

Segmentasi Pelanggan Menggunakan Metode K-Means Dalam Data Mining untuk Strategi Promosi UMKM Lily Cakes Pontianak

Edward Revaldo Danuwinata¹, Jimmy Tjen²

¹Program Studi Bisnis Digital, ²Program Studi Informatika,

Universitas Widya Dharma Pontianak, Jl. H.O.S. Cokroaminoto No. 445, Pontianak.

Email: edward.xmm1@gmail.com

ABSTRAK

Segmentasi pelanggan merupakan strategi penting dalam pemasaran untuk meningkatkan daya saing bisnis. Penelitian ini menerapkan metode K-Means untuk mengelompokkan pelanggan Lily Cakes berdasarkan karakteristik pembelian mereka. Data yang digunakan meliputi 679 transaksi selama tiga hari raya utama: Natal, Imlek, dan Idul Fitri. Jumlah klaster optimal ditentukan menggunakan metode *silhouette*. Hasil penelitian menunjukkan adanya empat klaster pelanggan dengan preferensi dan pola pembelian yang berbeda. Klaster 1 cenderung membeli kue premium untuk Lebaran dan Imlek, Klaster 2 berfokus pada Idul Fitri, Imlek, dan Natal, Klaster 3 membeli kue standar dan premium saat Imlek, sedangkan Klaster 4 lebih aktif berbelanja untuk Natal dan Lebaran. Berdasarkan hasil tersebut, disarankan strategi pemasaran yang disesuaikan dengan karakteristik tiap klaster, seperti promosi *pre-order*, potongan harga, serta pemasaran digital melalui Instagram dan WhatsApp. Penelitian ini berkontribusi terhadap penerapan *data mining* dalam konteks UMKM konvensional serta menjadi acuan praktis bagi pelaku usaha dalam mengoptimalkan strategi promosi berbasis data untuk meningkatkan penjualan dan daya saing sesuai dengan perkembangan zaman.

Kata Kunci : *Data Mining, Segmentation, Clustering, Business Intelegence, K-Means*

ABSTRACT

Customer segmentation is an important marketing strategy to increase business competitiveness. This study applies the K-Means method to group Lily Cakes customers based on their purchasing characteristics. The data used includes 679 transactions during three major holidays: Christmas, Chinese New Year, and Eid al-Fitr. The optimal number of clusters was determined using the silhouette method. The results show the existence of four customer clusters with different preferences and purchasing patterns. Cluster 1 tends to purchase premium cakes for Eid al-Fitr and Chinese New Year, Cluster 2 focuses on Eid al-Fitr, Chinese New Year, and Christmas, Cluster 3 purchases standard and premium cakes during Chinese New Year, while Cluster 4 is more active in shopping for Christmas and Eid. Based on these results, marketing strategies tailored to the characteristics of each cluster are recommended, such as pre-order promotions, discounts, and digital marketing through Instagram and WhatsApp. This study contributes to the application of data mining in the context of conventional MSMEs and serves as a practical reference for business actors in optimizing data-based promotional strategies to increase sales and competitiveness in accordance with current developments.

Keywords : *Data Mining, Segmentation, Clustering, Business Intelegence, K-Means*

1. PENDAHULUAN

Dalam menghadapi persaingan bisnis yang semakin kompetitif, Usaha Mikro, Kecil, dan Menengah (UMKM) dituntut untuk menerapkan strategi pemasaran yang efektif guna mempertahankan eksistensi sekaligus meningkatkan daya saing. Salah satu strategi yang dinilai efektif dalam memahami kebutuhan dan perilaku konsumen adalah segmentasi pasar, yaitu proses mengelompokkan konsumen ke dalam segmen-semen spesifik berdasarkan karakteristik yang serupa. Dengan pengelompokan pelanggan berdasarkan atribut seperti usia, jenis kelamin, dan pola pembelian, bisnis dapat merancang strategi pemasaran untuk memenuhi kebutuhan dan preferensi pelanggan (Mohsin, 2021; Mohassel & Trieu, 2020). Sebuah bisnis dapat menemukan peluang pasar, meningkatkan penjualan, membangun loyalitas pelanggan, dan mengalokasikan sumber daya secara efektif untuk meningkatkan profitabilitas melalui segmentasi yang tepat. Segmentasi pelanggan dapat membantu memprediksi hubungan pelanggan-penjual, mengevaluasi kepuasan pelanggan, dan mengoptimalkan layanan dengan sumber daya yang lebih sedikit, sehingga mendukung keberlanjutan pertumbuhan bisnis (Saxena et al., 2024).

Untuk mengimplementasikan segmentasi pelanggan secara lebih objektif dan berbasis data, berbagai metode analisis dapat digunakan. Salah satu metode yang paling populer di kalangan peneliti dan praktisi bisnis adalah metode K-Means (Ahmed et al., 2020; Mohassel et al., 2020; Cui, 2020). Pengelompokan dengan K-Means adalah metode pembelajaran tanpa pengawasan yang mengklasterisasi data berdasarkan kemiripan karakteristik (Zubair, 2024; Ghazar, 2021; Dalmaijer et al., 2022). Penentuan jumlah klaster yang ideal dilakukan terlebih dahulu untuk menentukan jumlah pembagian kelompok, Kemudian titik awal dipilih secara acak (Ikotun et al., 2021; Raeisin & Sesay, 2022). Setiap data akan dialokasikan ke klaster terdekat menurut jarak *Euclidean* atau metrik serupa lainnya. Dengan memahami perbedaan antar segmen, bisnis akan mampu menyusun kampanye pemasaran yang lebih terarah, menyesuaikan strategi komunikasi yang sesuai, dan menawarkan produk atau layanan yang relevan dengan kebutuhan khusus pelanggan di setiap segmen.

Berbagai studi telah memanfaatkan K-Means dalam segmentasi pasar di berbagai industri. Baiq et al. (2024) mengkategorikan pelanggan toko *Fast-Moving Consumer Goods* (FMCG) ke dalam tiga segmen rendah, sedang, dan tinggi untuk meningkatkan

layanan pelanggan dan strategi pemasaran. Al Qahtani & Sankar (2024) menggunakan K-Means untuk menganalisis klaster produktivitas di industri aluminium Bahrain, menghasilkan rencana strategis yang mengurangi biaya dan meningkatkan efisiensi rantai pasokan. Tabianan et al. (2024) menemukan bahwa kombinasi metode SAPK dan K-Means lebih akurat dalam mengelompokkan pelanggan *e-commerce* dibandingkan dengan K-Means murni. Omol et al. (2024) membagi pelanggan toko kelontong di Kenya menjadi empat segmen dan mengidentifikasi peluang peningkatan pengeluaran di segmen berpendapatan tinggi. Selain itu, Ong et al. (2024) menggunakan analisis conjoint dan K-Means untuk menganalisis preferensi pelanggan Samgyeopsal di Filipina, menemukan bahwa daging dan keju merupakan atribut yang paling penting.

Berdasarkan data transaksi Lily Cakes sebanyak 679 catatan pembelian yang seluruhnya terjadi selama tiga hari raya besar di Indonesia, yaitu Natal, Imlek, dan Idul Fitri, menunjukkan bahwa aktivitas penjualan hanya meningkat pada periode musiman, sementara pada hari biasa jarang terjadi transaksi. Kondisi ini menunjukkan bahwa Lily Cakes menghadapi tantangan dalam mempertahankan pelanggan di luar musim perayaan atau dengan kata lain memiliki potensi tingkat *customer churn* yang tinggi pada periode nonmusiman. Hal ini menegaskan pentingnya penerapan strategi segmentasi pelanggan untuk memahami pola pembelian serta merancang promosi yang lebih tepat sasaran sepanjang tahun.

Sebagai upaya menjawab permasalahan tersebut, penelitian ini menerapkan metode K-Means untuk mengelompokkan pelanggan Lily's Cake berdasarkan karakteristik tertentu. Hasil penelitian sebelumnya menunjukkan bahwa metode ini dapat meningkatkan efektivitas segmentasi pelanggan dan mendukung pengambilan keputusan strategis. Secara lebih terukur, penelitian ini bertujuan untuk mengidentifikasi kelompok pelanggan dengan karakteristik serupa berdasarkan pola pembelian selama tiga hari raya besar (Natal, Imlek, dan Idul Fitri), dengan tingkat akurasi klasterisasi dievaluasi menggunakan skor *silhouette* $\geq 0,5$ untuk menentukan jumlah klaster yang optimal. Tujuan utama dari studi ini adalah mengidentifikasi segmen pelanggan yang paling relevan, menentukan produk yang sesuai untuk setiap segmen, serta merancang strategi pemasaran yang lebih efisien.

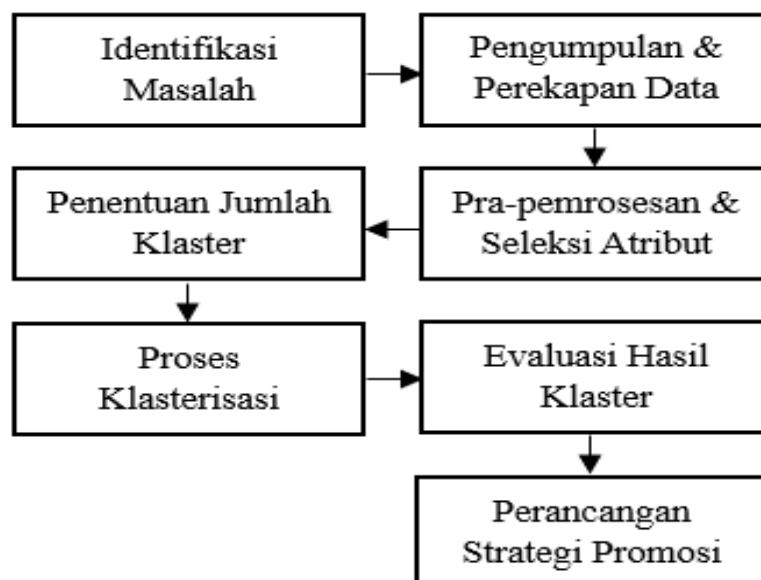
Hasil dari penelitian ini kemudian diimplementasikan dalam perancangan strategi promosi yang disesuaikan dengan karakteristik masing-masing kelompok pelanggan

melalui media sosial Instagram milik Lily Cakes. Sebagai kontribusi ilmiah, penelitian ini mengembangkan studi sebelumnya yang telah menggunakan metode K-Means dalam segmentasi pelanggan Lily Cakes. Dengan adanya studi ini, diharapkan dapat meningkatkan profitabilitas perusahaan dan memberikan wawasan yang lebih mendalam terkait penerapan *data mining* dalam strategi pemasaran untuk usaha mikro, kecil, dan menengah (UMKM) di Indonesia.

Penelitian ini disusun dalam empat bagian utama. Bagian pertama adalah latar belakang yang menjelaskan tujuan, permasalahan, dan solusi yang diusulkan. Bagian kedua menjelaskan metodologi penelitian, termasuk metode, pengolahan data dan penerapan K-Means. Bagian ketiga memaparkan hasil dan diskusi bagaimana hasil pengelompokan dapat memberikan efek terhadap penjualan produk kue Lily cakes, sedangkan bagian terakhir menyajikan kesimpulan dari penelitian ini.

2. METODE

Pada bagian ini akan dibahas secara mendalam mengenai penggunaan algoritma K-Means untuk pengelompokan preferensi pelanggan pada data penjualan Lily Cakes. Seluruh tahapan, metode dan proses yang digunakan dalam penelitian ini akan dijelaskan secara sistematis pada bagian ini. Secara ringkas, Tahapan penelitian dapat dilihat pada gambar 1.



Gambar 1: Tahapan penelitian menggunakan metode K-Means

Identifikasi Masalah

Seiring dengan perkembangan teknologi dan penggunaan media sosial yang semakin meluas, persaingan bisnis semakin meningkat, terutama bagi usaha yang tidak memanfaatkan platform digital sebagai saluran pemasaran. Lily Cakes menghadapi tantangan besar karena sistem penjualan tradisional yang mengandalkan komunikasi melalui WhatsApp dan telepon serta pencatatan transaksi manual. Masalah ini mengakibatkan pengelolaan data yang belum ideal, padahal data transaksi berpotensi memberikan wawasan penting dalam mengelompokkan preferensi pelanggan. Oleh karena itu, Penelitian ini bertujuan untuk melakukan segmentasi pelanggan dan merancang strategi pemasaran berbasis digital yang lebih efektif. Seluruh data manual kemudian dikonversi ke format spreadsheet untuk memudahkan pengolahan lebih lanjut menggunakan software analisis data.

Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan dengan pendekatan dokumentasi, yaitu dengan menghimpun catatan transaksi manual milik pemilik usaha Lily Cakes. Data yang diperoleh mencakup total 679 entri transaksi yang terjadi selama tiga momentum hari raya besar dalam satu tahun terakhir, yakni Natal, Tahun Baru Imlek, dan Idul Fitri. Ketiga momentum tersebut dipilih secara purposif karena menunjukkan lonjakan penjualan yang signifikan dibandingkan bulan-bulan biasa, sehingga dianggap representatif dalam mencerminkan preferensi utama pelanggan.

Setiap data transaksi mencakup sejumlah variabel seperti nama pelanggan, produk yang dibeli, jumlah pembelian, total harga, serta keterangan tambahan seperti lokasi pengiriman dan tujuan pembelian (untuk konsumsi sendiri atau sebagai hadiah). Meskipun sumber data bersifat manual, validitas data dijaga dengan cara melakukan pencocokan silang antara buku transaksi dan nota pesanan yang tersimpan. Tantangan yang ditemukan dalam tahap ini adalah ketidakkonsistenan dalam penamaan produk serta kekosongan pada beberapa kolom data. Untuk mengatasi hal ini, dilakukan proses *data cleaning* seperti normalisasi penamaan produk dan pengisian data kosong berdasarkan informasi relevan dari transaksi serupa.

Rekapitulasi Data

Proses pengolahan awal dimulai dengan mengubah data transaksi dari catatan manual ke dalam format spreadsheet untuk mempermudah analisis. Data kemudian direkap dalam bentuk tabel terstruktur yang memuat 12 atribut, seperti nama pelanggan, jenis kelamin, pekerjaan, domisili, produk yang dibeli, kategori produk, tipe produk, harga, kuantitas, hari raya, total pembelian, dan kelas produk. Penyusunan ini bertujuan untuk menyajikan data secara rapi dan konsisten agar mendukung proses analitik berbasis algoritma. Langkah berikutnya adalah seleksi atribut untuk menentukan variabel yang paling relevan terhadap segmentasi pelanggan. Dari proses ini, enam atribut utama yaitu jenis kelamin, pekerjaan, domisili, hari raya dan kelas produk dikodekan ke dalam format numerik agar sesuai dengan kebutuhan algoritma *K-Means*. Data yang telah dikodekan ini disimpan dalam sheet baru bernama “Konversi”, yang berfungsi sebagai basis data utama dalam proses klasterisasi. Sebelum klasterisasi, semua variabel dinormalisasi menggunakan metode *min-max scaling* sehingga berada pada rentang 0–1. Normalisasi ini memastikan setiap variabel memiliki kontribusi setara dalam perhitungan jarak Euclidean, sehingga variabel dengan skala lebih besar tidak mendominasi hasil klasterisasi.

Klasterisasi dengan K-Means

Setelah seluruh data selesai direkap, proses klasterisasi akan dilakukan menggunakan algoritma K-Means. Tahap pertama dalam metode K-Means adalah menentukan jumlah klaster yang paling sesuai. Penentuan ini dilakukan dengan metode *silhouette*, yang mengevaluasi *cohesion* (seberapa dekat data dengan pusat klasternya sendiri) dan *separation* (seberapa jauh data dari klaster lain). *Nilai silhouette* berkisar antara -1 hingga 1, dengan nilai mendekati 1 menunjukkan kualitas klaster yang lebih baik.

Secara matematis, nilai silhouette untuk data ke-*i* dapat dinyatakan sebagai:

$$s(i) = \begin{cases} \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} & \text{jika } |C_i| > 1 \\ 0 & \text{jika } |C_i| = 1 \end{cases} \quad (1)$$

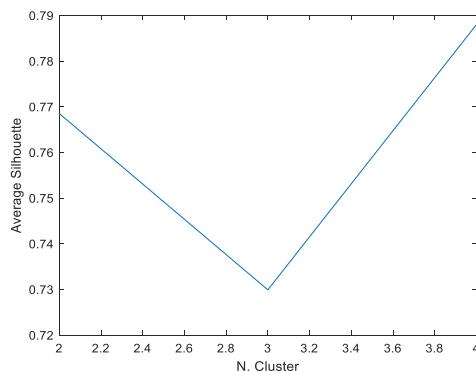
Dengan: *a(i)* merepresentasikan kohesi dari data ke-*i* atau rata-rata jarak dari data ke-*i* terhadap titik lain pada kelompok klaster yang sama, di mana:

$$a(i) = \frac{1}{|C_i| - 1} \sum_{\substack{j \in C_i \\ i \neq j}} d(i, j) \quad (2)$$

$b(i)$ merepresentasikan pemisahan data (i) dari data dalam klaster lainnya.

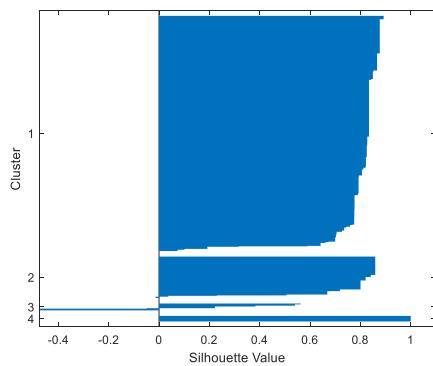
$$b(i) = \min_{i \neq j} \frac{1}{|C_j|} \sum_{j \in C_j} d(i, j) \quad (3)$$

Sedangkan $|C_j|$ merepresentasikan kardinalitas klaster ke-j yaitu jumlah elemen yang terdapat di dalamnya. Jika $|C_i| = 1$, maka nilai silhouette $s(i)$ ditetapkan sebesar 0. Hal ini disebabkan klaster hanya terdiri atas satu data ke-i, sehingga tidak ada titik data lain dalam klaster yang dapat digunakan untuk membandingkan jarak. Oleh karena itu, nilai siluet tidak dapat dihitung menggunakan rumus standar.

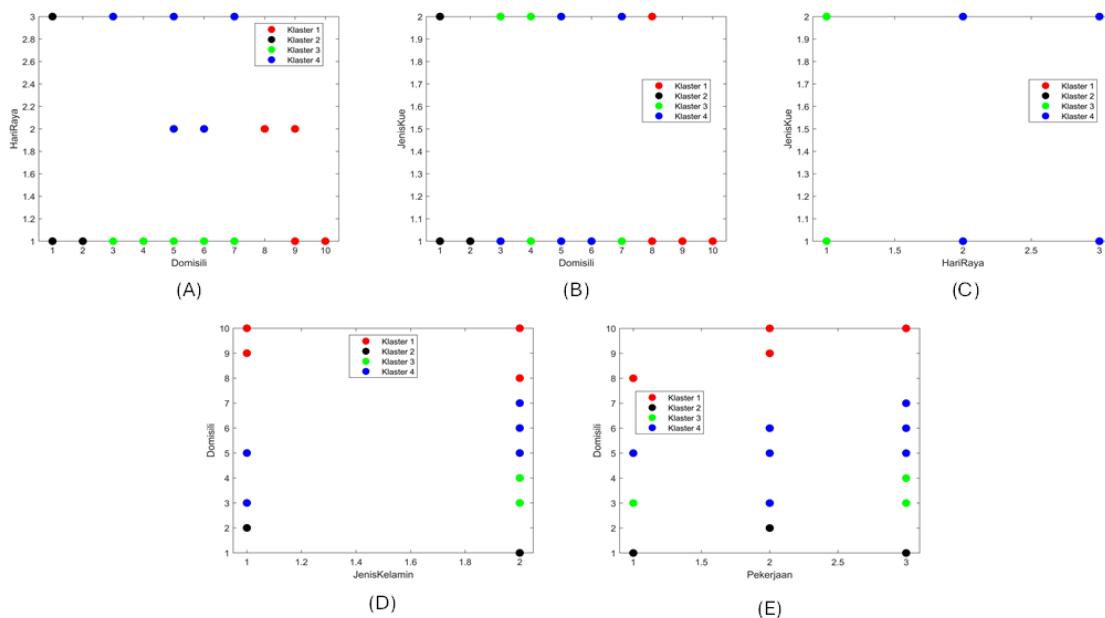


Gambar 2: Rata-Rata Silhouette

Penulis mencoba menggunakan 1 hingga 4 klaster untuk menentukan jumlah klaster yang ideal. Seperti ditunjukkan pada Gambar 2, dari empat jumlah klaster yang diuji, terdapat dua pilihan yang memungkinkan: 2 klaster dan 4 klaster. Secara keseluruhan, kedua pilihan ini dapat menghasilkan klasterisasi yang optimal, namun berdasarkan perhitungan rerata silhouette, 4 klaster merupakan hasil terbaik dengan nilai tertinggi sebesar 0,7880. Hasil distribusi silhouette untuk setiap klaster kemudian ditampilkan pada Gambar 3, di mana semua klaster umumnya memiliki nilai positif, dan Klaster 1 hingga 4 memiliki nilai silhouette yang tinggi atau mendekati 1. Meskipun terdapat beberapa nilai negatif pada Klaster 3, yang menunjukkan bahwa beberapa objek mungkin kurang tepat berada dalam klaster tersebut.

**Gambar 3:** Rata-Rata Silhouette

Proses klasterisasi dilakukan menggunakan algoritma K-Means dengan maksimal iterasi 100 dan toleransi perubahan centroid sebesar 10^{-4} . Algoritma akan berhenti apabila perubahan centroid lebih kecil dari toleransi atau jumlah iterasi maksimum tercapai. Seluruh variabel yang telah dinormalisasi dianggap memberikan kontribusi setara dalam perhitungan jarak Euclidean, sehingga tidak diperlukan pemberian bobot tambahan pada atribut tertentu. Berdasarkan perhitungan tersebut, dapat disimpulkan bahwa penggunaan 4 klaster memungkinkan pengelompokan yang lebih tepat dan terperinci. Scatter plot hasil klasterisasi ke dalam 4 kelompok ditampilkan pada Gambar 4.

**Gambar 4:** Scatter Plot Domisili terhadap Hari Raya (A), Scatter Plot Domisili terhadap Jenis Kue (B), Scatter Plot Hari Raya terhadap Jenis Kue (C), Scatter Plot Jenis Kelamin terhadap Domisili (D), Scatter Plot Pekerjaan terhadap Domisili (E)

3. HASIL DAN PEMBAHASAN

Setelah melakukan klasterisasi menggunakan K-Means, hasil dari pengelompokan dapat dilihat dari melalui *Scatter Plot* pada gambar 4, kemudian informasi dikelompokkan kedalam tabel sesuai dengan klasternya masing-masing. Informasi mengenai scatter plot dapat dilihat pada tabel 1

Tabel 1. Informasi pada *Scatter Plot*

Atribut	No	Keterangan
Jenis	1	Laki-Laki
Kelamin	2	Perempuan
Pekerjaan	1	Ibu Rumah Tangga
	2	Karyawan
	3	Wiraswasta
Domisili	1	Balai Bekuak
	2	Banjarmasin
	3	Jakarta
	4	Ketapang
	5	Pontianak
	6	Sampit
	7	Sepauk
	8	Singkawang
	9	Sintang
	10	Tayan
Hari Raya	1	Imlek
	2	Lebaran
	3	Natal
Jenis Kue	1	Premium
	2	Standar

Klaster 1 (Merah)

Tabel 2. Informasi pada *Scatter Plot* Klaster 1

No.	Karakteristik
1.	Pelanggan berjenis kelamin laki-laki berada di domisili Sintang, dan Tayan.
2.	Pelanggan berjenis kelamin perempuan berada di domisili Singkawang, dan Tayan.
3.	Pelanggan berdomisili di Singkawang dengan pekerjaan Ibu Rumah Tangga yang dominan melakukan pembelian kue untuk hari raya Idul Fitri dengan preferensi jenis kue standar dan premium.
4.	Pelanggan berdomisili di Sintang dengan pekerjaan Karyawan yang dominan melakukan pembelian kue untuk hari raya Idul Fitri dan Imlek dengan preferensi jenis kue premium.
5.	Pelanggan berdomisili di Tayan dengan pekerjaan Karyawan dan Wiraswasta yang dominan melakukan pembelian kue untuk hari raya Imlek dengan preferensi jenis kue premium.

Berdasarkan tabel 2, dapat dianalisa bahwa klaster ini terdiri dari pelanggan yang yang berdomisili di kota Singkawang, Sintang, dan Tayan. Dimana pelanggan laki-laki dominan di domisili Sintang dan Tayan, sedangkan perempuan berada di domisili Singkawang dan Tayan. Pekerjaan pelanggan dapat diketahui bahwa pelanggan yang berdomisili di Singkawang kebanyakan adalah ibu rumah tangga, domisili Sintang adalah karyawan, sedangkan domisili tayan adalah wiraswasta dan karyawan. Pelanggan dari domisili Singkawang kebanyakan membeli kue untuk Idul-fitri, pelanggan domisili Sintang kebanyakan membeli kue untuk Idul Fitri dan imlek, dan Pelanggan di Tayan kebanyakan membeli untuk hari raya Imlek. Pelanggan singkawang kebanyakan membeli kue berjenis premium dan standar sedangkan pelanggan dari Tayan dan Sintang rerata membeli kue kelas premium.

Fenomena ini dapat dijelaskan dari aspek sosial dan budaya masyarakat Kalimantan Barat, yang memiliki tradisi kuat dalam menyambut hari raya dengan berbagai hidangan khas, termasuk kue dan makanan manis sebagai simbol kebersamaan. Selain itu, Idul Fitri juga menjadi momen utama untuk berbagi makanan kepada tetangga dan kerabat, sehingga permintaan terhadap produk kue meningkat signifikan pada periode tersebut. Pola ini menunjukkan bahwa peningkatan penjualan tidak hanya dipengaruhi oleh kebutuhan konsumsi pribadi, tetapi juga oleh nilai sosial yang melekat pada budaya memberi dalam perayaan keagamaan.

Jika ditinjau dari jenis pekerjaan, sebagian besar pelanggan dalam klaster ini adalah ibu rumah tangga dan pegawai, dengan kecenderungan pembelian produk dalam jumlah sedang untuk kebutuhan keluarga. Ibu rumah tangga biasanya lebih memperhatikan variasi rasa dan harga, sedangkan pegawai cenderung memilih produk yang praktis dan siap saji karena keterbatasan waktu. Hal ini memperlihatkan hubungan antara kondisi sosial ekonomi dan preferensi produk, di mana keputusan pembelian dipengaruhi oleh kombinasi antara kebiasaan budaya, waktu perayaan, dan pertimbangan efisiensi. Dengan demikian, Klaster 1 mencerminkan segmen pelanggan yang berorientasi pada tradisi dan kebutuhan rumah tangga, yang menjadi pasar potensial pada periode menjelang Idul Fitri dan imlek.

Strategi promosi yang ideal untuk klaster ini ialah difokuskan pada momentum Idul Fitri dan Imlek dengan mempromosikan kue berjenis premium. Selain itu, benefit untuk pre-order juga dapat diterapkan, di mana pelanggan diberikan kesempatan untuk

memesan lebih awal dengan harga spesial. Strategi lain yang dapat diterapkan adalah layanan pengantaran dengan promo gratis ongkos kirim dengan minimal pembelian tertentu. Selain itu promosi yang disesuaikan untuk tiap kelompok pelanggan juga menjadi faktor penting dalam meningkatkan efektivitas penjualan. Penggunaan WhatsApp dan Instagram untuk pemasaran dapat menjadi pilihan utama, mengingat ibu rumah tangga di Singkawang lebih aktif di platform tersebut. Strategi pemasaran ini dapat memanfaatkan testimoni pelanggan meningkatkan kepercayaan dan jangkauan promosi. Berdasarkan deksripsi diatas, poster promosi yang dapat diciptakan dari karakteristik yang telah disimpulkan, seperti pada gambar 5.



Gambar 5: Media Promosi Klaster 1

Klaster 2 (Hitam)

Tabel 3. Informasi pada *Scatter Plot* Klaster 2

No	Karakteristik
1.	Pelanggan berjenis kelamin laki-laki berada di domisili Banjarmasin.
2.	Pelanggan berjenis kelamin Perempuan berada di domisili Balai Bekuak.
3.	Pelanggan berdomisili di Balai Bekuak dengan pekerjaan Ibu Rumah Tangga dan Wiraswasta yang dominan melakukan pembelian kue untuk Imlek dan Natal dengan preferensi jenis kue standard dan premium.
4.	Pelanggan berdomisili di Banjarmasin dengan pekerjaan Karyawan yang dominan melakukan pembelian kue untuk Imlek dengan preferensi jenis kue premium.

Berdasarkan tabel 3, klaster ini terdiri atas pelanggan yang berasal dari kota Banjarmasin & Balai Bekuak, dimana pelanggan Banjarmasin berjenis kelamin lelaki

dan pelanggan dari Balai Bekuak adalah perempuan. Pelanggan Balai Bekuak memiliki jenis pekerjaan yaitu Ibu rumah tangga dan wiraswasta. Sedangkan pelanggan Banjarmasin adalah Karyawan. Pelanggan domisili kebanyakan membeli kue saat Imlek dan Natal, sedangkan pelanggan Banjarmasin sering membeli kue saat imlek. Pelanggan domisili Balai Bekuak menyukai jenis kue premium maupun standar sedangkan pelanggan Banjarmasin menyukai kue premium.

Pola pembelian pada klaster ini menunjukkan peningkatan menjelang Imlek dan Natal, dua perayaan yang erat kaitannya dengan tradisi saling memberi dan berkumpul bersama keluarga. Pada beberapa daerah di Indonesia terutama Kalimantan Barat, kue sering dijadikan sebagai bentuk penghormatan dan ucapan terima kasih kepada kerabat dan tamu, sedangkan di Banjarmasin, pembelian kue premium lebih banyak dilakukan untuk keperluan perayaan keluarga dan relasi kerja. Fenomena ini menggambarkan bagaimana aspek budaya dan sosial berperan besar dalam membentuk perilaku konsumsi masyarakat di kedua wilayah tersebut.

Jika dilihat dari segi pekerjaan, pelanggan wiraswasta dan ibu rumah tangga di Balai Bekuak cenderung membeli dalam jumlah sedang dengan mempertimbangkan variasi rasa dan tampilan produk. Kue biasanya dibeli untuk disajikan kepada tamu atau dijadikan buah tangan bagi kerabat yang berkunjung. Sementara itu, pelanggan yang bekerja sebagai karyawan di Banjarmasin lebih cenderung memilih produk premium yang praktis dan siap saji, menyesuaikan dengan ritme kehidupan kota yang lebih cepat. Perbedaan ini menunjukkan adanya kaitan antara latar belakang pekerjaan dan preferensi produk. Dengan kata lain, pelanggan dalam Klaster 2 membeli bukan hanya untuk konsumsi pribadi, tetapi juga sebagai bagian dari tradisi berbagi dan menjaga hubungan sosial saat perayaan hari besar.

Strategi promosi yang tepat untuk Kluster 2 (Hitam) difokuskan pada setiap perayaan seperti Imlek, Natal, dan Idul Fitri, dengan menyesar segmen pasar yang terdiri dari produk premium serta standar. Beberapa promosi yang dapat diterapkan mencakup promo *cashback* atau skema *buy 2 get 1*, yang bertujuan untuk meningkatkan fleksibilitas pilihan bagi pelanggan dalam memilih jenis kue. Berdasarkan deksripsi diatas, poster promosi yang dapat diciptakan dari karakteristik yang telah disimpulkan, seperti pada gambar 6.



Gambar 6: Media Promosi Klaster 2

Klaster 3 (Hitam)

Tabel 4. Informasi pada *Scatter Plot* Klaster 3

No	Karakteristik
1.	Pelanggan berjenis kelamin Perempuan berada di domisili Jakarta dan Ketapang
2.	Pelanggan berdomisili di Jakarta dengan pekerjaan Ibu Rumah Tangga dan Wiraswasta yang dominan melakukan pembelian kue untuk Imlek dengan preferensi jenis kue standard
3.	Pelanggan berdomisili di Ketapang dengan pekerjaan Wiraswasta yang dominan melakukan pembelian kue untuk Imlek dengan preferensi jenis kue premium dan standard
4.	Pelanggan berdomisili Pontianak, Sampit, Sepauk yang tidak diketahui jenis kelamin dan pekerjaanya yang dominan melakukan pembelian kue untuk Imlek dengan preferensi jenis kue premium untuk pelanggan Sepauk

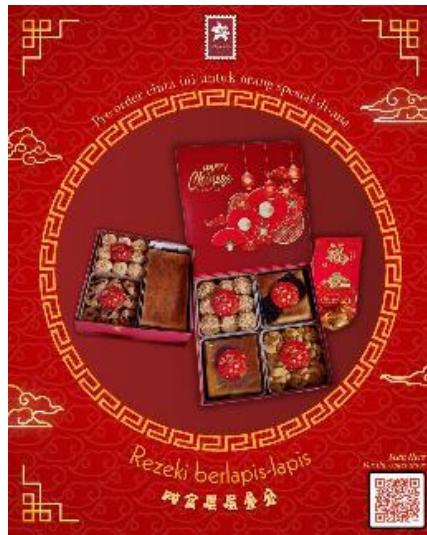
Berdasarkan tabel 4, klaster ini terdiri atas pelanggan berjenis kelamin perempuan berdomisili Jakarta dan Ketapang, dimana ada pelanggan Jakarta merupakan Ibu rumah tangga dan ada juga yang merupakan wiraswasta. Sedangkan pelanggan Ketapang kebanyakan adalah wiraswasta. Lalu juga merupakan klaster pelanggan kelompok Pontianak, Sampit, dan Sepauk yang tidak diketahui jenis kelaminnya. Pelanggan domisili Jakarta menyukai jenis kue standar, sedangkan pelanggan Ketapang menyukai jenis kue standar maupun premium. Dan pelanggan Sepauk menyukai jenis kue premium. Dimana pelanggan klaster ini sering membeli kue untuk imlek berjenis standar dan premium.

Fenomena dominasi pembelian saat imlek ini dapat dijelaskan dari faktor budaya dan geografis. Kalimantan Barat adalah salah satu wilayah dengan jumlah diaspora tionghua terbesar di Indonesia terutama kota Ketapang dan Pontianak. Hal ini membuat perayaan imlek menjadi momen utama peningkatan permintaan produk pangan, terutama kue yang biasa digunakan sebagai hadiah atau bingkisan bagi keluarga atau rekan bisnis. Pola serupa juga terlihat pada pelanggan yang berdomisili di Jakarta. Kemungkinan hal ini dipengaruhi oleh masyarakat asal Pontianak yang menetap di ibu kota, serta gaya hidup perkotaan yang identik dengan tren konsumsi musiman. Pada periode menjelang hari raya, pelanggan di perkotaan cenderung mencari produk dengan kemasan menarik untuk kebutuhan hampers, sehingga meningkatkan penjualan produk premium.

Dari sisi pekerjaan, pelanggan yang berprofesi sebagai wiraswasta umumnya membeli dalam jumlah lebih besar karena kue sering digunakan sebagai bentuk apresiasi atau simbol hubungan bisnis kepada mitra kerja. Sementara itu, ibu rumah tangga lebih banyak melakukan pembelian untuk konsumsi pribadi atau keluarga. Pola ini menunjukkan adanya keterkaitan antara jenis pekerjaan dengan preferensi produk, di mana perbedaan tujuan pembelian dan kemampuan ekonomi turut memengaruhi perilaku konsumen dalam memilih jenis dan jumlah produk yang dibeli.

Strategi promosi untuk Kluster 3 (Hijau) difokuskan pada perayaan Imlek dengan menawarkan kombinasi produk premium dan standar. Mengingat preferensi pelanggan di Jakarta dan Ketapang yang menyukai kedua jenis kue, inovasi dalam varian rasa dapat menjadi strategi yang menarik untuk meningkatkan daya tarik produk. Untuk meningkatkan daya tarik produk di kota besar seperti Jakarta, konsep *pre-order* dengan tema spesial juga bisa diterapkan, misalnya melalui kemasan unik yang sesuai dengan nuansa Imlek. Dalam pendekatan kepada kelompok pelanggan yang terdiri dari wiraswasta dan ibu rumah tangga, pemanfaatan strategi digital menjadi aspek penting. Mengingat pelanggan urban di Jakarta lebih sering berbelanja secara online, promosi melalui *Instagram* dapat menjadi metode yang efektif untuk menjangkau mereka. Selain itu, testimoni dari pelanggan loyal yang telah terbiasa membeli kedua jenis kue dapat dimanfaatkan sebagai konten pemasaran untuk membangun kredibilitas dan meningkatkan kepercayaan calon pelanggan baru. Berdasarkan deksripsi diatas, poster

promosi yang dapat diciptakan dari karakteristik yang telah disimpulkan, seperti pada gambar 7.



Gambar 7: Media Promosi Klaster 3

Klaster 4 (Biru)

Tabel 5. Informasi pada *Scatter Plot* Klaster 4

No	Karakteristik
1.	Pelanggan berdomisili di Pontianak dengan pekerjaan Ibu Rumah Tangga, Karyawan, dan Wiraswasta yang dominan melakukan pembelian kue untuk hari raya Idul Fitri dan Natal dengan preferensi jenis kue premium dan standard
2.	Pelanggan berdomisili di Jakarta, dengan pekerjaan Karyawan yang dominan melakukan pembelian kue untuk hari raya Natal dengan preferensi jenis kue premium
3.	Pelanggan berdomisili di Sepauk, dengan pekerjaan Wiraswasta yang dominan melakukan pembelian kue untuk hari raya Natal dengan preferensi jenis kue standard
4.	Pelanggan berdomisili di Sampit dengan pekerjaan Karyawan dan wiraswasta yang dominan melakukan pembelian kue untuk hari raya Idul Fitri dengan preferensi jenis kue premium
5.	Pelanggan berdomisili di Pontianak dengan pekerjaan Ibu Rumah Tangga, Karyawan, dan Wiraswasta yang dominan melakukan pembelian kue untuk hari raya Idul Fitri dan Natal dengan preferensi jenis kue premium dan standard
6.	Pelanggan berdomisili di Jakarta, dengan pekerjaan Karyawan yang dominan melakukan pembelian kue untuk hari raya Natal dengan preferensi jenis kue premium

Berdasarkan tabel 5, klaster ini terdiri atas pelanggan berdomisili Jakarta yang berjenis kelamin laki-laki, pelanggan berdomisili Sampit, Sepauk berjenis kelamin perempuan, dan pelanggan Pontianak yang berjenis laki-laki dan perempuan pelanggan di berbagai kota memiliki preferensi dan kebiasaan pembelian yang berbeda. Di Pontianak, mayoritas pelanggan adalah ibu rumah tangga, sementara di Jakarta, Pontianak, dan Sampit, banyak pelanggan yang bekerja sebagai karyawan. Adapun pelanggan yang berprofesi sebagai wiraswasta tersebar di Pontianak, Sampit, dan Sepauk. Dari pola pembelian, pelanggan di Jakarta cenderung membeli kue untuk perayaan Natal, sedangkan di Pontianak, permintaan tinggi terjadi saat Idul fitri dan Natal. Pelanggan di Sampit lebih banyak membeli kue untuk Idul fitri, sementara di Sepauk, pembelian lebih dominan saat Natal. Dalam hal preferensi jenis kue, pelanggan di Jakarta dan Sampit lebih menyukai kue premium. Di Pontianak, pelanggan memilih baik kue premium maupun standar, sementara di Sepauk, kue standar lebih diminati. Untuk perayaan Idul fitri dan Natal, pelanggan secara umum menyukai kedua jenis kue tersebut.

Pola pembelian menunjukkan variasi yang dipengaruhi oleh faktor geografis dan tradisi perayaan di masing-masing daerah. Di Jakarta, pembelian kue meningkat menjelang Natal, sejalan dengan gaya hidup perkotaan yang identik dengan budaya memberi hadiah atau hampers pada akhir tahun. Sementara itu, di Pontianak dan Sampit, peningkatan permintaan terjadi pada Idul Fitri dan Natal, dua momen penting bagi masyarakat yang menjunjung nilai kebersamaan keluarga. Adapun di Sepauk, preferensi pembelian lebih dominan pada Natal, dengan kecenderungan memilih kue berjenis standar yang lebih ekonomis dan mudah dibagikan.

Jika ditinjau dari pekerjaan, pelanggan ibu rumah tangga umumnya berasal dari Pontianak dan berperan besar dalam pengambilan keputusan pembelian untuk kebutuhan keluarga, sementara pelanggan karyawan dan wiraswasta tersebar di Jakarta, Sampit, dan Sepauk. Pelanggan karyawan lebih menyukai kue premium karena dianggap lebih praktis dan cocok dijadikan hadiah formal, sedangkan pelanggan ibu rumah tangga dan wiraswasta lebih mempertimbangkan variasi rasa dan harga untuk disajikan kepada tamu. Pola ini menunjukkan adanya keterkaitan antara jenis pekerjaan, domisili, dan preferensi produk. Dengan demikian, Klaster 4 mencerminkan segmen

pelanggan yang berorientasi pada momen keluarga dan perayaan keagamaan, di mana nilai sosial dan estetika produk berperan penting dalam keputusan pembelian.

Dalam strategi promosi yang sesuai, Klaster 4 (bunga) difokuskan pada perayaan Natal dan Idul Fitri dengan kombinasi produk premium dan standar. Salah satu strategi utama adalah menawarkan kue edisi khusus Natal dan Idul Fitri yang dapat disesuaikan dengan dekorasi khas hari raya, sehingga memberikan sentuhan personal bagi pelanggan. Selain itu, promo “Beli untuk Keluarga” dapat menjadi daya tarik utama, terutama bagi mereka yang ingin mengirimkan paket kue dalam jumlah besar kepada sanak saudara di kota lain. Selain itu, unggahan *feed* yang menampilkan paket eksklusif untuk keluarga besar dapat menjadi daya tarik utama bagi pelanggan yang mencari pilihan hadiah berkualitas untuk momen spesial. Berdasarkan deksripsi diatas, poster promosi yang dapat diciptakan dari karakteristik yang telah disimpulkan, seperti pada gambar 8.



Gambar 8: Media Promosi Klaster 4

4. SIMPULAN

Penelitian ini bertujuan untuk melakukan segmentasi pelanggan Lily Cakes menggunakan algoritma K-Means berdasarkan data penjualan pada tiga hari raya besar, yaitu Natal, Imlek, dan Idul Fitri. Hasil klasterisasi menunjukkan bahwa empat klaster pelanggan dengan karakteristik yang berbeda dapat diidentifikasi, dan hasil evaluasi menggunakan nilai silhouette menunjukkan bahwa empat klaster memberikan hasil optimal dengan tingkat pemisahan data yang baik. Klaster 1 (Merah) mencakup

pelanggan di Singkawang, Sintang, dan Tayan yang cenderung membeli kue premium untuk Idul Fitri dan Imlek. Klaster 2 (Hitam) terdiri atas pelanggan dari Banjarmasin dan Balai Bekuak dengan pola pembelian untuk Imlek dan Natal. Klaster 3 (Hijau) meliputi pelanggan di Jakarta, Ketapang, Pontianak, Sampit, dan Sepauk dengan preferensi pada kue standar dan premium untuk Imlek. Sementara itu, Klaster 4 (Biru) terdiri atas pelanggan di Jakarta, Pontianak, Sampit, dan Sepauk yang membeli kue untuk Natal dan Idul fitri. Analisis ini menunjukkan adanya perbedaan preferensi produk dan waktu pembelian antar klaster, sehingga strategi promosi yang disesuaikan untuk setiap segmen seperti pre-order menjelang hari raya, promo gratis ongkir, dan optimalisasi promosi digital melalui WhatsApp serta Instagram dapat meningkatkan efektivitas pemasaran.

Dari sisi operasional, pelaku usaha disarankan untuk mulai mengelola data pelanggan secara lebih terstruktur dengan memanfaatkan perangkat sederhana yang mudah digunakan. Salah satu langkah yang dapat dilakukan adalah membangun sistem input penjualan menggunakan software seperti Microsoft Access atau aplikasi spreadsheet, sehingga data transaksi dapat dicatat secara rutin dan mudah dianalisis. Hal sederhana ini akan membantu pelaku usaha dalam menyesuaikan strategi promosi berdasarkan segmen pelanggan yang telah diidentifikasi.

Keterbatasan penelitian ini terletak pada cakupan data yang masih terbatas pada periode waktu tertentu serta belum mempertimbangkan pengaruh musim dan variasi produk terhadap pola pembelian pelanggan. Untuk penelitian selanjutnya, disarankan agar analisis diperluas dengan periode waktu yang lebih panjang serta mengintegrasikan metode market basket analysis untuk setiap klaster dan musim penjualan, sehingga dapat diketahui kombinasi produk yang sering dibeli bersamaan pada periode tertentu. Dengan demikian, hasil segmentasi tidak hanya menggambarkan perbedaan karakteristik pelanggan, tetapi juga memberikan informasi strategis mengenai perilaku pembelian lintas produk, yang dapat dimanfaatkan pelaku usaha dalam merancang paket promosi dan strategi penjualan yang lebih tepat sasaran.

5. DAFTAR PUSTAKA

- Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. In *Electronics (Switzerland)* (Vol. 9, Issue 8). <https://doi.org/10.3390/electronics9081295>

- Al Qahtani, H., & P. Sankar, J. (2024). The cluster analysis in the aluminium industry with K-means method: an application for Bahrain. *Cogent Business & Management*, 11(1). <https://doi.org/10.1080/23311975.2024.2361475>
- Baiq Nikum Yulisasih, Herman, H., & Sunardi, S. (2024). K-Means clustering method for Customer Segmentation Based On Potential Purchases. *Jurnal ELTIKOM*, 8(1), 83–90. <https://doi.org/10.31961/eltikom.v8i1.1137>
- Cui, M. (2020). Introduction to the k-means clustering algorithm based on the elbow method. <https://doi.org/10.23977/accaf.2020.010102>
- Dalmaijer, E. S., Nord, C. L., & Astle, D. E. (2022). Statistical power for cluster analysis. *BMC Bioinformatics*, 23(1). <https://doi.org/10.1186/s12859-022-04675-1>
- Ghazal, T. M., Hussain, M. Z., Said, R. A., Nadeem, A., Hasan, M. K., Ahmad, M., Khan, M. A., & Naseem, M. T. (2021). Performances of k-means clustering algorithm with different distance metrics. *Intelligent Automation and Soft Computing*, 30(2), 735–742. <https://doi.org/10.32604/iasc.2021.019067>
- Ikotun, A. M., Almutari, M. S., & Ezugwu, A. E. (2021). K-means-based nature-inspired metaheuristic algorithms for automatic data clustering problems: Recent advances and future directions. In *Applied Sciences (Switzerland)* (Vol. 11, Issue 23). MDPI. <https://doi.org/10.3390/app112311246>
- Mohassel, P., Rosulek, M., & Trieu, N. (2020). Practical privacy-preserving k-means clustering. *Proceedings on Privacy Enhancing Technologies*, 2020(4), 414–433. <https://doi.org/10.2478/popets-2020-0080>
- Mohsin Abdulazeez, A. (2021). PJAEE, 17 (7) (2021) Meta-heuristic algorithms for k-means clustering: a review meta-heuristic algorithms for k-means clustering: a review meta-heuristic algorithms for k-means clustering: a review. <https://www.researchgate.net/publication/349054880>
- Omol, E., Onyangor, D., Mburu, L., & Abuonji, P. (2024). Application of k-means clustering for customer segmentation in grocery stores in kenya. in international journal of science. <http://ijstm.inarah.co.id>
- Ong, A. K. S., Prasetyo, Y. T., Esteller, A. J. D., Bruno, J. E., Lagorza, K. C. O., Oli, L. E. T., Chuenyindee, T., Thana, K., Persada, S. F., & Nadlifatin, R. (2023). Consumer preference analysis on the attributes of samgyeopsal Korean cuisine and its market segmentation: Integrating conjoint analysis and K-means clustering. *PLOS ONE*, 18(2), e0281948. <https://doi.org/10.1371/journal.pone.0281948>
- Raeisi, M., & Sesay, A. B. (2022). A distance metric for uneven clusters of unsupervised k-means clustering algorithm. *ieee access*, 10, 86286–86297. <https://doi.org/10.1109/access.2022.3198992>
- Saxena, A., Agarwal, A., Pandey, B. K., & Pandey, D. (2024). examination of the criticality of customer segmentation using unsupervised learning methods.

- circular economy and sustainability*, 4(2), 1447–1460.
<https://doi.org/10.1007/s43615-023-00336-4>
- Sudirjo, F. (2023). marketing strategy in improving product competitiveness in the global market. *journal of contemporary administration and management (adman)*, 1(2), 63–69. <https://doi.org/10.61100/adman.v1i2.24>
- Tabianan, K., Velu, S., & Ravi, V. (2022). K-means clustering approach for intelligent customer segmentation using customer purchase behavior data. *sustainability*, 14(12), 7243. <https://doi.org/10.3390/su14127243>
- Zubair, Md., Iqbal, MD. A., Shil, A., Chowdhury, M. J. M., Moni, M. A., & Sarker, I. H. (2024). An improved k-means clustering algorithm towards an efficient data-driven modeling. *annals of data science*, 11(5), 1525–1544. <https://doi.org/10.1007/s40745-022-00428-2>