , vol. 15, no. 2, hal. 25, Agu 2021, doi: 10.33365/jtk.v15i2.1162.
- [9] E. Muningsih dan S. Kiswati, "Penerapan metode K-Means untuk clustering produk online shop dalam penentuan stok barang," *J. Bianglala Inform.* , vol. 3, no. 1, 2015.
- [10] A. P. Windarto, "Penerapan data mining pada ekspor buah-buahan menurut negara tujuan menggunakan K-Means clustering," *Techno.COM* , vol. 16, no. 4, hal. 348–357, 2017.
- [11] S. Hajar dan A. A. Novani, "Penerapan K-Means clustering pada ekspor minyak kelapa sawit menurut negara tujuan," in *Seminar Nasional* , Feb 2020, hal. 314–318.
- [12] F. Buttle, *Customer relationship management: Concepts and technologies* , 3 ed. Routledge, 2014.
- [13] R. Venkatesan dan V. Kumar, "A customer lifetime value framework for customer selection and resource allocation strategy," *J. Mark.* , vol. 68, no. 4, hal. 106–125, 2004.
- [14] C.-W. Chung dan H.-L. Chen, "A modified RFM model for customer segmentation," *Expert Syst. Appl.* , vol. 39, no. 10, hal. 11550–11556, 2016.
- [15] D. Oktarian, R. Firmansyah, dan A. Rachman, "Implementasi algoritma K-Means untuk segmentasi data konsumen dalam perancangan strategi pemasaran," *J. Teknol. dan Sist. Inf.* , vol. 1, no. 1, hal. 24–31, 2020.
- [16] M. Fakhri, S. Nurhaeni, dan A. Prasetyo, "Penerapan metode clustering dalam analisis data konsumen menggunakan algoritma K-Means," *J. Sist. Inf.* , vol. 17, no. 2, hal. 45–52, 2021.
- [17] K. L. Keller, *Strategic brand management: Building, measuring, and managing brand equity* . Pearson Education, 2003.
- [18] D. A. Aaker, *Building strong brands* . Free Press, 1996.
- [19] M. E. Porter, *Competitive advantage: Creating and sustaining superior performance* . Free Press, 1985.
- [20] M. Christopher, *Logistics & supply chain management* , 5 ed. Pearson Education, 2016.
- [21] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognit. Lett.* , vol. 31, no. 8, hal. 651–666, 2010, doi: 10.1016/j.patrec.2009.09.011.

Conflict of Interest Statement:



The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © Universitas Muhammadiyah Sidoarjo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Copyright © Universitas Muhammadiyah Sidoarjo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.