



Analisis Preferensi Pelanggan Semanis Coffe And Resto terhadap Kombinasi Menu Makanan dan Minuman Menggunakan Data Mining (Association Rule Dan Clustering)

Adit Septian Saepul Millah^{1*}, Hendi Suhendi²

¹⁻²Sistem informasi, Universitas Adhirajasa Reswara Sanjaya, Indonesia

*Penulis korespondensi: aditseptiansaepulmillah242@gmail.com¹

Abstract. *The coffee shop industry in Indonesia is experiencing rapid growth that requires business owners to optimize data-driven strategies. This study aims to analyze customer preferences at Semanis Coffee and Resto using data mining methods to support more effective business decision-making. The method used is Market Basket Analysis with the FP-Growth algorithm for association rule mining and the K-Means algorithm for customer segmentation. The research data consists of 672 sales transactions during the March-May 2025 period. The results of the association analysis with a minimum support of 0.004 and a minimum confidence of 0.2 resulted in five valid rules with a lift ratio above 1. The strongest rule is the combination of Americano→Milk Choco with a confidence of 42.9% and an elevator ratio of 5.229, indicating a strong linkage between products. The most popular products are Milk Choco (10.8%) and Americano (8.5%). Customer segmentation analysis identified three clusters: Cluster 0 (Loyal Customers) 80% with high frequency but low transaction value; Cluster 1 (Occasional Customers) 10% with low activity; and Cluster 2 (Large Buyers) 10% with high transaction value but low frequency. This study concludes that product bundling strategies, loyalty programs, reactivation campaigns, and premium services can be applied to increase the effectiveness of coffee shop businesses.*

Keywords: *Coffee Shop; Customer Preferences; Data Mining; FP-Growth; K-Means*

Abstrak. Industri kedai kopi di Indonesia mengalami pertumbuhan pesat yang menuntut pemilik bisnis untuk mengoptimalkan strategi berbasis data. Penelitian ini bertujuan menganalisis preferensi pelanggan di Semanis Coffee and Resto dengan metode *data mining* guna mendukung pengambilan keputusan bisnis yang lebih efektif. Metode yang digunakan adalah *Market Basket Analysis* dengan algoritma FP-Growth untuk penambangan aturan asosiasi serta algoritma K-Means untuk segmentasi pelanggan. Data penelitian terdiri atas 672 transaksi penjualan selama periode Maret–Mei 2025. Hasil analisis asosiasi dengan minimum support 0,004 dan minimum confidence 0,2 menghasilkan lima aturan valid dengan rasio lift di atas 1. Aturan terkuat adalah kombinasi Americano→Milk Choco dengan confidence 42,9% dan rasio lift 5,229, menunjukkan keterkaitan kuat antarproduk. Produk paling populer adalah Milk Choco (10,8%) dan Americano (8,5%). Analisis segmentasi pelanggan mengidentifikasi tiga klaster: Cluster 0 (Pelanggan Setia) 80% dengan frekuensi tinggi namun nilai transaksi rendah; Cluster 1 (Pelanggan Sesekali) 10% dengan aktivitas rendah; dan Cluster 2 (Pembeli Besar) 10% dengan nilai transaksi tinggi namun frekuensi rendah. Penelitian ini menyimpulkan bahwa strategi bundling produk, program loyalitas, kampanye reaktivasi, dan layanan premium dapat diterapkan untuk meningkatkan efektivitas bisnis kedai kopi.

Kata kunci: Data Mining; FP-Growth; Kedai Kopi; K-Means; Preferensi Pelanggan

1. LATAR BELAKANG

Industri makanan dan minuman, khususnya sektor kafe, mengalami perkembangan pesat di Indonesia. Tidak hanya sekadar tempat untuk menikmati minuman dan makanan, kafe juga telah menjelma menjadi ruang sosial, tempat bekerja, dan sarana rekreasi ringan. Perubahan fungsi ini menandai bergesernyaperilaku konsumen serta meningkatnya tuntutan terhadap kenyamanan dan pelayanan. Menurut (Lim Nathanael 2023), nilai penjualan sektor spesialis kopi dan teh di Asia Pasifik mencapai USD 26,9 miliar dengan tingkat pertumbuhan tahunan sebesar 8 persen pada periode 2017 hingga 2022. Di Indonesia sendiri, pola konsumsi masyarakat cenderung beralih dari konsumsi di rumah ke luar rumah. Channel foodservice

seperti kafe dan restoran menunjukkan potensi pertumbuhan yang lebih besar dibandingkan saluran ritel rumahan.

Hal ini menunjukkan bahwa sektor kafe memiliki peluang besar untuk terus berkembang, terutama jika didukung dengan inovasi pelayanan dan pemanfaatan teknologi. Meskipun peluangnya besar, tidak semua pelaku usaha mampu memaksimalkan potensi tersebut. Banyak kafe yang belum memiliki strategi bisnis berbasis data, sehingga pengambilan keputusan seperti pengembangan menu, penentuan harga, dan promosi dilakukan secara konvensional tanpa mempertimbangkan preferensi pelanggan secara spesifik. Padahal, data transaksi yang dimiliki dapat dianalisis untuk mengetahui pola konsumsi pelanggan secara akurat. Di Semanis Coffee and Resto misalnya, belum tersedia sistem segmentasi pelanggan berdasarkan preferensi dan perilaku pembelian. Selain itu, kombinasi produk yang sering dibeli bersamaan juga belum dianalisis, sehingga potensi dalam pembuatan paket menu atau strategi bundling belum dimanfaatkan secara optimal. Di tengah persaingan yang semakin kompetitif, pengambilan keputusan yang tidak berbasis data akan menyulitkan kafe untuk bertahan dan berkembang.

Oleh karena itu, dibutuhkan pendekatan sistematis dan berbasis teknologi untuk membantu proses pengambilan keputusan yang lebih tepat sasaran. Penelitian ini hadir untuk menjawab kebutuhan tersebut, yaitu dengan memanfaatkan data transaksi pelanggan untuk menganalisis preferensi mereka menggunakan pendekatan data mining. Metode data mining menawarkan solusi strategis melalui pendekatan association rule dan clustering. Association rule dapat digunakan untuk menemukan kombinasi produk yang sering dibeli bersamaan, seperti yang ditemukan oleh (Valencia and Tisno Atmojo 2024) dalam penelitian mereka pada CanNgopi, di mana kombinasi produk dengan confidence tinggi dapat menjadi dasar pembuatan paket menu. Sementara itu, metode clustering seperti yang digunakan oleh (Shiddiq Ahmad Fadhilah 2022; Padli, Yasir, and Hasugian 2023) dapat mengelompokkan pelanggan berdasarkan karakteristik perilaku konsumsi, sehingga strategi promosi dapat disesuaikan secara lebih personal. Dengan menggabungkan kedua pendekatan tersebut, pelaku usaha dapat menyusun strategi yang lebih efektif dan berbasis data.

2. KAJIAN TEORITIS

Preferensi pelanggan didefinisikan sebagai keputusan menyukai atau tidak menyukai seseorang terhadap suatu produk barang atau jasa yang dikonsumsi [1]. Dalam konteks bisnis kafe, preferensi pelanggan menjadi faktor krusial yang menentukan keberhasilan strategi pemasaran dan pengembangan produk. Perilaku konsumen merupakan analisis mengenai

bagaimana individu, cluster, dan organisasi menentukan pilihan, menyediakan, memanfaatkan, serta bagaimana produk, layanan, gagasan, atau pengalaman digunakan untuk memenuhi kebutuhan dan keinginan mereka [2].

Data mining merupakan proses mengekstraksi informasi berharga yang tersembunyi di dalam database menggunakan berbagai metode, mulai dari statistik dan matematika, hingga kecerdasan buatan dan machine learning. Dalam konteks analisis preferensi pelanggan, dua metode data mining yang paling relevan adalah association rules dan clustering [3]. Association rules adalah aturan asosiatif yang memenuhi syarat minimum untuk confidence dan support dalam sebuah database. Metode ini menggunakan tiga parameter utama: support yang mengukur seberapa dominan suatu item dalam seluruh transaksi, confidence yang mencerminkan keterkaitan antara dua elemen, dan lift ratio yang mengukur kekuatan hubungan antara item [4].

Clustering adalah aspek data mining yang bersifat tanpa pengawasan, melibatkan pemisahan data ke dalam kategori berdasarkan tingkat kesamaan [5]. Dalam penelitian ini, algoritma K-Means dipilih karena kesederhanaannya, kemudahan implementasi, serta kemampuannya menangani data berukuran besar dengan kompleksitas waktu linear $O(nKT)$ [6]. Untuk association rules, algoritma FP-Growth digunakan karena efisiensinya yang hanya memerlukan dua kali pemindaian data untuk membangun struktur pohon, berbeda dengan algoritma apriori yang sering menghasilkan kombinasi berlebihan [7].

Market Basket Analysis (MBA) merupakan aplikasi praktis dari association rules yang digunakan untuk memperkirakan produk mana yang diminati konsumen berdasarkan pola pembelian tidak disengaja [8]. Fadilah [9] melakukan analisis perilaku pembelian konsumen menggunakan metode association rule dan clustering pada Jore Coffee & Eatery. Penelitian tersebut menemukan 21 aturan asosiasi dengan pola terkuat menunjukkan bahwa pembelian kategori non-coffee, tea, dan rice bowl seringkali diikuti oleh pembelian snack. Segmentasi pelanggan menggunakan K-Means menghasilkan tiga klaster: Loyal Customers (77,65%), Big Customers (0,19%), dan Impulsive Customers (22,16%). Arrasyid et al. [16] menggunakan algoritma FP-Growth dan K-Means untuk rekomendasi paket menu pada Kopi Mesra Abadi dengan membagi waktu menjadi tiga periode, menghasilkan rekomendasi paket menu yang spesifik untuk setiap periode.

Shafira dan Anggreini [17] menganalisis pola pembelian menggunakan algoritma Apriori pada Coffee Shop Gelora Fantasi, menemukan keterkaitan signifikan antara menu Cappuccino dan Kopi Susu Aren dengan dukungan tertinggi 14,5%. Valencia dan Atmojo [10] menggunakan algoritma FP-Growth pada CanNgopi dengan hasil support 31,9% dan confidence

96,5% untuk kombinasi produk snack, signature, dan paket pasta. Nugraheni dan Nugroho [11] menerapkan Market Basket Analysis dengan algoritma Apriori untuk menganalisis pembelian jajanan khas lebaran, menemukan pola signifikan dengan confidence tertinggi 85,7% untuk kombinasi NASTAR dan MONDE menuju KASTANGEL.

Padli et al. [12] menerapkan K-Means untuk mengelompokkan menu makanan dan minuman berdasarkan tingkat penjualan, berhasil membagi menu ke dalam beberapa klaster yang merepresentasikan cluster menu dengan karakteristik performa penjualan yang serupa. Penelitian-penelitian tersebut menunjukkan efektivitas penerapan association rule dan clustering dalam analisis bisnis kafe, namun masih terbatas pada implementasi terpisah atau fokus pada aspek tertentu saja.

3. METODE PENELITIAN

Objek Penelitian

Objek dari penelitian ini adalah informasi mengenai transaksi penjualan yang diambil dari Semanis Coffee and Resto, sebuah kedai kopi dan restoran yang terletak di Jl. Panyadap Majalaya No.25-26, Ds. Sukamanah, Kec. Majalaya Kab. Bandung. Data yang dianalisis meliputi semua catatan transaksi penjualan yang berlangsung dalam rentang waktu dari Maret 2025 sampai Mei 2025. Catatan transaksi ini mencakup berbagai fitur yang berkaitan dengan pola pembelian produk, termasuk daftar menu makanan dan minuman yang dipesan oleh pelanggan, jumlah pembelian, frekuensi pembelian, serta tanggal dan waktu transaksi. Pemilihan data ini bertujuan untuk memperoleh gambaran menyeluruh mengenai kebiasaan pelanggan, sehingga bisa digunakan sebagai bahan analisis untuk penerapan metode Association Rule–Market Basket Analysis dan Clustering.

Alur Penelitian

Proses kerja dalam penelitian ini dijelaskan melalui diagram alur pada Gambar 1. Sesuai dengan alur penelitian yang telah dirancang, untuk memudahkan pemahaman, dibuatlah penjelasan rinci mengenai langkah-langkah sebagai berikut: Langkah awal penelitian difokuskan pada observasi lapangan guna memetakan kondisi nyata serta menggali isu-isu krusial yang melandasi studi ini. Melalui proses identifikasi masalah, akar persoalan ditemukan untuk kemudian dirumuskan menjadi sasaran utama serta ruang lingkup batasan penelitian.

Studi Literatur: Mencakup rujukan dan konsep yang diperlukan untuk menyelesaikan laporan penelitian. Referensi yang digunakan berasal dari artikel jurnal dan buku yang relevan dengan isu yang diteliti, khususnya yang berkaitan dengan pemanfaatan algoritma FP-Growth dalam kerangka *Market Basket Analysis* (MBA), serta penggunaan teknik *K-Means*

untuk proses klasifikasi kelompok data

Studi Lapangan: Dilaksanakan untuk mendapatkan pemahaman langsung tentang cara kerja bisnis, jenis transaksi, dan preferensi pelanggan di Semanis Coffee and Resto. Temuan dari studi lapangan digunakan sebagai tambahan dan dukungan untuk analisis data transaksi menggunakan metode Association Rules (FP-Growth) dan Clustering (K-Means).

Pengumpulan Data: Dilakukan menggunakan data historis transaksi penjualan dalam rentang waktu Maret hingga Mei 2025 dan melakukan wawancara dengan manajemen serta staf. Wawancara dilakukan untuk mengonfirmasi pola belanja yang terdeteksi dari data transaksi, kebutuhan bisnis yang belum dapat diukur dengan angka, serta kendala yang dihadapi dalam pengelolaan produk dan pengembangan strategi promosi.

Preprocessing Data

Untuk rentang waktu Maret hingga Mei 2025, data transaksi penjualan dari Semanis Coffee and Resto perlu diseleksi terlebih dahulu. Dalam proses data mining, preprocessing data menjadi tahapan penting, mengingat tidak semua data atau atributnya relevan untuk digunakan. Berikut adalah langkah-langkah preprocessing data yang dilakukan:

Input Data Selection: Sebelum memulai penggalian informasi dalam data mining, pemilihan data dari himpunan data operasional wajib dilakukan. Proses reduksi data dilakukan guna mengeliminasi indikator yang dianggap kurang berkaitan dengan fokus studi. Setelah melalui tahap seleksi, data tersebut akan dipisahkan ke dalam format dokumen khusus, terlepas dari basis data operasional, untuk kemudian dianalisis pada fase penggalian data (*data mining*). Transformation Data: Proses ini melibatkan perubahan format data asli agar memiliki skala yang konsisten. Hal tersebut bertujuan agar dataset memenuhi kriteria atau persyaratan teknis tertentu, sehingga hasil analisis yang diperoleh menjadi lebih valid dan reliabel.

Metode Data Mining

Data mining adalah proses pencarian pola atau informasi menarik dalam data yang sudah dipilih, dengan menerapkan teknik atau metode khusus. Dalam penelitian ini RapidMiner digunakan sebagai perangkat lunak untuk mengolah data transaksi penjualan dari Semanis Coffee and Resto.

Implementasi Aturan Asosiasi Melalui FP-Growth: Dalam penelitian ini, analisis pola keterkaitan dilakukan menggunakan teknik *Market Basket Analysis* berbasis algoritma FP-Growth. Proses pengolahan data transaksi tersebut dijalankan dengan bantuan aplikasi RapidMiner guna meningkatkan efisiensi dan akurasi dalam tahap ekstraksi informasi.

Clustering dengan Algoritma K-Means: Dengan RapidMiner sebagai perangkat lunak

yang dipakai, pengolahan data berikutnya dilakukan dengan metode pengelompokan menggunakan algoritma K- Means dengan data histori transaksi penjualan selama 3 bulan terakhir yang dijadikan acuan untuk menyelesaikan isu penelitian.

Alat dan Bahan Penelitian

Alat Penelitian: Dalam penelitian ini digunakan komputer dengan spesifikasi RedmiBook 15, Windows 10 Home Single Language 64-bit, Prosesor Intel(R) Core(TM) i3-1115G4 @ 3.00 GHz, RAM 8 GB, Intel UHD Graphics (terintegrasi), DirectX 12, dan SSD sebagai penyimpanan. Program aplikasi yang digunakan meliputi: RapidMiner Studio untuk mengolah data dan menjalankan algoritma FP-Growth (Association Rule) dan K-Means (Clustering); Microsoft Excel untuk menghimpun dan menyiapkan data transaksi sebelum diolah; dan Microsoft Word untuk menyusun dan menulis laporan akhir penelitian.

Bahan Penelitian: Meliputi data history transaksi penjualan Semanis Coffee and Resto dari bulan Maret sampai Mei 2025 dalam format file Excel, berbagai bacaan dan sumber acuan seperti artikel, buku, dan jurnal ilmiah yang berhubungan untuk menunjang pembahasan serta analisis, dan catatan observasi dan hasil wawancara langsung di lokasi studi guna melengkapi data dan memperkuat hasil diskusi.

4. HASIL DAN PEMBAHASAN

Pengumpulan data

Pada penelitian ini, data yang digunakan adalah data lama perusahaan, yaitu data *history* transaksi penjualan selama tiga bulan yang dimulai dari Maret 2025 hingga Mei 2025. Proses pengumpulan data dilakukan melalui sistem POS (*Point of Sale*) yang sudah tercatat dalam database internal perusahaan dengan melakukan ekspor dalam format Excel sesuai dengan data yang diperlukan. Terdapat 22 *item* menu makanan dan 31 *item* menu minuman Dataset tersebut mencakup daftar produk yang dirangkum dalam tabel berikut:

Tabel 1. Menu dan produk.

No	Kategori Menu	Produk
1.	Nasi & mie	Nasi Goreng,Rice Bowl Ayam,Rice Bowl Beef,Mie Geulis Original,Spaghetti Aglio e Olio
2.	Ayam & daging	Chicken Katsu,Ayam Geprek, Ayam Teriyaki,Sate Taichan,Tenderloin Steak
3.	Snack & cemilan	Cireng Bumbu,Tahu Lada Garam,French Fries & Sausage,Chicken Burger,Pisang Goreng

4.	Bakso aci neng geulis (bacineng)	Bacineng Manja,Bacineng	Lucu,Bacineng Wagyu,Bacineng Siomay
5.	Cake	Croffle, Cheese Cake	
6.	Mojito & tea	Strawberry Mojito,Mojito Lychee, Mojito Strawberry, Lemon Tea,Strawberry Tea, Lychee Tea,Sweet Tea	
1.	Manual brew (hot)	V60, Japanese, Vietnam Drip	
2.	Hot latte	Latte, Mocca Latte	
3.	Black coffee	Long Black,AmericanoEspresso	
4.	Ice coffee	Caramel Latte,Hazelnut Latte, Creme Brulee, Es Kopi Susu, Charcoal Latte, Choco Latte Lemonpresso, Cappucino	
5.	Milk base	Red Velvet, Milk Charcoal, Matcha, Taro, Milk Choco	

Pada data history transaksi penjualan sudah di ubah menjadi data excel yang di ambil dari system POS (point of sale) yang digunakan di Semanis coffee and resto. Berikut merupakan contoh data transaksi penjualan yang sudah di ubah menjadi data excel:

No_Meja	Tanggal	Waktu	Nama_Pesanan	Jumlah	Harga
8	02/03/2025	12:34:00	Croffle	2	20000
1	02/03/2025	13:12:00	Chicken Katsu	5	150000
7	02/03/2025	14:25:00	Rice Bowl Ayam	1	25000
4	02/03/2025	15:47:00	Rice Bowl Beef	2	60000
9	02/03/2025	16:30:00	Mojito - Strawberry	2	50000
2	02/03/2025	17:15:00	Matcha	1	22000
5	02/03/2025	18:00:00	Lemon Tea	1	18000
6	02/03/2025	18:45:00	Spagetty Aglio e Olio	1	30000
10	03/03/2025	13:26:00	Croffle	2	40000
3	03/03/2025	13:26:00	Chicken Katsu	5	150000
2	03/03/2025	13:33:00	Rice Bowl Ayam	1	25000
7	03/03/2025	14:45:00	Rice Bowl Beef	2	60000
3	03/03/2025	15:51:00	Mie Geulis Original	1	18000
5	03/03/2025	16:13:00	Mojito - Strawberry	2	50000
1	03/03/2025	16:22:00	Milk Choco	1	22000
9	03/03/2025	16:30:00	Matcha	1	22000
8	03/03/2025	16:40:00	Lemon Tea	1	18000
4	03/03/2025	17:44:00	Strawberry Tea	3	54000
6	03/03/2025	18:13:00	Sweet Tea	1	10000

Gambar 1. Contoh data transaksi penjualan.

Berdasarkan data diatas, dapat diidentifikasi beberapa kolom penting untuk dilakukan analisis bisnis seperti no_meja, tanggal, waktu, nama_pesanan, jumlah dan harga. Sebelum memproses data ini, diperlukan tahap *pre-processing* untuk menentukan jenis data yang bisa

diproses menggunakan *software* Rapidminer. Ada sejumlah perbedaan dalam tahap *pre-processing data* untuk *association rule* dan *K-means clustering*.

Market Basket Analysis (MBA)

Merupakan pendekatan analisis data yang bertujuan untuk menemukan pola keterkaitan antar item yang sering muncul bersamaan dalam suatu transaksi. Teknik ini umumnya digunakan dalam bidang ritel dan pemasaran untuk memahami perilaku konsumen berdasarkan data historis pembelian.[13]. Dalam konteks penelitian ini, MBA digunakan sebagai dasar untuk menganalisis hubungan antar menu makanan dan minuman yang dipesan oleh pelanggan di suatu tempat usaha.

Transormasi data

Dalam penelitian ini, data transaksi penjualan yang digunakan terlebih dahulu dihilangkan informasinya pada kolom-kolom yang tidak relevan [14]. Misalnya No_meja, Tanggal, Waktu, Nama_pesanan, jumlah, harga. Untuk analisis hubungan antar item, kolom No_meja dan Tanggal digabung menjadi satu entitas yang dinamakan Transaksi_ID, sehingga setiap transaksi memiliki identifikasi yang unik. Kemudian, kolom nama_pesanan diubah menjadi *Item* agar sesuai dengan format input yang digunakan dalam algoritma *FP-Growth*. Berikut adalah hasil transformasi data untuk *FP-GROWTH*:

Transaksi_ID	Item
8_02/03/2025	Croffle
1_02/03/2025	Chicken Katsu
7_02/03/2025	Rice Bowl Ayam
4_02/03/2025	Rice Bowl Beef
9_02/03/2025	Mojito - Strawberry
2_02/03/2025	Matcha
5_02/03/2025	Lemon Tea
6_02/03/2025	Spagetty Aglio e Olio
10_03/03/2025	Croffle
3_03/03/2025	Chicken Katsu
2_03/03/2025	Rice Bowl Ayam
7_03/03/2025	Rice Bowl Beef
3_03/03/2025	Mie Geulis Original
5_03/03/2025	Mojito - Strawberry
1_03/03/2025	Milk Choco
9_03/03/2025	Matcha
8_03/03/2025	Lemon Tea
4_03/03/2025	Strawberry Tea
6_03/03/2025	Sweet Tea

Gambar 2. Hasil transformasi data untuk FP-GROWTH.

Transformasi ini bertujuan untuk mempersiapkan struktur data yang sesuai dengan kebutuhan *algoritma FP-Growth*, yang berfungsi untuk menemukan frekuensi kemunculan gabungan item dalam dataset transaksi. *Format pivot biner* memungkinkan perhitungan tersebut dilakukan dengan cara yang efisien karena sistem dapat dengan mudah membaca keterkaitan antar item berdasarkan pola kemunculan angka 1 dalam baris yang sama. Berikut

merupakan contoh data yang di ubah ke bentuk biner:

Transaksi_ID	Americano	Ayam Gorep	Ayam Teriyaki	Bacang L	Bacang P	Bacang C	Bacang V	Banana Sn	Cappuccino	Caramel L	Caramel U	Charcoal L	Charcoal U	Cheese Ca	Chicken B	Chicken K	Chicken K	Choco Latte	Direng Bur
1_01/05/2025	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
1_02/03/2025	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
1_02/04/2025	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
1_02/05/2025	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1_03/03/2025	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1_03/04/2025	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1_03/05/2025	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
1_04/04/2025	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
1_04/05/2025	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
1_05/03/2025	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
1_05/04/2025	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
1_05/05/2025	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
1_06/03/2025	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Gambar 3. Hasil transformasi data biner.

Tahapan ini diawali dengan proses reduksi data guna menyederhanakan struktur dataset mentah agar sesuai dengan format numerik yang dibutuhkan dalam proses klusterisasi [15]. Informasi-informasi transaksional seperti Nama_Pesanan, Waktu, dan Harga per item tidak digunakan secara langsung, melainkan diolah menjadi data ringkasan per pelanggan. Kolom Jumlah dan Harga dikalikan untuk membentuk atribut baru bernama Total_Harga, yang merepresentasikan nilai transaksi aktual. Kemudian, melalui proses agregasi berbasis No_meja, diperoleh indikator perilaku pelanggan seperti Total_Pembelian, Total_Item, Frekuensi_Transaksi, dan Jumlah_Jenis_Menu. Keempat atribut ini digunakan sebagai fitur utama dalam algoritma *K-Means*. Berikut merupakan hasil Transformasi data untuk *K-Means*.

Row Labels	Total_Pembelian	Total_Item	Frekuensi_Transaks	Jumlah_Jenis_Menu
1	15413000	240	129	129
2	267219000	345	131	131
3	270777000	351	131	131
4	18980000	226	130	130
5	943771000	484	130	130
6	266625000	340	131	131
7	283091000	349	130	130
8	289737000	338	128	128
9	1842798000	700	130	130
10	17488000	242	131	131

Gambar 4. Hasil transformasi data untuk K-Means.

Setelah data pelanggan diubah menjadi format numerik agregatif melalui proses transformasi, tahapan selanjutnya adalah normalisasi data. Metode yang digunakan dalam penelitian ini adalah *Z-Transformation*, yaitu dengan mengubah nilai setiap atribut berdasarkan rata-rata (mean) dan standar deviasi (standard deviation) dari kolom tersebut. Hasil transformasi ini menjadikan nilai-nilai berada dalam distribusi standar, dengan rata-rata mendekati nol dan standar deviasi mendekati satu. Proses normalisasi dilakukan menggunakan operator Normalize dalam aplikasi RapidMiner, dengan memilih seluruh atribut numerik hasil agregasi sebagai target (Whendasmoro and Joseph 2022). Berikut adalah hasil akhir dari proses normalisasi data.

Row No.	Total_Pemb...	Total_item	Frekuensi_T...	Jumlah_Jeni...
1	-0.717	-0.863	-1.106	-1.106
2	-0.272	-0.117	0.905	0.905
3	-0.266	-0.075	0.905	0.905
4	-0.710	-0.962	-0.101	-0.101
5	0.921	0.870	-0.101	-0.101
6	-0.273	-0.153	0.905	0.905

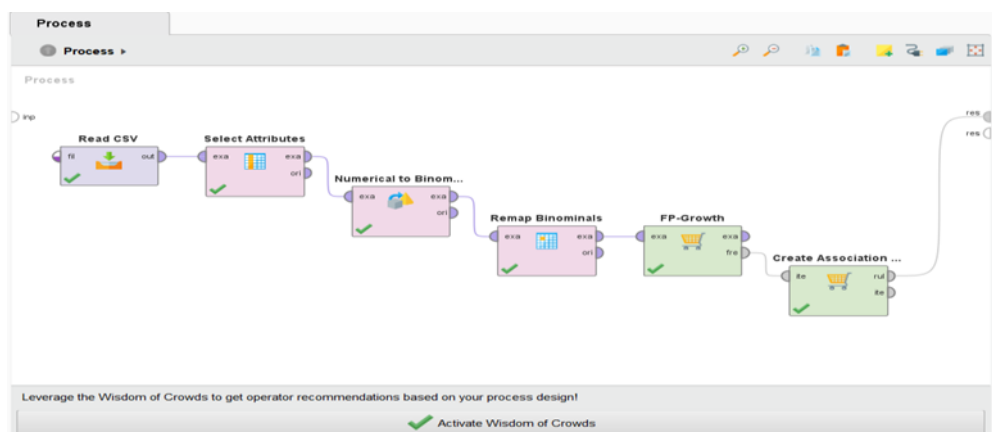
ExampleSet (10 examples, 0 special attributes, 4 regular attributes)

Gambar 5. Hasil akhir dari proses normalisasi data.

Setiap fitur numerik dikonversi ke dalam bentuk skor standar deviasi, di mana nilai positif menunjukkan pelanggan yang berada di atas rata-rata, dan nilai negatif menunjukkan sebaliknya. Transformasi ini bertujuan untuk menyetarakan skala antar fitur numerik agar algoritma K-Means dapat mengelompokkan data secara seimbang tanpa bias terhadap fitur tertentu.

Metode Association Rule FP-Growth

Data historis dari transaksi pelanggan yang dimanfaatkan dalam *Association rule* telah mengalami proses transformasi menjadi bentuk *pivot biner* [17]. Setelah tahap pembersihan dan penggabungan item untuk setiap transaksi, didapatkan total 672 baris data yang masing-masing menunjukkan satu transaksi yang berbeda. Semua data tersebut dipakai sebagai landasan uji menggunakan algoritma *FP-Growth* dengan *minimum support* sebesar 0,004 dan *confidence* sebesar 0,2. Analisis ini dilakukan dengan menggunakan perangkat lunak *RapidMiner Studio*, yang menghasilkan aturan asosiasi berdasarkan seberapa sering kombinasi item muncul dalam transaksi pelanggan. Berikut merupakan proses analisis menggunakan *RapidMiner Studio*:



Gambar 6. Proses pada RapidMiner untuk FP-Growth.

Sebagaimana diilustrasikan pada gambar tersebut, tahapan pengolahan data menggunakan algoritma ini diawali dengan memasukkan dataset yang telah ditransformasi ke dalam sistem. Pengaturan parameter dilakukan pada operator *Select Attribute* dengan

memilih mode 'Subset' agar variabel 'Transaksi_ID' dapat dipisahkan dari analisis. Selanjutnya, dilakukan konversi data melalui fungsi *Numerical to Binomial* dan *Remap Binomials* (0 untuk negatif, 1 untuk positif). Untuk menghasilkan pola asosiasi, algoritma FP- Growth diintegrasikan dengan pengaturan *Min Support* di angka 0,004. Akhirnya, pembentukan aturan dilakukan oleh operator *Create Association Rules* yang disetel pada batas *Confidence* minimum 0,2.

Analisis Hasil Association Rule Mining

Implementasi algoritma FP-Growth pada penelitian ini menghasilkan pola hubungan antar-item dengan tingkat kepercayaan 0,2 dan dukungan minimal 0,004. Nilai lift yang melebihi ambang batas 1 pada aturan yang terbentuk menunjukkan bahwa kemunculan satu item memperkuat probabilitas kemunculan item lainnya. Adapun data lengkap association rule disajikan pada tabel berikut.

Tabel 2. Hasil aturan asosiasi.

No	Premise	Conclusion	Support	Confidence	Lift
1	Bacineng Lucu	Milk Choco	0.004	0.231	2.815
2	Strawberry Tea	Tahu Lada Garam	0.004	0.273	1.270
3	ayam goreng	Taro	0.004	0.300	1.346
4	Ayam Teriyaki	Milk Charcoal	0.004	0.333	1.393
5	Americano	Milk Choco	0.009	0.429	5.229

Dari hasil pengujian yang dilakukan menggunakan algoritma FP-Growth dengan *minimum support* sebesar 0,004 serta *confidence* sebesar 0,2, diperoleh sejumlah aturan asosiasi yang mempunyai nilai lift di atas 1, yang berarti terdapat keterkaitan positif antar item yang dianalisis. Salah satu aturan dengan nilai lift tertinggi adalah:

Americano \Rightarrow *Milk Choco* > *Confidence: 42,9% | Lift: 5,229*

Artinya, pelanggan yang membeliAmericano memiliki kemungkinan lebih besar untuk juga membeli Milk Choco, dengan kekuatan hubungan lebih dari lima kali dibandingkan peluang acak. Hal ini menunjukkan bahwa kedua item tersebut sering dibeli bersama dalam pola konsumsi nyata, dan berpotensi tinggi untuk dijadikan bundling atau promosi gabungan.

Analisis Frekuensi Pembelian Produk

Setelah dilakukan pembersihan melalui teknik *data reduction*, diperoleh gambaran mengenai intensitas belanja pada periode bulan maret 2025 sampai mei 2025:

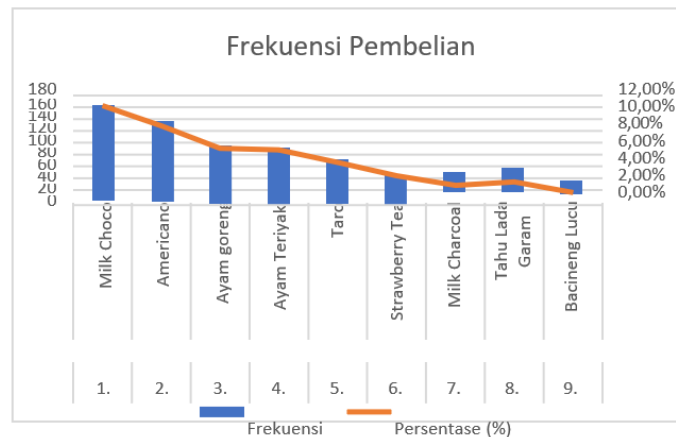
Tabel 3. Frekuensi kemunculan item pada transaksi.

No	Item	Frekuensi	Persentase (%)
1.	Milk Choco	162	10,8%
2.	Americano	136	8,5%
3.	Ayam goreng	96	6,0%
4.	Ayam Teriyaki	93	5,8%
5.	Taro	71	4,4%
6.	Strawberry Tea	47	2,9%
7.	Milk Charcoal	29	1,8%
8.	Tahu Lada Garam	35	2,2%
9.	Bacineng Lucu	17	1,0%

Dari tabel tersebut dapat dijelaskan bahwa:

- Produk Milk Choco menjadi item dengan frekuensi pembelian tertinggi, yaitu sebanyak 162 kali atau 10,8% dari keseluruhan transaksi, menjadikannya produk paling dominan dalam periode tersebut.
- Americano menyusul di posisi kedua dengan 136 transaksi (8,5%), dan kerap muncul sebagai bagian dari aturan asosiasi.
- Produk makanan utama seperti Ayam Goreng dan Ayam Teriyaki menunjukkan preferensi stabil dari pelanggan, dengan kontribusi gabungan lebih dari 11% dari total pembelian.
- Item seperti Taro, Strawberry Tea, dan Milk Charcoal menempati kategori menengah, sedangkan Bacineng Lucu meskipun memiliki frekuensi rendah tetap masuk dalam aturan asosiasi berkat nilai lift yang signifikan.

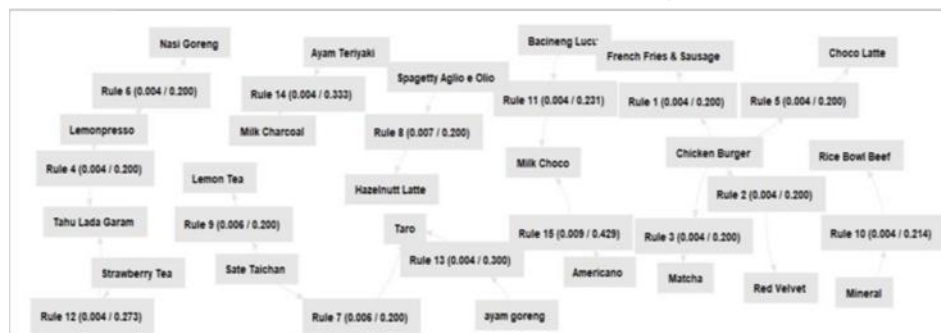
Berikut adalah hasil analisis yang menunjukkan seberapa sering masing-masing produk dibeli. Data ini kemudian diurutkan, dimulai dari produk yang paling sering dibeli hingga yang paling jarang, dan disajikan dalam bentuk diagram berikut:



Gambar 6. Frekuensi pembelian.

Output Association Rule

Penentuan strategi pemasaran dalam penelitian ini didasarkan pada pola perilaku belanja pelanggan yang terekstraksi melalui aturan asosiasi. Hasil tersebut diperoleh dari pemrosesan dataset dengan menetapkan ambang batas *support* minimal 0,004 serta tingkat *confidence* di angka 0,2. Parameter ini dipilih karena menghasilkan aturan asosiasi dengan nilai lift ratio terbesar mencapai 5,229, yang mengindikasikan kekuatan asosiasi yang sangat signifikan antar item dalam satu transaksi. Pola keterkaitan yang terbentuk dari hasil pengolahan dataset tersebut ditampilkan secara visual dalam grafik aturan berikut ini:



Gambar 7. Hasil Tree.

Hanya hasil *rules* dengan indeks *lift ratio* > 1 yang diikutsertakan dalam proses analisis. Aturan- aturan valid tersebut dipaparkan dalam rincian di bawah ini:

Rule 1 terbentuk antara item *Americano* dengan *Milk Choco*. *Americano* merupakan item pendahulu, sedangkan *Milk Choco* menjadi item pengikut. Kombinasi ini muncul dalam 0,9% transaksi, dengan *confidence* sebesar 42,9%. Aturan ini dinilai sangat valid karena memiliki *lift ratio* tertinggi sebesar 5,229, menunjukkan bahwa pelanggan yang membeli *Americano* berpeluang lima kali lebih tinggi membeli *Milk Choco* dalam satu transaksi dibandingkan peluang acak.

Rule 2 terbentuk antara *Ayam Teriyaki* dengan *Milk Charcoal*, yang muncul bersama dalam 0,4% dari total transaksi. Nilai *confidence* sebesar 33,3% menunjukkan bahwa satu dari tiga pembeli *Ayam Teriyaki* juga membeli *Milk Charcoal*. Dengan *lift* sebesar 1,393, aturan ini dinyatakan valid dan relevan untuk pengemasan menu makanan dan minuman dalam satu paket.

Rule 3 terbentuk antara *Ayam Goreng* dan *Taro*. Kedua item ini muncul dalam 0,4% transaksi dengan nilai *confidence* sebesar 30,0% dan *lift* sebesar 1,346. Artinya, hampir sepertiga pelanggan yang membeli *Ayam Goreng* juga memilih *Taro*, mengindikasikan kecenderungan konsumen untuk mengombinasikan makanan utama dengan minuman manis.

Rule 4 menunjukkan hubungan antara *Strawberry Tea* sebagai pendahulu dan *Tahu Lada Garam* sebagai pengikut, dengan *confidence* sebesar 27,3% dan *lift* 1,270. Meski nilai

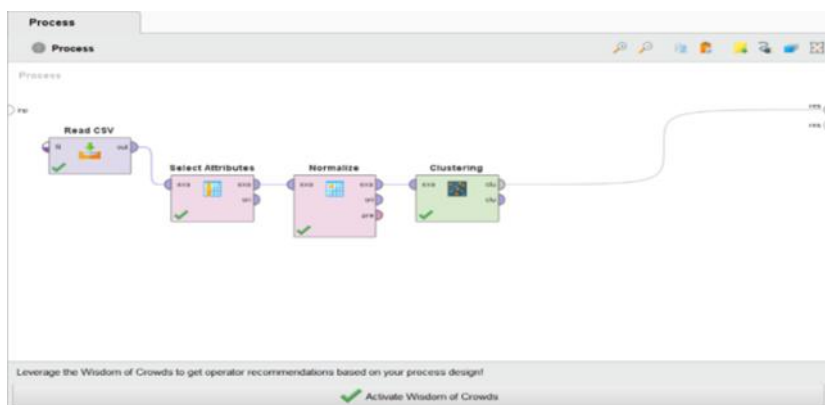
Analisis Preferensi Pelanggan Semanis Coffe And Resto terhadap Kombinasi Menu Makanan dan Minuman Menggunakan Data Mining
dukungan rendah (0,4%), aturan ini tetap valid dan bisa dimanfaatkan dalam konteks promosi makanan ringan.

Rule 5 terbentuk dari kombinasi *Bacineng Lucu* dengan *Milk Choco*, yang muncul dalam 0,4% transaksi, memiliki confidence 23,1% dan lift 2,815. Meski confidence relatif lebih rendah dari aturan lainnya, lift ratio yang tinggi menunjukkan korelasi kuat, yang berpotensi digunakan dalam strategi upselling produk minuman setelah pembelian camilan.

Berdasarkan hasil uji, setiap aturan yang terbentuk membuktikan adanya keterkaitan kuat antara produk awal (*premise*) dengan produk lanjutannya (*conclusion*). Indikator *lift ratio* di sini berperan dalam mengukur seberapa besar kecenderungan item-item tersebut dibeli dalam satu paket transaksi. 5 aturan yang telah dianalisis dapat dimanfaatkan dalam proses penyusunan rekomendasi untuk perusahaan.

Metode Clustering (K-Means)

Setelah proses transformasi dan normalisasi selesai, data dianalisis menggunakan algoritma K- Means dengan jumlah kluster (k) sebanyak tiga. Proses ini menghasilkan tiga cluster pelanggan berdasarkan kemiripan perilaku transaksional mereka terhadap atribut numerik, yaitu: Total_Pembelian, Total_Item, Frekuensi_Transaksi, dan Jumlah_Jenis_Menu. Proses klasterisasi ini diimplementasikan menggunakan perangkat lunak RapidMiner, yang menyediakan serangkaian operator modular untuk membangun alur pemrosesan data secara terstruktur. Berikut merupakan gambar alur klasterisasi di RapidMiner.



Gambar 8. Alur klasterisasi di RapidMiner.

Berdasarkan gambar di atas, terlihat bahwa pada proses analisis klasterisasi menggunakan algoritma *K- Means*, terdapat serangkaian tahapan perhitungan algoritma yang dilakukan dalam lingkungan *main process* di RapidMiner. Proses diawali dengan memasukkan data hasil transformasi agregatif melalui operator Read CSV. Data ini merupakan hasil pengolahan transaksi pelanggan yang telah direkap menjadi atribut numerik seperti Total_Pembelian, Total_Item, Frekuensi_Transaksi, dan Jumlah_Jenis_Menu. Selanjutnya,

operator *Select Attributes* dimasukkan ke dalam alur utama. Pada bagian *parameters*, opsi attribute filter type yang semula bernilai “all” diubah menjadi “subset”, dan dipilih atribut numerik relevan untuk analisis, dengan mengecualikan atribut identifikasi seperti No_meja yang tidak digunakan dalam perhitungan jarak. Setelah itu, operator *Normalize* dimasukkan ke dalam *main process* untuk melakukan normalisasi data. Metode yang digunakan adalah *Z-Transformation*, yang mengubah nilai numerik menjadi skor standar deviasi. Langkah ini bertujuan menyetarakan skala antar fitur numerik, sehingga tidak terjadi dominasi satu fitur tertentu dalam proses klasterisasi. Operator *K-Means* kemudian digunakan untuk menjalankan proses klasterisasi. Parameter k diatur menjadi 3, yang artinya data akan diclusterkan ke dalam tiga klaster. Selain itu, dilakukan penyesuaian pada parameter lainnya, yaitu *Max runs* = 10 dan *Max optimization steps* = 100, untuk memastikan proses iterasi berjalan optimal dan dalam mencapai pembentukan klaster. Berikut merupakan tabel hasil klasterisasi yang sudah di proses di *RapidMiner*:

Tabel 4. Hasil nilai rata-rata.

Attribute	Cluster_0	Cluster_1	Cluster_2
Total_Pembelian	-0.413	-0.475	1.715
Total_item	-0.374	-0.515	1.637
Frekuensi_Transaksi	0.570	-1.609	-0.101
Jumlah_Jenis_Menu	0.570	-1.609	-0.101

Dari hasil perhitungan pada tabel di atas diketahui nilai rata-rata (centroid) dari setiap atribut dalam masing-masing cluster. Berdasarkan atribut *Total Pembelian*, nilai rata-rata tertinggi terdapat pada cluster 2, sedangkan nilai terendah terdapat pada cluster 1. Ditinjau dari parameter Total Item, Cluster 2 memiliki nilai rata-rata yang melampaui kelompok lainnya. Namun, jika dilihat dari intensitas transaksi dan variasi menu yang dipesan, Cluster 0 tampil sebagai kelompok yang paling unggul. Adapun Cluster 1 merupakan segmen dengan tingkat aktivitas dan volume terendah dibandingkan kelompok lainnya. Hasil tersebut menunjukkan adanya perbedaan karakteristik pada setiap klaster yang terbentuk. Angka-angka tersebut masih merupakan nilai hasil normalisasi (*Z-score*), sehingga perlu dilakukan proses profiling cluster lebih lanjut guna memahami pola-pola perilaku pelanggan secara deskriptif berdasarkan cluster yang terbentuk.

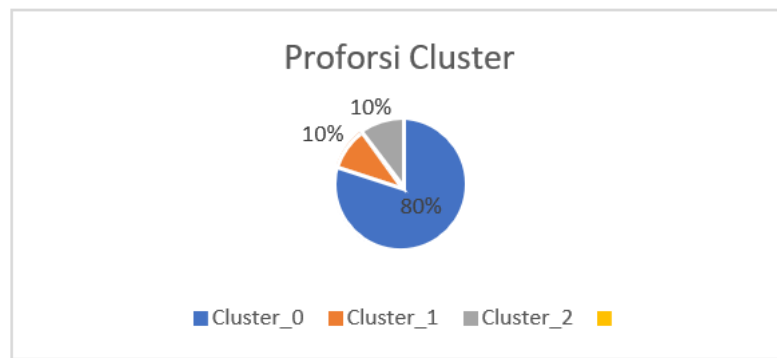
Distribusi Data pada Setiap Cluster

Untuk memahami sebaran data pada masing-masing cluster yang terbentuk, dilakukan analisis distribusi berdasarkan jumlah dan proporsi data dalam setiap cluster. Tabel berikut menunjukkan distribusi data transaksi pada setiap cluster:

Tabel 5. Distribusi data dan proporsi setiap cluster.

Cluster	Jumlah Data	Proporsi	Persentase
Cluster_0	8	0.800	80%
Cluster_1	1	0.100	10%
Cluster_2	1	0.100	10%
Total	10	1.000	100%

Berdasarkan tabel distribusi di atas dapat diketahui bahwa cluster_0 mendominasi dengan 8 data transaksi atau 80% dari keseluruhan data. Sementara itu, cluster_1 dan cluster_2 masing-masing hanya memiliki 1 data transaksi atau 10% dari total data. Jumlah total transaksi yang tercantum dalam tabel di atas bisa dihitung dalam bentuk persentase sebagai berikut:



Gambar 9. Diagram proporsi cluster.

Distribusi yang tidak seimbang ini mengindikasikan bahwa mayoritas pelanggan (80%) memiliki karakteristik perilaku yang serupa dan tergrouping dalam cluster_0. Meskipun cluster_0 memiliki nilai rata-rata yang moderat pada sebagian besar atribut, namun cluster ini merepresentasikan pola perilaku pelanggan mayoritas. Di sisi lain, cluster_1 dan cluster_2 yang masing-masing hanya memiliki 10% data, merepresentasikan segmen pelanggan dengan karakteristik unik - dimana cluster_1 cenderung memiliki nilai rendah pada hampir semua atribut, sedangkan cluster_2 menunjukkan nilai tinggi terutama pada atribut *Total Pembelian* dan *Total Item*.

Analisis ciri-ciri cluster

Dalam analisis ciri-ciri cluster, kita bisa menilai dengan mengamati hasil perhitungan rata-rata untuk setiap cluster. Langkah ini bertujuan untuk mengekstrak makna dari pengelompokan yang telah terbentuk guna mengenali kecenderungan tiap segmen. Adapun karakteristik mendetail dari setiap klaster yang dihasilkan adalah sebagai berikut:

Tabel 6. Karakteristik cluster.

Attribute	Cluster_0 Loyal Customers	Cluster_1 Occasional Customers	Cluster_2 Big Spenders
Total_Pembelian	-0.413	-0.475	1.715
Total_item	-0.374	-0.515	1.637
Frekuensi_Transaksi	0.570	-1.609	-0.101
Jumlah_Jenis_Menu	0.570	-1.609	-0.101

Merujuk pada data rata-rata nilai di tiap cluster, poin-poin kesimpulan yang dapat ditarik adalah sebagai berikut:

a. Cluster 0 - Loyal Customers

Analisis pada tabel karakteristik menunjukkan bahwa cluster pada cluster 0 memiliki nilai Total Pembelian sebesar -0.413, Total Item sebesar -0.374, Frekuensi Transaksi sebesar 0.570, dan Jumlah Jenis Menu sebesar 0.570. Pada cluster 0 keunggulan cluster ini terletak pada aspek intensitas transaksi serta keragaman menu yang dipesan yang melampaui nilai rata-rata keseluruhan, sedangkan total pembelian dan total item belum mencapai angka rata-rata keseluruhan.

Ditemukan adanya pola pada cluster 0 di mana para pembeli cenderung memiliki karakteristik melakukan pembelian dengan frekuensi tinggi dan mencoba berbagai jenis menu, namun dengan total pembelian dan jumlah item per transaksi yang relatif rendah. Segmentasi pada cluster 0 dapat diartikan sebagai konsumen loyal yang melakukan pembelian secara konsisten dimana melakukan transaksi secara terus menerus dengan variasi menu yang beragam.

Perhatian lebih perlu diberikan kepada cluster loyal customer ini, mengingat rendahnya angka pembelian per kunjungan berdampak pada pemasukan yang tidak optimal. Cluster ini pun direkomendasikan sebagai subjek utama untuk langkah-langkah perbaikan ke depan

b. Cluster 1 - Occasional Customers

Berdasarkan tabel karakteristik dapat dijabarkan bahwa cluster pada cluster 1 memiliki nilai Total Pembelian sebesar -0.475, Total Item sebesar -0.515, Frekuensi Transaksi sebesar -1.609, dan Jumlah Jenis Menu sebesar -1.609. Pada cluster 1 Secara keseluruhan, tiap atribut pada cluster ini belum mampu melampaui standar rata-rata yang ditetapkan, menunjukkan aktivitas pembelian yang sangat rendah di semua aspek.

Pada cluster 1 diketahui bahwa konsumen memiliki karakteristik melakukan pembelian dengan frekuensi sangat rendah, jumlah item sedikit per transaksi, total

Analisis Preferensi Pelanggan Semanis Coffe And Resto terhadap Kombinasi Menu Makanan dan Minuman Menggunakan Data Mining
pembelian rendah, dan variasi menu yang sangat terbatas. Klasifikasi yang terbentuk pada cluster 1 menunjukkan kecenderungan occasional customers yaitu konsumen yang tidak rutin bertransaksi secara berkala dan cenderung tidak aktif dalam berbelanja.

Langkah perbaikan difokuskan pada cluster 1 karena mencatatkan performa pembelian terendah di antara semua segmen. Untuk mengatasi hal tersebut, perusahaan perlu menjalankan program interaksi yang lebih intens dan upaya aktivasi kembali pelanggan guna mengoptimalkan hasil yang sebelumnya kurang maksimal.

c. Cluster 2 - Big Spenders

Merujuk pada tabel karakteristik dapat dijabarkan bahwa cluster pada cluster 2 memiliki nilai Total Pembelian sebesar 1.715, Total Item sebesar 1.637, Frekuensi Transaksi sebesar -0.101, dan Jumlah Jenis Menu sebesar -0.101. Cluster 2 menunjukkan performa unggul pada beberapa indikator yang nilainya berhasil melewati standar rata-rata total, yakni total pembelian dan total item, sedangkan frekuensi transaksi dan jumlah jenis menu sedikit berada di bawah ambang rata-rata total.

Pada cluster 2 menunjukkan bahwa konsumen memiliki karakteristik melakukan pembelian dengan nilai transaksi sangat tinggi dan jumlah item terbanyak per transaksi, namun dengan frekuensi transaksi yang tidak terlalu tinggi dan variasi menu yang terbatas. Berdasarkan tinjauan aktivitas transaksi, konsumen dalam cluster 2 merupakan tipe *big spenders*. Mereka memiliki kecenderungan untuk memborong produk dalam kuantitas yang paling banyak jika dikomparasikan dengan cluster lainnya

Nilai beli yang lebih besar dibandingkan cluster lain ditemukan pada cluster 2, namun hal ini tidak dibarengi dengan intensitas transaksi yang rendah sehingga perlu adanya peningkatan untuk cluster ini dalam hal frekuensi kunjungan. Cluster ini termasuk ke dalam big spenders yang memiliki potensi besar untuk dikembangkan. Penggunaan analisis clustering pada penelitian ini sudah berjalan dengan baik, karena mampu dalam mengelompokkan data konsumen berdasarkan karakteristik pembelian yang berbeda-beda. Pada penelitian ini clustering berhasil mengidentifikasi tiga segmen konsumen yang jelas dan bermakna. Capaian ini menegaskan efektivitas analisis klaster yang diterapkan, sehingga penggunaannya sangat direkomendasikan untuk membantu penyusunan strategi promosi yang lebih tertarget dan akurat.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis data transaksi penjualan di Semanis Coffee and Resto

E-ISSN : 3031-8912; P-ISSN : 3031-8904, Hal. 13-32 selama periode Maret hingga Mei 2025 menggunakan metode implementasi algoritma FP-Growth dalam *Market Basket Analysis* (MBA) serta penggunaan K-Means untuk proses *clustering* menjadi inti dari pendekatan analisis ini, penelitian ini berhasil mencapai tujuan yang diharapkan sebagaimana dinyatakan dalam bagian "Introduction".

Dari analisis Association Rule Mining menggunakan FP-Growth, penelitian ini berhasil mengidentifikasi 5 aturan asosiasi yang valid dari 672 transaksi dengan nilai lift ratio di atas 1. Aturan asosiasi terkuat ditemukan pada kombinasiAmericano→Milk Choco dengan confidence 42,9% dan lift ratio 5,229, menunjukkan korelasi positif yang signifikan antar item. Produk terpopuler adalah Milk Choco (162 kali/10,8%) danAmericano (136 kali/8,5%), mengindikasikan preferensi konsumen terhadap minuman berbasis kopi dan susu.

Analisis Customer Segmentation dengan K-Means berhasil mengidentifikasi tiga segmen pelanggan yang berbeda: Loyal Customers (80%) dengan frekuensi transaksi tinggi namun nilai pembelian rendah, Occasional Customers (10%) dengan aktivitas pembelian sangat rendah, dan Big Spenders (10%) dengan nilai transaksi tertinggi namun frekuensi kunjungan rendah. Distribusi ini menunjukkan basis pelanggan yang stabil namun memerlukan strategi khusus untuk optimalisasi revenue.

Penelitian ini memberikan kontribusi berupa wawasan mendalam tentang perilaku konsumen melalui pendekatan data mining yang komprehensif. Kombinasi antara association rule mining dan customer segmentation memberikan perspektif holistik yang dapat digunakan sebagai dasar pengambilan keputusan strategis dalam operasional bisnis.

Prospek pengembangan hasil penelitian mencakup implementasi bundling products berdasarkan aturan asosiasi, program cross-selling yang targeted, dan strategi pemasaran yang disesuaikan dengan karakteristik setiap segmen pelanggan. Penelitian selanjutnya dapat diperluas dengan periode analisis yang lebih panjang, integrasi data eksternal, dan analisis profitabilitas untuk memberikan rekomendasi yang lebih komprehensif dalam meningkatkan performa bisnis Semanis Coffee and Resto.

DAFTAR REFERENSI

- Ananda, R., & Yamani, A. Z. (2020). Penentuan centroid awal K-means pada proses clustering data evaluasi pengajaran dosen. *Jurnal*, 1(3), 544–550.
- Aqliyah, Z. H., Kurniawan, R., & Suprati, T. (2025). FP-growth algorithm for association model optimization in household sales data. *Journal of Artificial Intelligence and Engineering Applications*. Retrieved from <https://ioinformatic.org/>
- Dio, R., Hidayati, J., Arifin, R., Putera, D. A., & Dermawan, A. A. (2023). Analisis data mining pembelian dengan association rule market basket analysis menggunakan algoritma

- Fadhilah, S. A. (2022). Analisis perilaku pembelian konsumen menggunakan metode association rule–market basket analysis dan clustering analysis (Studi kasus: Jore Coffee & Eatery).
- Gohae, M., Dakhi, P., & Duha, T. (2021). Pengaruh perilaku konsumen terhadap keputusan pembelian pada UD Anisa Kecamatan Telukdalam Kabupaten Nias Selatan.
- Irawan, Y. (n.d.). Penerapan data mining untuk evaluasi data penjualan menggunakan metode clustering dan algoritma hirarki divisive.
- Komitek, J. (n.d.). Application of the K-means clustering algorithm in mapping the regional voter strategy for the legislative candidates for the DPR RI.
- Laksono, B., Syahidin, Y., & Yunengsih, Y. (2024). Implementasi data mining klusterisasi data pasien rawat inap dengan algoritma K-means clustering. *Jurnal Teknologi Sistem Informasi dan Aplikasi*, 7(2), 621–627. <https://doi.org/10.32493/jtsi.v7i2.39354>
- Minarni, P. A., Hasibuan, M. S., & Prasetya, F. D. (n.d.). Implementasi data mining dalam analisa pola peminjaman buku di perpustakaan Universitas Muhammadiyah Pringsewu menggunakan metode association rule.
- Munanda, E., & Monalisa, S. (2021). Penerapan algoritma FP-growth pada data transaksi penjualan untuk penentuan tata letak barang. *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, 7(2), 173–184.
- Nugraheni, W., & Nugroho, A. (2023). Penerapan metode market basket analysis (MBA) dengan algoritma. *Jurnal Teknologi Informasi dan Komunikasi*, 7(4). <https://doi.org/10.35870/jti>
- Padli, M., Yasir, A., & Hasugian, B. S. (2023). Computer science and information technology. *Computer Science and Information Technology*, 4(2), 182–191.
- Purwati, N., & Karnila, S. (2023). Strategi peningkatan penjualan produk menggunakan market basket analysis. *Jurnal Sistem Informasi Bisnis*, 13(2), 96–103. <https://doi.org/10.21456/vol13iss2pp96-103>
- Srikandi, A. E., & Sunarti, K. (2016). Analisis faktor-faktor preferensi pelanggan dan pengaruhnya terhadap keputusan pembelian (Studi terhadap pelanggan McDonald's di Indonesia dan Malaysia). Retrieved from <http://www.aboutmcdonalds.com>
- Sugianto, C. A., & Sukmawati, D. (2023). Penerapan algoritma FP-growth untuk mengetahui pola pada data transaksi percetakan (Studi kasus Java Printing Batujajar). *Jurnal*, 5, 20–26.
- Valencia, S., & Atmojo, W. T. (2024). Analisis pola pembelian pada data penjualan CanNgopi menggunakan algoritma FP-growth. *Jurnal Komtika (Komputasi dan Informatika)*, 8(2), 214–224. <https://doi.org/10.31603/komtika.v8i2.12672>
- Whendasgoro, R. G., & Joseph, J. (2022). Analisis penerapan normalisasi data dengan menggunakan Z-score pada kinerja algoritma K-NN. *JURIKOM (Jurnal Riset Komputer)*, 9(4), 872. <https://doi.org/10.30865/jurikom.v9i4.4526>