

INTEGRASI ALGORITMA APRIORI DAN K-MEANS DALAM ANALISIS POLA PEMBELIAN UNTUK MENINGKATKAN STRATEGI PEMASARAN

Violita Eka Putri¹⁾, Hindriyanto Dwi Purnomo²⁾

1. Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana, Indonesia
2. Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana, Indonesia

Article Info

Kata Kunci: Apriori, association rules, clustering, k-means, market basket analysis, rapidminer

Keywords: Apriori, association rules, clustering, k-means, market basket analysis, rapidminer

Article history:

Received 11 October 2024

Revised 16 November 2024

Accepted 23 December 2025

Available online 1 March 2025

DOI :

<https://doi.org/10.29100/jipi.v10i1.5768>

* Corresponding author.

Corresponding Author

E-mail address:

violitaep123@gmail.com

ABSTRAK

UMKM pada bidang usaha kuliner sedang mengalami peningkatan yang signifikan sehingga, muncul persaingan dalam dunia bisnis yang semakin tidak terelakkan. Selain itu, kebiasaan pelanggan dalam melakukan pembelian yang membutuhkan waktu lama menjadi perhatian khusus bagi pemilik bisnis Premium Salad.co untuk dapat membuat penawaran produk yang lebih sesuai dengan keinginan para pelanggan. Oleh karena itu, penelitian ini bertujuan untuk membentuk sebuah strategi pemasaran dalam bentuk rekomendasi paket menu atau dapat juga digunakan sebagai paket bundling produk dengan memperhatikan produk apa saja yang memiliki frekuensi penjualan yang sering dibeli secara bersamaan oleh pelanggan, hal ini bertujuan untuk meningkatkan daya tarik pelanggan pada saat memilih dan membeli produk, meningkatkan profit penjualan, pemerataan penjualan produk, sekaligus inovasi baru untuk mengimbangi adanya persaingan bisnis kuliner. Data transaksi yang sebelumnya tidak dimanfaatkan secara optimal oleh Premium Salad.co kini dapat dimanfaatkan untuk mencari pengetahuan lebih dalam mengenai gambaran penjualan produk yang terjadi secara keseluruhan dengan bantuan data mining. Pada penelitian ini metode data mining yang digunakan yaitu clustering dan association rules. Algoritma k-means berperan untuk mengelompokkan data dalam 4 cluster dengan nilai uji validitas Davies Bouldin Index (DBI) sebesar 0,465. Algoritma apriori berperan dalam pencarian aturan asosiasi pada cluster. Tujuan dari menggabungkan dua metode ini agar menghasilkan aturan asosiasi yang lebih variatif dan lebih sesuai dengan penyelesaian masalah yang dibutuhkan. Dengan menetapkan minimum support sebesar 0,01 dan minimum confidence sebesar 0,5. Pada cluster 0 dengan dataset 321 transaksi menghasilkan 1 rules dengan tingkat kepercayaan 57%, lalu pada cluster 1 dengan dataset sejumlah 228 mampu menghasilkan 3 rules dengan tingkat kepercayaan tertinggi sebesar 75%. Cluster 3 dengan dataset paling sedikit yaitu 127 transaksi mampu menghasilkan sejumlah 16 rules dengan tingkat kepercayaan tertinggi mencapai 100%.

ABSTRACT

UMKM in the culinary business sector are experiencing a significant increase so that competition in the business world is increasingly inevitable. Apart from that, the customer's habit of making purchases that take a long time is a special concern for the Premium Salad.co business owner to be able to make product offerings that are more in line with customers' wishes. Therefore, this research aims to form a marketing strategy in the form of menu package recommendations or can also be used as a product bundling package by paying attention to what products have a sales frequency that is often purchased simultaneously by customers. This aims to increase customer attraction. when selecting and purchasing products, increasing sales profits, equalizing product sales, as well as new innovations to balance the competition in the culinary business. Transaction data that was previously not utilized optimally by Premium Salad.co can now be used to gain deeper knowledge regarding the overall picture of product sales with the help of data mining. In this research, the data mining methods used are clustering and association rules. The k-means algorithm plays a role in grouping data into 4 clusters with a Davies Bouldin Index (DBI) validity test value of 0.465. The a priori algorithm plays a role in

searching for association rules in clusters. The aim of combining these two methods is to produce association rules that are more varied and more suitable for solving the required problem. By setting a minimum support of 0.01 and a minimum confidence of 0.5. In cluster 0 with a dataset of 321 transactions, it produces 1 rule with a confidence level of 57%, then in cluster 1 with a dataset of 228 it is able to produce 3 rules with the highest confidence level of 75%. Cluster 3 with the smallest dataset, namely 127 transactions, was able to produce 16 rules with the highest level of confidence reaching 100%.

I. PENDAHULUAN

USAHA Mikro Kecil dan Menengah (UMKM) merupakan salah satu sektor industri yang memiliki peran penting dalam mempertahankan sekaligus mendukung peningkatan perekonomian Negara Indonesia [1]. Berdasarkan data dari Kementerian Perekonomian Republik Indonesia bernomor HM.4.6/103/SET.M.EKON.3/05/2021 untuk saat ini UMKM sedang dalam trend yang cukup positif tidak hanya dalam peningkatan jumlah setiap tahunnya, tetapi kehadiran para pelaku UMKM juga telah berkontribusi terhadap Produk Domestik Bruto (PDB) sebesar 61,07% atau 8.573,89 triliun rupiah dan mampu menyerap sebanyak 97% dari total tenaga kerja [2]. Airlangga Hartanto yang merupakan Menteri Koordinator Bidang Perekonomian Indonesia mengungkapkan bahwa, bidang usaha makanan dan minuman telah banyak menyumbang pertumbuhan UMKM di Indonesia [3]. Perlu diketahui bahwasanya, Provinsi Jawa Tengah memiliki jumlah sektor industri UMKM yang mencapai angka 1.457.126. Kota Salatiga yang merupakan bagian dari Provinsi Jawa Tengah memiliki jumlah UMKM sebanyak 22.995 unit yang didominasi dalam bidang makanan dan minuman atau bisnis kuliner sekitar 80% yang dilansir dari webiste KRJogja.com [4].

Premium Salad.co merupakan salah satu sektor industri UMKM yang bergerak dalam bisnis kuliner di Kota Salatiga dan sudah berdiri sejak tahun 2018. Bisnis dalam bidang kuliner pastinya para pelaku usaha senantiasa berusaha untuk menyajikan berbagai pilihan menu makanan dan minuman semata-mata untuk memberikan daya tarik kepada para pelanggan agar membeli produknya. Sementara itu, berdasarkan hasil wawancara bersama pemilik bisnis tidak jarang pelanggan mengeluhkan waktu orderan yang lama, hal ini dikarenakan para pelanggan yang melakukan pembelian pada *offline store* kerap merasa bingung dalam memilih menu yang akan dibeli terutama bagi para pelanggan yang pertama kali berkunjung sehingga, membuat antrian pesanan yang memanjang. Adanya peningkatan UMKM yang setiap tahunnya semakin bertambah terkhusus dalam dunia bisnis kuliner memberikan sedikit penekanan kepada pemilik bisnis untuk mencari strategi yang tepat agar dapat mempertahankan keberlangsungan bisnis dan para pelanggan setianya [5]. Salah satu strategi pemasaran penting yang harus dilakukan pada suatu tempat bisnis kuliner yaitu dikenal dengan istilah *cross-selling*, biasanya kegiatan pemasaran dengan menerapkan *cross selling* ini pemilik bisnis akan menawarkan produk tambahan kepada para pelanggan dan pastinya produk yang ditawarkan merupakan produk yang relevan dan sesuai dengan apa yang sudah pernah dibeli oleh pelanggan [6]. Akan tetapi, seringkali suatu tempat bisnis melakukan pemasaran produknya tidak sesuai dengan apa yang dibutuhkan pelanggan, hal ini dikarenakan lemahnya pelaku bisnis dalam menentukan strategi pemasaran secara tepat dan akurat [7].

Berdasarkan permasalahan mengenai strategi pemasaran, sebenarnya dapat ditangani dengan memanfaatkan riwayat data transaksi yang merupakan salah satu aset penting bagi suatu tempat bisnis karena, selama kegiatan operasional berjalan maka akan semakin banyak data transaksional yang disimpan [8]. Namun, seringkali terjadi pada beberapa tempat bisnis tidak memanfaatkan kumpulan data tersebut dan hanya menjadikan riwayat data transaksi sebagai arsip penjualan yang tidak berguna, seperti yang terjadi pada Premium Salad.co. Dengan memanfaatkan data transaksional dapat dilakukan penggalian informasi yang mendalam untuk mengidentifikasi fenomena penjualan yang terjadi pada suatu tempat bisnis sekaligus memperoleh informasi yang dapat dimanfaatkan bagi pemilik bisnis [9].

Oleh karena itu, adanya penelitian ini untuk memberikan kontribusi pada Premium Salad.co dalam memecahkan permasalahannya mengenai pembentukan strategi pemasaran. Jika memperhatikan penggunaan dataset yang dikumpulkan dalam penelitian ini cukup besar maka, untuk meningkatkan hasil analisis data yang lebih akurat dan relevan penggunaan metode data mining (Knowledge Discovery in Database (KDD)) ini sangat membantu pada proses identifikasi pola pembelian pelanggan sekaligus mendeteksi hubungan atau korelasi antar item yang terindikasi memiliki frekuensi pembelian yang sering dilakukan secara bersamaan oleh pelanggan dalam satu waktu [10]. Dengan memanfaatkan teknik data mining dapat dilakukan eksplorasi informasi yang tersembunyi secara lebih kompleks serta dapat menggunakan data yang berukuran besar [6]. Hasil dari proses pengolahan data transaksi ini mampu menghasilkan pengetahuan atau insight baru yang tidak dapat diketahui secara manual oleh

manusia sehingga, dapat digunakan sebagai bahan pendukung pengambilan keputusan bisnis yang berbasis data besar untuk meningkatkan efektivitas strategi pemasaran [11]. Konsep dari market basket analysis ini didasarkan pada teori bahwasannya, jika terdapat seseorang yang membeli item tertentu memiliki kemungkinan besar membeli item yang lainnya juga [12].

Dalam melakukan analisis data transaksi ini menerapkan dua algoritma yang berbeda yaitu dengan menggunakan algoritma k-means dan algoritma apriori. Salah satu pendekatan kuantitatif dalam melakukan penambangan data secara mendalam dan menyeluruh menggunakan data mining yaitu melalui pengimplementasian algoritma apriori. Pada dasarnya algoritma apriori sudah terbukti memiliki kinerja yang baik dalam menangani permasalahan terkait market basket analysis, tetapi terdapat kelemahan pada penggunaan dataset yang berukuran besar pada proses iterasi pencarian aturan asosiasi. Penurunan performa yang sering terjadi pada saat proses pencarian aturan asosiasi adalah apabila terdapat suatu kumpulan item besar dalam data transaksi yang cenderung diabaikan pada saat proses iterasi untuk menemukan frequent itemset dan informasi yang diperlukan menjadi tidak tersedia dalam hal ini yaitu rules atau aturan asosiasi, hal ini dikawatirkan dapat membuat aturan asosiasi yang dihasilkan tidak akurat untuk dijadikan sebagai rekomendasi bisnis dalam mengatur strategi pemasaran [13] [14]. Oleh karena itu, pada penelitian ini juga menggabungkan algoritma k-means clustering yang berperan dalam melakukan pengelompokan data secara keseluruhan kedalam beberapa cluster berdasarkan kemiripan karakteristiknya dengan menggunakan perhitungan jarak sebagai pembanding setiap objek dari cluster satu dengan yang lainnya [15]. Sedangkan, algoritma apriori berperan dalam pembentukan aturan asosiasi pada cluster yang sudah terbentuk pada proses sebelumnya [16]. Dengan menggabungkan kedua metode ini bertujuan untuk mempersempit ruang pencarian pembentukan aturan asosiasi sehingga, rules yang dihasilkan menjadi lebih tepat dan akurat dengan mempertimbangkan ambang batas minimum nilai support dan confidence [8].

Peneliti sudah melakukan kajian literatur bersumber dari penelitian terdahulu yang dipakai sebagai landasan penelitian sekaligus pemilihan algoritma yang digunakan untuk proses pengolahan market basket analysis pernah yang dilakukan oleh Sania Fitri Octavia, Mustakim, Inggih Permana, dan Siti Monalisa dengan judul penelitian "Penerapan Algoritma Association Rules Dalam Penentuan Pola Pembelian Berdasarkan Hasil Clustering" berhasil menyusun strategi penjualan dengan mengatur penataan barang pada toko sekaligus mengelola stok barang yang bertujuan untuk meningkatkan pelayanan kepada para pelanggan. Penelitian ini melakukan perbandingan hasil dari pencarian aturan asosiasi melalui algoritma k-medoid dan apriori, k-medoid dan fp-growth, serta pencarian aturan asosiasi yang hanya menggunakan satu algoritma association rules. Dan hasil yang didapatkan dengan menggabungkan 2 metode clustering dan association rules dapat menghasilkan sejumlah aturan asosiasi yang bervariatif, sedangkan jika hanya menggunakan metode association rules tidak dapat menghasilkan rules sama sekali meskipun nilai support dan nilai confidence yang diterapkan sama. Jika dilihat dari uji akurasi pada penelitian ini menggabungkan algoritma k-medoid dengan fp-growth menghasilkan nilai akurasi mencapai 120% lebih tinggi dari k-medoid dan apriori [13]. Penelitian lainnya yang sejalan literatur sebelumnya pernah dilakukan oleh Fitri Nuraeni, Dewi Tresnawati, Yoga Handoko Agustin, dan Gisna Fauzian Dermawan dalam judul "Optimization Of Market Basket Analysis Using Centroid-Based Clustering Algorithm And FP-Growth Algorithm" yang berhasil memberikan solusi atas permasalahan bisnis terkait pembentukan paket menu untuk inovasi pemasaran baru, untuk membantu pelanggan mempercepat proses orderan, sekaligus mengatur stok produksi. Dengan menggunakan metode DBI dan Silhouette index dari beberapa percobaan pada algoritma clustering (fuzzy c-means, k-means, k-medoids) dipilih algoritma k-means untuk proses pengelompokan data karena hasil aturan asosiasi yang terbentuk menunjukkan persentase nilai lift ratio lebih dari satu yang menunjukkan algoritma k-means memiliki kemampuan dalam mengelompokkan data dengan tingkat keimipiran yang tinggi untuk setiap cluster nya. Dari hasil penelitian yang didapatkan masih dibutuhkan pengembangan mengenai penggunaan algoritma clustering yang bersifat soft-clustering untuk memperbaiki proses pengelompokan [9].

Penelitian mengenai market basket analysis yang menerapkan dua metode clustering dan association rule sudah pernah dilakukan oleh Natalia Mamahit dan Anita Qoriah dengan judul penelitian "Penerapan Algoritma FP-Growth dan K-Means pada Data Transaksi Minimarket" yang berhasil memberikan solusi atas permasalahan mengenai paket produk yang tidak tersedia pada minimarket dan penataan barang yang masih berantakan. Dengan menggunakan 312 data penjualan didapatkan 10 cluster yang masing-masing ditetapkan nilai minimum support sebesar 20% dan minimum confidence 50%. Dari proses pencarian aturan asosiasi mampu menghasilkan rekomendasi paket produk bundle sebanyak 21 dan melakukan penataan pada 10 rak minimarket. Pada penelitian ini didapatkan bahwa dengan menggabungkan dua metode mampu membentuk rules lebih banyak daripada menggunakan satu metode association rules saja. Hasil penelitian ini tidak memberikan persentase hasil nilai lift ratio ataupun tingkat kepercayaan terhadap rules yang terbentuk untuk memberikan gambaran kepada para pembaca seberapa besar tingkat kevalidan rules yang terbentuk [17]. Penelitian lainnya yang sejalan dengan literatur sebelumnya pernah dilakukan oleh N P Dharshinni, Elvina Bangun, Sarah Karunia, Ruth Damayanti, Gabriel Rophe, Roy Pandapotan dengan judul penelitian "Menu Package Recommendation using Combination of

K-Means and FP-Growth Algorithms at Bakery Stores” yang berhasil membuktikan bahwa menerapkan dua metode clustering dan association rule pada 6017 data penjualannya mampu membentuk rekomendasi paket menu yang akurat dan relevan, hal ini dapat dilihat dari 3 cluster yang terbentuk dilakukan pencarian aturan asosiasi dengan minimum support dan confidence sama pada masing-masing cluster yaitu 20% dan 60%. Hasil paket rekomendai menu dapat dikatakan valid dengan dukungan persentase nilai lift ratio lebih dari satu dan tingkat kepercayaan pada rentang 79%-92% [18].

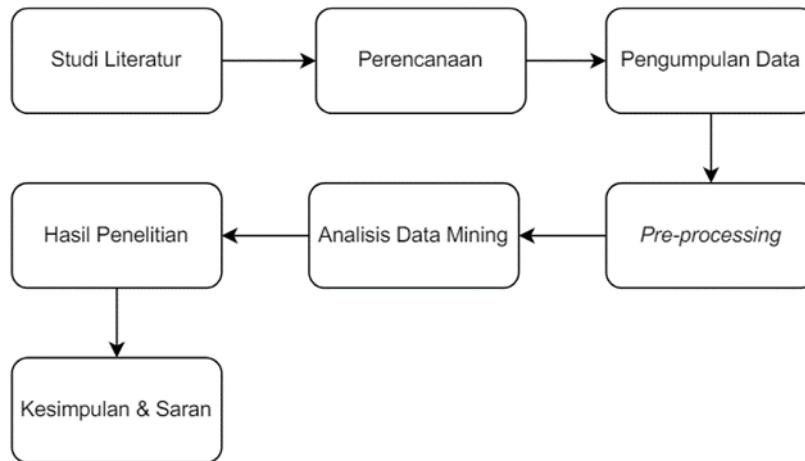
Beberapa kajian literatur yang menerapkan metode clustering terlebih dahulu, lalu disusul dengan menerapkan association rule juga sudah pernah dilakukan pada penelitian yang berjudul “Analisis Keranjang Pasar Menggunakan Algoritma K-Means dan FP-Growth pada PT.Citra Mustika Pandawa” oleh Sindhy Genjang Setyorini, Ega Kurnia Sari, Lutfy Rahma Elita, Shinta Ayunda Putri yang berhasil mengelompokkan data menjadi lima cluster dengan menetapkan minimum support yang pertama sebesar 30%, minimum support yang kedua sebesar 50%, dan untuk minimum confidence masing-masing ditetapkan sebesar 50%. Dari hasil uji pencarian aturan asosiasi pada setiap cluster tidak menjamin dapat membentuk rules dengan jumlah dan kombinasi yang sama pada masing-masing cluster, bahkan jika peneliti menetapkan nilai minimum support yang tinggi yaitu sebesar 50% terdapat cluster yang tidak menghasilkan rules sama sekali [19]. Hal ini sejalan dengan penelitian yang dilakukan oleh Imaduddin Syukra, Assa Hidayat, dan Muhammad Zakiy Fauzi dengan judul penelitian “Implementation of K-Medoids and FP-Growth Algorithms for Grouping and Product Offering Recommendation” yang berhasil membentuk tiga cluster dimana, pada masing-masing cluster ditetapkan minimum support pada range 3% - 5% dan minimum confidence ditetapkan sebesar 50%. Sementara itu, hasil uji pencarian aturan asosiasi pada penelitian ini hanya berhasil membentuk rules pada cluster 3 saja, tetapi jika peneliti menetapkan nilai minimum support sebesar 9% maka pada cluster 3 ini tidak terbentuk rules sama sekali [20]. Keenam sumber literatur yang digunakan sebagai landasan penelitian ini memiliki tujuan yang sama dalam melakukan identifikasi terhadap pola pembelian pelanggan dengan mengimplementasikan teknik clustering terlebih dahulu, lalu menerapkan association rules untuk mencapai satu tujuan yaitu memberikan kontribusi positif mengenai strategi penjualan bisnis seperti membentuk rekomendasi paket, penataan barang, hingga pengelolaan stok barang. Berdasarkan kajian literatur yang sudah dilakukan dengan menggabungkan dua metode ini dapat menghasilkan aturan asosiasi yang lebih bervariatif, akurat, dan lebih relevan dengan berbasis database transaksional berukuran besar.

Berdasarkan hasil dari pemaparan diatas, dengan menggabungkan dua metode clustering dan association rules pada suatu bisnis UMKM di bidang kuliner diharapkan dapat memberikan kontribusi secara positif dalam membentuk strategi pemasaran terkhusus pada Premium Salad.co sehingga, mampu memberikan penawaran yang menarik kepada para pelanggan untuk membeli produknya. Fokus pada penelitian ini adalah pengolahan riwayat data transaksi dengan menggunakan teknologi untuk membantu dalam menentukan strategi pemasaran. Sementara itu, jika melihat data transaksi yang dimiliki berjumlah cukup besar dengan tipe data atribut yang bervariatif pula maka algoritma k-means ini berperan dalam mengelompokkan data kedalam cluster secara homogen berdasarkan kemiripan karakteristik setiap atributnya. Hasil dari proses clustering sebelumnya dapat mempermudah peneliti dalam mengidentifikasi pola belanja pelanggan secara spesifik [21]. Cluster yang sudah terbentuk ini dapat memberikan informasi mengenai distribusi penjualan secara keseluruhan sehingga, pada tahap ini peneliti dapat menentukan langkah penyelesaian masalah secara tepat sesuai dengan kebutuhan atau tujuan dari penelitian ini. Selanjutnya, pada proses pencarian aturan asosiasi menggunakan apriori untuk penelitian ini difokuskan pada cluster yang terindikasi membutuhkan solusi permasalahan. Sehingga, apabila terdapat cluster yang teridentifikasi memiliki frekuensi penjualan rendah baik secara pendapatan ataupun frekuensi jumlah penjualan produknya maka, ditetapkan pada cluster yang terpilih tersebut untuk diberikan solusi permasalahan dalam bentuk kombinasi item yang dapat dijadikan sebagai rekomendasi paket menu ataupun paket *bundling* sebagai salah satu langkah dalam melakukan strategi pemasaran dengan syarat kombinasi item yang terbentuk harus menenuhi ambang batas nilai minimum support dan minimum support yang sudah ditetapkan, serta persentase nilai lift ratio mampu menunjukkan angka nilai sebagai alat ukur kuat atau tidaknya kombinasi produk yang terbentuk. Penentuan nilai ambang batas dan syarat nilai lift ratio diberlakukan dengan tujuan untuk memberikan kombinasi produk yang valid agar dapat digunakan oleh pemilik bisnis.

II. METODE PENELITIAN

2.1. Tahapan Penelitian

Penelitian kali ini melalui dua proses yang pertama yaitu, analisis clustering menggunakan algoritma K-Means dan proses analisis kedua yaitu analisis association rules menggunakan algoritma apriori dengan tujuan akhir yaitu pembentukan rules. Rancangan penelitian dapat dilihat seperti pada gambar dibawah ini :



Gambar 1. Alur Penelitian

Tahap penelitian dapat dijelaskan sebagai berikut :

a. Studi Literatur

Studi literatur kali ini peneliti menggunakan beberapa bahan literatur yang bersumber dari jurnal dan artikel sebagai referensi dalam menguatkan topik penelitian dan metode yang digunakan. Berdasarkan dari hasil studi literatur, penelitian ini akan mengintegrasikan dua metode data mining yaitu, *clustering* dan *association rules* dengan memakai *tools rapidminer* sebagai alat pengolahan data.

b. Perencanaan

Pada tahap perencanaan peneliti melakukan identifikasi masalah secara luas sebagai topik dari penelitian ini yang kemudian dijabarkan secara spesifik pada rumusan masalah sehingga, terbentuklah poin-poin permasalahan untuk dicari solusinya.yaitu, bagaimana cara membentuk rekomendasi menu atau membentuk produk *bundling* yang tepat. Berangkat dari rumusan masalah yang ada ditetapkan tujuan penelitian ini dan batasan dari penelitian ini sehingga, proses penelitian dapat berjalan secara sistematis dan dapat dituangkan dalam bentuk desain penelitian sebagai patokan selama proses penyelesaian penelitian. Pada tahap ini menjadi proses penentuan dalam pemilihan metode dan algoritma yang akan digunakan.

c. Pengumpulan Data

Pengumpulan data transaksi penjualan didapatkan dari sistem Go-Biz yang dipakai oleh Premium Salad.co melalui proses unduh data mentah dalam format csv sebanyak 1386 data *record* terhitung sejak bulan Juni 2023 hingga Desember 2023.

d. Pre-processing

Pada tahap persiapan data diperlukan beberapa proses yang harus dilakukan pada data mentah yang digunakan untuk penelitian ini. Pertama, seleksi data yang bertujuan untuk memilih atribut yang akan digunakan selama pengolahan data dimana pada masing-masing algoritma memiliki kebutuhan penggunaan atribut yang berbeda pula. Kedua, pembersihan data untuk menghapus beberapa data yang kurang relevan dan data kosong sesuai dengan apa yang ada pada data mentah. Ketiga, transformasi data untuk mengubah format skala pengukuran data menyesuaikan dengan algoritma pengolahan data yang digunakan. Semua rangkaian proses ini bertujuan untuk menyiapkan data siap pakai sebelum mengimplementasikan algoritma clustering dan association rules lebih lanjut.

1. Seleksi Data. Pada proses ini perlu adanya pemahaman mengenai data hingga pemilihan atribut yang akan digunakan selama proses pengolahan data karena, data mentah yang ada tidak semuanya cocok untuk digunakan dalam analisis data mining, sehingga disesuaikan dengan tujuan dari proses analisis ini. Dengan bantuan *python* menggunakan fungsi '*drop()*' untuk menghapus kolom atribut yang tidak digunakan pada masing-masing algoritma kemudian disimpan pada dua file yang berbeda.

2. Pembersihan Data. Proses ini berfungsi untuk melakukan pembuangan terhadap beberapa data yang kurang relevan, data null, ataupun data duplikat sebelum dilakukan analisis data mining. Hanya saja pada dataset ini tidak memiliki data duplikat, sehingga data kosong dan data yang kurang relevan saja yang dihapus dengan menerapkan *conditional formatting* yang ada pada *microsoft excel*.

3. Transformasi Data. Pada tahap ini bertujuan untuk mengubah skala pengukuran data ke dalam format yang sesuai dengan metode analisis yang digunakan dalam penelitian ini. Pada proses *clustering* semua value pada setiap atribut yang memiliki tipe data polynominal diubah ke tipe data integer dengan

bantuan fungsi yang ada pada *microsoft excel*, sedangkan pada proses *association rules* setelah dilakukan inisialisasi jumlah produk per transaksi kemudian dengan bantuan *python* menggunakan fungsi library *mlxtend*, *value* pada dataset akan diubah menjadi bilangan biner 0 dan 1.

e. Analisis Data Mining

Analisa data mining merupakan proses utama dari penyelesaian masalah atau pencarian solusi terbaik untuk penelitian. Peneliti menerapkan konsep market basket analysis dengan mengintegrasikan dua metode yaitu k-means *clustering* dan *apriori*. Secara singkatnya, dataset akan dikelompokkan terlebih dahulu berdasarkan kemiripan karakteristik atribut dengan menerapkan metode DBI sebagai uji validitas *cluster* ideal yang terbentuk. Kemudian, dalam melakukan pencarian aturan asosiasi menggunakan *apriori* hanya ditujukan kepada *cluster-cluster* yang membutuhkan perbaikan permasalahan berdasarkan hasil analisa *descriptive statistics* yang dilakukan pada proses *clustering* sebelumnya.

f. Hasil Penelitian

Hasil penelitian ini menjabarkan tentang hasil dari penyelesaian masalah terkait data transaksi menggunakan data mining. Hasil penelitian dimuat oleh peneliti kedalam format laporan, sebagai bentuk dokumentasi penelitian dan bisa diberikan kepada pemilik bisnis jika lau ingin menggunakannya.

g. Kesimpulan dan Saran

Kesimpulan dan saran berisi review singkat terkait penelitian yang sudah selesai dilakukan dengan menarik kesimpulan sekaligus saran yang berdasarkan fakta pada saat proses penelitian dilakukan hingga hasil akhir penelitian didapatkan.

2.2. Market Basket Analysis

Analisis keranjang belanja atau market basket analysis merupakan metode data mining yang berfungsi dalam pencarian pola pembelian pelanggan dengan mengidentifikasi item-item yang sering dibeli secara bersamaan dan berulang dalam satu waktu transaksi pada *shopping basket* [22].

$$\{\text{Teh, nasi}\} \longrightarrow \{\text{ayam}\} \quad (\text{support} = 40\%, \text{confidence} = 50\%) \quad (1)$$

Misalkan, seorang pelanggan membeli teh dan nasi memiliki kemungkinan sebesar 50% untuk membeli ayam, hal tersebut didasarkan pada 40% dari data transaksi memuat tiga item tersebut. Pengolahan data transaksi menggunakan metode ini mampu menghasilkan informasi yang dapat memberikan dampak positif bagi pemilik bisnis dalam hal penguatan strategi pemasaran seperti, memanfaatkan informasi mengenai produk yang paling banyak diminati, personalisasi pemasaran (rekomendasi produk, promo *bundling*), dan optimisasi stok produk [9]. Market basket analysis sering disebut dengan metode *association rules* dimana, metode ini berfungsi untuk menentukan kombinasi item atau aturan asosiasi. Pencarian aturan asosiasi juga harus memperhatikan parameter support (converage) dan confidence (accuration) [23].

Rumus perhitungan nilai support item :

$$\text{Support (A)} = \frac{\sum \text{transaksi A}}{\sum \text{transaksi}} \times 100\% \quad (2)$$

Rumus perhitungan nilai support kombinasi item :

$$\text{Support (A} \cap \text{B)} = \frac{\sum \text{transaksi A} \cap \text{B}}{\sum \text{transaksi}} \times 100\% \quad (3)$$

Rumus perhitungan nilai confidence :

$$\text{Support (A | B)} = \frac{\sum \text{transaksi A}}{\sum \text{transaksi B}} \times 100\% \quad (4)$$

Nilai penunjang (support) merupakan persentase kejadian dari item (itemset) berdasarkan jumlah transaksi [24], sedangkan nilai kepercayaan (confidence) berfungsi dalam mengukur seberapa kuat korelasi atau hubungan yang terjadi dari setiap item dalam sebuah rules [12]. Dalam membentuk *rules* atau aturan asosiasi terdapat teknik lift ratio yang berfungsi untuk mengukur kevalidan sebuah aturan asosiasi. Jika lift ratio bernilai lebih dari satu hal ini menunjukkan bahwasannya *rules* tersebut dapat dikatakan valid, maka semakin tinggi nilai lift ratio kekuatan aturan asosiasi semakin besar pula [9]. Namun, jika nilai lift ratio kurang dari satu, dapat dikatakan *rules* yang terbentuk kurang valid untuk dapat digunakan dalam membentuk strategi pemasaran [25]. Berikut ini cara perhitungan nilai lift ratio :

$$\text{Lift ratio} = \frac{\text{Support (A} \cap \text{B)}}{\text{Support (A) Support (B)}} \times 100\% \quad (5)$$

2.3. Algoritma Apriori

Algoritma apriori menjadi salah satu algoritma yang sering dipakai dalam pencarian aturan asosiasi. Algoritma ini bekerja dengan membentuk pola *frequent itemset* tinggi pada iterasi data [24]. Iterasi data pada algoritma ini bekerja dengan melalui dua tahap yaitu, membangkitkan kandidat (*candidate generation*) dengan

memperhatikan minimum *support*. Kemudian, kandidat yang terpilih disandingkan dengan *antecedent* untuk mendapatkan nilai *confidence* dan membentuk aturan asosiasi dengan syarat nilai *confidence* harus lebih dari minimum *confidence* yang sudah ditentukan sebelumnya [25]. Berikut langkah pencarian aturan asosiasi :

1. Menentukan nilai minimum support
2. Membangkitkan kandidat mulai dari 1-itemset hingga k-itemset dengan memperhatikan minimum support. Proses iterasi akan berlanjut hingga k-itemset sudah tidak terbentuk lagi
3. Menetapkan nilai confidence dari *final itemset* yang sudah terbentuk

2.4. Algoritma K-Means

K-means merupakan salah satu teknik dalam *clustering* yang berfungsi untuk mempartisi objek kedalam sejumlah *cluster* berdasarkan kedekatan jarak atau kemiripan atribut [26]. Untuk mengetahui jarak kedekatan objek dalam suatu data terhadap titik pusat dapat menggunakan rumus perhitungan *euclidean distance*, sebagai berikut [18] :

$$D(i,j) = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ni} - x_{nj})^2} \quad (6)$$

Algoritma ini terkenal sederhana dan cepat dalam melakukan pengolahan data [19], serta mampu menangani permasalahan pengolahan data seperti tipe atribut yang berbeda, data *noise* dan dapat diterjemahkan dengan mudah [23]. Adapun alur pengolahan data menggunakan algoritma k-means, sebagai berikut [27] :

1. Menetapkan nilai k cluster
2. Menghitung jarak *centroid* (titik pusat) pada setiap *cluster* menggunakan rumus *euclidean distance*
3. Menetapkan item ke setiap *cluster* dengan memperhatikan kedekatan jarak objek dengan *centroid*
4. Proses sebelumnya akan terus melakukan iterasi, hingga objek sudah tidak berpindah *cluster* lagi.

III. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Objek penelitian market basket analysis ini dilakukan pada salah satu bisnis UMKM di Kota Salatiga yaitu Premium Salad.co. Pada tahap pengumpulan data, peneliti mengunduh riwayat data transaksi dari sistem Go-Biz yang terhitung dari bulan Juni 2023 hingga Desember 2023 dan untuk data yang terkumpul terdiri dari 1386 transaksi, 12 kategori menu dengan 60 produk. Berikut ini contoh data mentah transaksi dan tabel produk yang dijual pada setiap kategori menu :

	Order_ID	Detail_Item	Transaction_ID	Amount	Transaction time	Payment Type	Payment Date	GO-PAY Transactions ID	Customer ID
0	'F-2235073894	Veggie Salad, Alpukat Juice	'0802b39b-7884-319b-ac14-220bc40c2672	50500	2023-09-30T14:23:18.000+07:00	GO-PAY	20230930	0520230930072318iGh02mZNnZID	NaN
1	'F-2235040306	The Bowl Teriyaki (Free Egg), The Bowl Teriyak...	'1864ad8b-66b2-3ffe-b80e-872cf3d6b035	100500	2023-09-30T13:40:12.000+07:00	GO-PAY	20230930	0520230930064012LVqt9Er0nID	NaN
2	'F-2234970199	Bento With Teriyaki, Salad Buah 500ml	'2e598a7e-58d8-30b6-b9f8-5777e73c3207	82000	2023-09-30T12:57:37.000+07:00	GO-PAY	20230930	0220230930055737p7Jdx9S8zyID	NaN
3	'F-2234882698	Salad Buah 500ml, Chicken Steak Blackpepper	'5ae95d1b-7950-3bcc-afd8-1fcfa0fecc97c	70500	2023-09-30T12:07:04.000+07:00	GO-PAY	20230930	0220230930050704aMxHUIWWuXID	NaN

Gambar 2. Data Mentah Transaksi Premium Salad.co

TABEL I
KATEGORI PRODUK

Departemen	Kategori	Produk
Departemen 1	Conshake (Coconut Shake)	Orginal Conshake, Berry Conshake, Avocado Conshake, Dragon Fruit Conshake, Mango Conshake
Departemen 2	Fruit Salad	Salad Buah 200ml / Xs, Salad Buah 350ml, Salad Buah 500ml, Salad Buah 650ml, Salad Buah 750ml, Salad Buah 1000ml
Departemen 3	Super Premium Fruit Salad	Super Premium Fruit Salad 200ml, Super Premium Fruit Salad 350ml, Super Premium Fruit Salad 500ml, Super Premium Fruit Salad 650ml, Super Premium Fruit Salad 1000ml

Departemen	Kategori	Produk
Departemen 4	Steak	Chicken Steak Blackpepper, Mix Platter, French Fries, Chicken Steak Crispy Blackpepper
Departemen 5	Vegetable Salad	Veggie Salad, Chicken Veggie Salad, Egg Veggie Salad, Premium Sausage Veggie Salad, Beef Veggie Salad, Tuna Veggie Salad, Prawn Veggie Salad
Departemen 6	Wrap Salad	Egg Salad Wrap, Chicken Salad Wrap, Saussage Salad Wrap, Beef Salad Wrap, Tuna Salad Wrap
Departemen 7	Sandwich	Chicken Sandwich, Beef Sandwich, Egg Sandwich, Chicken Katsu Sandwich Bento Matah, Bento Blackpepper, Bento With Teriyaki, The Bowl Matah (Free Egg), The Bowl Teriyaki (Free Egg), The Bowl Blackpepper (Free Egg), Beef Bimbimbap, Pecel
Departemen 8	Healthy Foodies	Hydro Power, Green Energy Full, Antioxidant Fusion, Immune Booster, For More Energy
Departemen 9	Hydro Juice Cup	Melon Juice, Apel Juice, Alpukat Juice, Semangka Juice, Banana Juice, Naga Juice, Mango Juice
Departemen 10	Original Juice	Berry Smoothies Bowl 450ml, Avocado Smoothies Bowl 450ml, Mango Smoothies Bowl 450ml
Departemen 11	Smoothies Bowl	Banana Penyet, Banana Crispy
Departemen 12	Ban Ban Banana	

3.2. Pre-processing Data

Tahap *pre-processing* data perlu dilakukan sebelum masuk dalam pengolahan data mining, tahap ini bertujuan untuk mempersiapkan data yang sudah memenuhi kriteria analisa data mining. Pada penelitian ini memiliki perbedaan pada proses *pre-processing* menyesuaikan dengan kebutuhan pada metode *clustering* dan *association rules*. Berikut ini hasil dari proses *pre-processing* data pada masing-masing metode :

a. Seleksi Data

Pada proses analisis data mining menggunakan *clustering*, dari 9 kolom yang ada pada data mentah transaksi penjualan dipilih 4 kolom yang akan digunakan terlihat pada tabel 2 di bawah ini :

TABEL II
ATRIBUT DATA PADA PROSES ANALISA CLUSTERING

Atribut	Keterangan
Order_ID	No struk setiap transaksi
Detail_Item	Nama produk setiap transaksi
Amount	Total harga transaksi
Transaction time	Kategori produk

Pada proses analisis data mining menggunakan *association rules*, berikut ini atribut data yang akan digunakan :

TABEL III
ATRIBUT DATA PADA PROSES ASSOCIATION RULES

Atribut	Keterangan
Detail_Item	Nama produk setiap transaksi
Order ID	No struk setiap transaksi

b. Pembersihan Data

Dataset pada proses *clustering*, terdapat beberapa data teridentifikasi kurang relevan dan data kosong, sehingga dataset awal yang terdiri dari 1386 transaksi berkurang menjadi 1324 transaksi. Berikut sampel data yang tidak relevan dan data kosong :

TABEL IV
ATRIBUT DATA PADA PROSES ASSOCIATION RULES

Order ID	Detail Item
'c6394aad-5916-11ee-8367-db69d567488f-C81BC	1111111-1111-0111-0111-111111111111Pesanan manual
'FP-14016948	Null

c. Transformasi Data

Transformasi data dilakukan untuk mengubah format data sesuai dengan metode penelitian yang digunakan. Transformasi data untuk analisis *clustering* dan association rule memiliki perbedaan format, yang dapat dilihat pada gambar dibawah ini :

	Order ID	Jumlah Produk	Amount	Kategori Produk	Jam
0	'F-2344569473	3	91000	5	18
1	'F-2344551240	1	35500	2	18
2	'F-2344489277	1	52500	5	17
3	'F-2344486862	2	71000	6	17
4	'F-2344276412	1	35500	2	13

Gambar 3. Sampel Dataset Hasil Transformasi untuk Proses Clustering

Pada gambar 2 merupakan hasil transformasi data untuk proses *clustering*. Terdapat perbedaan pada format atribut jumlah produk merupakan hasil dari jumlah produk yang dibeli dari setiap transaksi. Lalu, pada atribut kategori produk didapatkan dari median atau nilai tengah dari setiap transaksi per departemen. Dan atribut terakhir yang berubah adalah atribut jam yang menunjukkan *transaction time* yang digunakan yaitu satuan angka per jam.

Detail_Item	Alpukat	Alpukat	Antioxidant	Antioxidant	Apel	Apel	Avocado	Avocado	Smoothies	Banana	... Blackpepper (Free Egg)
	Juice	Juice	Fusion	Fusion	Juice	Conshake	Conshake		Bowl 450ml	Crispy	
Order_ID											
'F-2119372101	0	0	0	0	0	0	0	0	0	0	0
'F-2119389128	0	0	0	0	0	0	0	0	0	0	0
'F-2119599185	0	0	0	0	0	0	0	0	0	0	0
'F-2119981399	0	0	0	0	0	0	0	0	0	0	0

Gambar 4. Sampel Dataset Hasil Encoding untuk Proses Association Rules

Tabel transaksi yang digunakan untuk analisis *association rules* dilakukan inisialisasi produk apa saja yang dibeli pada setiap ID transaksi untuk kemudian masuk dalam tahap encoding data untuk mengubah jumlah produk yang dibeli dari setiap transaksi menjadi bilangan biner atau nol dan satu. Hal ini bertujuan untuk memudahkan dalam pembentukan *frequent itemsets* untuk mencari *rules* terbaik.

3.3. Analisis Data Mining

Penyelesaian permasalahan pada penelitian ini dilakukan dengan mengintegrasikan dua metode dari data mining yaitu, *clustering* dan *association rules*. Sebelum dilakukan proses analisis lebih dalam perlu adanya pemahaman mengenai gambaran dari seluruh data transaksi sehingga, memudahkan untuk mencari informasi penting sesuai dengan kebutuhan atau tujuan dari penelitian ini.

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
JumlahProduk	1324	1	18	1.82	1.165
Amount	1324	17500	375000	61922.58	35275.067
KategoriProduk	1324	1	12	4.97	2.770
Jam	1324	7	21	14.00	3.377
Valid N (listwise)	1324				

Gambar 5. Hasil Analisa Deskripsi Statistik

Berdasarkan pada gambar diatas dapat disimpulkan sebagai gambaran mengenai data transaksi yang khususnya terjadi selama tujuh bulan pada Premium Salad.co melalui proses *descriptive statistics* dengan bantuan *tools* SPSS. Dapat diketahui para pelanggan mampu melakukan pembelian dengan jumlah produk maksimal 18 dan minimal pembelian produk 1 dengan rata-rata sebesar 1.821. Untuk *amount* atau total harga transaksi mampu menembus maksimal hingga Rp.375.000,00 dan minimal total transaksi Rp.17.500,00 dengan rata-rata sebesar Rp.61.922,583. Selanjutnya, untuk kategori produk yang dibeli maksimal pada departemen 12 dan minimal pada departemen 1 dengan rata-rata kategori produk 4.967. Dan pada jam atau waktu order pelanggan maksimal terjadi pada pukul 21.00 WIB dan minimal waktu order pada pukul 07.00 WIB dengan rata-rata waktu order pada pukul 14.00 WIB. Gambaran mengenai data transaksi ini dapat dinyatakan valid yang dapat dilihat dari jumlah Valid N (*listwise*) sebesar 1324 data transaksi. Dari hasil analisa deskripsi statistik pada dataset transaksi secara keseluruhan ini nantinya akan dijadikan sebagai patokan dalam melakukan proses analisis selanjutnya.

a. Algoritma K-Means

Dalam menentukan jumlah *cluster* terbaik dataset diproses menggunakan *tools* rapidminer dengan beberapa percobaan nilai K *cluster* mulai dari 3 hingga 6 dengan masing-masing *max runs* 10 kali dan *max optimization steps* 100. Peneliti menerapkan parameter *performance cluster distance* untuk melihat performa atau nilai *Davies Bouldin Index* (DBI) yang dihasilkan pada setiap nilai K. Metode DBI merupakan metode untuk menetapkan jumlah *cluster* terbaik yang ditunjukkan dengan nilai K optimum pada nilai *Davies Bouldin Index* (DBI) yang paling kecil. Berdasarkan pada tabel dibawah ini dapat diketahui *cluster* terbaik pada nilai K = 4 dengan perolehan nilai sebesar 0.465.

TABEL V
 HASIL NILAI DAVIES BOULDIN INDEKS

Cluster	DBI
3	0.537
4	0.465
5	0.476
6	0.504

Adapun hasil dari 4 model *cluster* yang dihasilkan dapat diketahui jumlah data yang masuk pada setiap *cluster* yaitu, pada *cluster* 0 menghasilkan 475 transaksi, *cluster* 1 menghasilkan 674 transaksi, *cluster* 2 menghasilkan 13 transaksi, dan *cluster* 3 menghasilkan 162 transaksi.

Attribute	cluster_0	cluster_1	cluster_2	cluster_3
Jumlah Produk	1.909	1.441	5.769	2.827
Amount	69254.632	38460.757	249096.154	123016.975
Kategori Produk	5.265	4.657	5.154	5.364
Jam	13.777	14.282	13.769	13.525

Gambar 6. Rata-rata Karakteristik Cluster

Setelah *cluster* ideal sudah terbentuk maka, dapat diketahui informasi mengenai rata-rata nilai atribut pada setiap *cluster* seperti pada gambar diatas. Menurut informasi diatas dapat disimpulkan mengenai pola pembelian pelanggan sebagai berikut :

1. Cluster 0

Berdasarkan pada tabel karakteristik *cluster* dapat diketahui dengan data sebanyak 475 transaksi memiliki rata-rata jumlah produk yang dibeli sebesar 1.909, untuk rata-rata total harga pembelian 69254.632, lalu rata-rata kategori produk yang dibeli sebesar 5.265, dan rata-rata jam atau waktu order sebesar 13.777. Pada *cluster* 0 semua atribut berada diatas rata-rata kecuali pada atribut jam. *Cluster* ini menyatakan bahwa rata-rata kedatangan para pelanggan pada pukul 13.00 hingga 14.00 dengan membelanjakan uangnya rata-rata hingga mampu mencapai Rp.69.254.632 pada kategori produk yang dibeli kisaran departemen 5 hingga 6. Meskipun, memiliki jumlah konsumen cukup tinggi *cluster* ini masih memerlukan perhatian karena, dilihat dari jumlah produk yang dibeli rata-rata hanya 1 hingga 2 produk saja. Maka, *cluster* 0 ini dapat dipilih sebagai solusi permasalahan.

2. Cluster 1

Berdasarkan pada tabel karakteristik *cluster* dapat diketahui dengan data sebanyak 674 transaksi memiliki rata-rata jumlah produk yang dibeli sebesar 1.441, untuk rata-rata total harga pembelian 38460.757, lalu rata-rata kategori produk yang dibeli sebesar 4.657, dan rata-rata jam atau waktu order sebesar 14.282. Jika melihat rata-rata semua atribut pada *cluster* ini menghasilkan nilai dibawah rata-rata dari hasil *descriptive statistics* kecuali nilai rata-rata pada atribut jam. Dapat dilihat rata-rata proses orderan berjalan secara aktif pada saat pukul 14.00 hingga 15.00 dengan membelanjakan uangnya rata-rata sebesar Rp.38.460.757 dan kategori produk yang dibeli kisaran departemen 4 hingga 5. Namun, jumlah produk yang dibeli rata-rata hanya 1 produk saja. Cluster ini dapat direkomendasikan sebagai solusi permasalahan melihat *cluster* ini memiliki konsumen paling tinggi, tetapi rata-rata total pendapatan dan jumlah pembelian paling rendah.

3. Cluster 2

Berdasarkan pada tabel karakteristik *cluster* dapat diketahui dengan data sebanyak 13 transaksi memiliki rata-rata jumlah produk yang dibeli sebesar 5.769, untuk rata-rata total harga pembelian 249096.154, lalu rata-rata kategori produk yang dibeli sebesar 5.154, dan rata-rata jam atau waktu order sebesar 13.769. Jika melihat hasil analisa *descriptive statistics* dengan informasi pada *cluster* ini dapat disimpulkan bahwa semua atribut menghasilkan nilai diatas rata-rata kecuali perolehan rata-rata pada atribut jam. Biasanya dalam *cluster* ini para pelanggan mulai melakukan pembelian pada pukul 13.00 hingga 14.00 dengan membelanjakan uangnya rata-rata mampu mencapai Rp.249.096,154 dan kategori produk yang dibeli kisaran departemen 5 hingga 6 dengan jumlah produk yang dibeli rata-rata 5 hingga 6 produk. Melihat deskripsi statistik transaksi konsumen pada *cluster* 2 ini sudah dapat diartikan bahwa para pelanggan ini bersifat *impulsive customer*, yang artinya pelanggan ini mampu berbelanja tanpa berpikir terlebih dahulu.

4. Cluster 3

Berdasarkan pada tabel karakteristik *cluster* dapat diketahui dengan data sebanyak 162 transaksi memiliki rata-rata jumlah produk yang dibeli sebesar 1.827, untuk rata-rata total harga pembelian 123016.975, lalu rata-rata kategori produk yang dibeli sebesar 5.488, dan rata-rata jam atau waktu order sebesar 13.525. Dapat disimpulkan bahwasannya pada *cluster* ini semua atribut menghasilkan nilai rata-rata dibawah hasil *descriptive statistics* kecuali nilai rata-rata pada atribut jam. Pada *cluster* 3 ini biasanya proses orderan mulai berjalan secara aktif terjadi pada saat jam 14.00 hingga 15.00 dengan membelanjakan uangnya rata-rata mampu mencapai Rp.123.016,975 dan kategori produk yang dibeli kisaran departemen 5 hingga 6. Pada *cluster* 3 ini memiliki rata-rata pembelian tertinggi kedua, tetapi para pelanggan hanya membeli produk 1 hingga 2 item saja. Dapat simpulkan pada *cluster* ini masih memerlukan adanya peningkatan jumlah pembelian produk agar semakin meningkatkan nilai penjualan, maka *cluster* 3 ini dapat dipilih sebagai solusi permasalahan.

b. Algoritma Apriori

Berdasarkan hasil dari proses analisis *clustering* yang sudah dilakukan sebelumnya, didapatkan tiga *cluster* yang akan menjadi fokus utama dalam pencarian aturan asosiasi untuk mencari solusi permasalahan penelitian ini terkait pembentukan strategi pemasaran pada Premium Salad.co berupa rekomendasi paket menu atau produk *bundling*. Dataset yang dipakai dalam proses pencarian aturan asosiasi merupakan transaksi yang mengandung dua produk atau lebih yang dibeli oleh pelanggan, sehingga komposisi dataset pada masing-masing *cluster* yaitu 321 transaksi untuk *cluster* 0, 228 untuk *cluster* 1, dan 127 transaksi untuk *cluster* 3.

TABEL VI
 HASIL PENGUJIAN KOMBINASI MINIMAL SUPPORT DAN MINIMAL CONFIDENCE

Min Support	Min Confidence	Rules (Cluster 0)	Rules (Cluster 1)	Rules (Cluster 3)
1%	40%	2	3	36
1%	50%	1	3	25
1%	60%	0	1	17
2%	40%	0	0	7
2%	50%	0	0	3
2%	60%	0	0	1

Berdasarkan pada tabel diatas, setelah melakukan serangkaian percobaan didapatkan untuk minimum *support* terbaik adalah 0.01 dan minimum *confidence* terbaik sebesar 50%. Peneliti dapat menentukan nilai minimum *support* 0,01 karena dari keseluruhan transaksi *frequent itemsets* paling banyak sebesar 0.016. Apabila nilai minimum *support* dinaikkan dengan minimal *confidence* 50% maka, untuk *cluster* 0 dan *cluster* 1 tidak terbentuk aturan asosiasi, dimana sesuai hasil analisis *clustering* pada ketiga *cluster* tersebut membutuhkan penyelesaian permasalahan dalam bentuk aturan asosiasi. Nilai minimum *confidence* dipilih sebesar 50% agar mampu memberikan

aturan asosiasi terbaik, karena melihat nilai minimum *support* yang digunakan sudah cukup rendah. Aturan asosiasi yang terbentuk diharapkan memiliki keterkaitan antar produk dengan kemungkinan suatu produk dapat dibeli secara bersamaan yaitu minimal sebesar 50%. Pernyataan ini didukung dengan hasil penelitian oleh Adie Wahyudi Oktavia Gama, I ketut Gede Darma Putra, I Putu Agung Bayupati [28].

1. Aturan Asosiasi Cluster 0

TABEL VII
 HASIL PENCARIAN ATURAN ASOSIASI PADA CLUSTER 0

Premises	Conclusion	Support	Confidence	Lift Ratio
Beef Sandwich	Avocado Smoothies Bowl 450ml	0.012	0.571	7.337

2. Aturan Asosiasi Cluster 1

TABEL VIII
 HASIL PENCARIAN ATURAN ASOSIASI PADA CLUSTER 1

Premises	Conclusion	Support	Confidence	Lift Ratio
The Bowl Blackpepper (Free Egg)	The Bowl Matah (Free Egg)	0.013	0.500	7.125
Avocado Conshake	Salad Buah 500ml	0.013	0.750	3.054
Chicken Katsu Sandwich	Egg Veggie Salad	0.013	0.500	16.286

3. Aturan Asosiasi Cluster 3

TABEL IX
 HASIL PENCARIAN ATURAN ASOSIASI PADA CLUSTER 3

Premises	Conclusion	Support	Confidence	Lift Ratio
Beef Veggie Salad	Chicken Veggie Salad	0.016	0.500	2.190
Veggie Salad	Chicken Sandwich	0.039	0.500	3.175
Tuna Veggie Salad	Antioxidant Fusion	0.016	0.500	7.056
Tuna Salad Wrap	Chicken Salad Wrap	0.016	0.500	7.938
Dragon Fruit Conshake	Mango Conshake	0.016	0.500	15.875
Mango Conshake	Dragon Fruit Conshake	0.016	0.500	15.875
The Bowl Matah (Free Egg)	Bento With Teriyaki	0.039	0.556	6.414
Super Premium Fruit Salad 350ml	Chicken Veggie Salad	0.016	0.667	2.920
Salad Buah 500ml, Chicken Sandwich	Veggie Salad	0.016	0.667	8.467
Salad Buah 500ml, Immune Booster	Antioxidant Fusion	0.016	0.667	9.407
Semangka Juice	Bento Matah	0.024	0.750	13.607
Apel Juice	Salad Buah 500ml	0.016	1	5.080
Apel Juice	Antioxidant Fusion	0.016	1	14.111
Berry Conshake	Mango Conshake	0.016	1	31.750
Apel Juice	Salad Buah 500ml, Antioxidant Fusion	0.016	1	42.333
Immune Booster, Salad Buah 350ml	Mango Juice	0.016	1	15.875

Berdasarkan hasil uji pencarian aturan asosiasi yang sudah selesai dilakukan dengan menetapkan syarat minimum *support* dan minimum *confidence* yang sama pada masing-masing *cluster* menghasilkan jumlah aturan asosiasi yang berbeda. Pada *cluster 0* dengan dataset sebanyak 321 transaksi menghasilkan 1 *rules* dengan tingkat kepercayaan 57% pada kombinasi produ jika seorang pelanggan membeli beef sandwich maka, kemungkinan pelanggan akan membeli avocado smoothies bowl 450ml . Pada *cluster 1* dengan dataset sebanyak 228 transaksi menghasilkan 3 *rules* dengan tingkat kepercayaan 75% pada kombinasi produ jika seorang pelanggan membeli avocado conshake maka pelanggan akan membeli salad buah 500ml. Dan pada *cluster 3* dengan dataset paling sedikit sejumla 127 transaksi mampu menghasilkan 16 *rules* dengan tingkat kepercayaan paling tinggi sebesar 100% dan terdapat satu kombinasi itemset yang memiliki tingkat kepercayaan 100% sekaligus nilai *lift ratio* paling tinggi diantara itemset yang terbentuk, dimana seorang pelanggan yang membeli apel juice memiliki kemungkinan akan melakukan pembelian pada produk salad buah 500ml dan antioxidant fusion. Dari ketiga tabel diatas dapat disimpulkan bahwa, hasil yang didapatkan bersifat kuat untuk menangani strategi pemasaran. Hal ini dapat dibuktikan dengan melihat nilai lift ratio dari semua *rules* yang terbentuk bernilai lebih dari satu. Namun, untuk syarat

minimum *support* yang ditetapkan pada penelitian kali ini dapat dikatakan cukup rendah, tetapi nilai nilai *confidence* yang tinggi memiliki kemungkinan besar produk tersebut akan dibeli secara bersamaan seperti penelitian yang sudah dilakukan oleh Griya Jitri Pabutungan dan Hindriyanto Dwi Purnomo [22].

Mengkaji dari penelitian terdahulu seperti yang sudah dilakukan oleh Sindhy Genjang Setyorini, Ega Kurnia Sari, Lutfy Rahma Eliya, Shinta Ayunda Putri [19] dan Imaduddin Syukra, Assa Hidayat, Muhammad Zakiy Fauzi [20] memiliki kesamaan mengenai penentuan nilai minimum terbaik yang berada pada range nilai support yang terbilang rendah karena, jika nilai minimum support berada pada range lebih tinggi maka terdapat beberapa cluster yang tidak terbentuk rules. Jika dilihat dari aturan asosiasi yang dihasilkan memiliki nilai persentase lift ratio lebih dari sama dengan 1 dan tingkat kepercayaan berada pada range 50% - 100%. Mengamati tinjauan literatur yang digunakan sebagai landasan pada penelitian ini memiliki perbedaan pada proses pencarian aturan asosiasi, dapat dilihat pada hasil penelitian ini hanya fokus pada cluster yang membutuhkan perhatian khusus untuk diberikan solusi atas permasalahan yang terjadi sehingga, tidak semua cluster diterapkan proses association rules. Pemilihan cluster yang membutuhkan perhatian khusus ini didapatkan dengan bantuan analisis deskriptif statistics pada masing-masing cluster yang dapat dilihat pada Gambar 5, lalu dilakukan analisis karakteristik setiap cluster seperti pada Gambar 6. Dapat ditarik kesimpulan dari hasil penelitian yang sudah dilakukan hasilnya terbilang memiliki insight yang cukup baik untuk dapat digunakan oleh pemilik bisnis sebagai pendukung pengambilan keputusan terkait inovasi strategi pemasaran yang baru. Sebagai saran bagi penelitian selanjutnya terkait market basket analysis menggunakan metode data mining disarankan dapat memperluas area analisis yang tidak hanya berfokus pada rekomendasi paket menu makanan atau paket bundling secara keseluruhan, tetapi dapat mengembangkan analisis lebih dalam untuk membentuk rekomendasi paket menu yang dapat disesuaikan berdasarkan waktu makan pelanggan pagi, siang, sore, dan malam atau dapat disesuaikan berdasarkan hari dengan tujuan untuk mengoptimalkan referensi menu berdasarkan perilaku belanja pelanggan sekaligus lebih mengembangkan kreatifitas pemasaran bisnis.

3.4. Evaluasi

Evaluasi yang didapatkan setelah melakukan pengujian data transaksional menggunakan metode *clustering* dan dilanjutkan dengan menggunakan metode *association rules* adalah :

1. Pencarian aturan asosiasi dilakukan pada *cluster* yang dianggap membutuhkan perhatian khusus untuk diberikan perbaikan atas permasalahan yang terjadi.
2. Tidak semua produk yang dibeli pada setiap transaksi memiliki relasi satu sama lain, meskipun terdapat kecenderungan frekuensi pembelian yang tinggi pada suatu departemen tertentu.
3. Hasil uji pencarian aturan asosiasi rata-rata kombinasi produk yang didapatkan berasal dari salah satu produk pada departemen tertentu yang terindikasi rendah dari segi jumlah pembelian.
4. Semakin banyak dataset yang digunakan, maka nilai *support* yang dipakai cenderung lebih rendah. *Cluster* 3 dengan jumlah dataset paling sedikit diantara *cluster* 0 dan *cluster* 1 yaitu sebanyak transaksi mampu membentuk aturan asosiasi paling banyak yaitu berjumlah 18 *rules*.

IV. KESIMPULAN

Berdasarkan hasil pengolahan data transaksi pada periode tujuh bulan terakhir tahun 2023 ini menghasilkan 4 *cluster* ideal dengan nilai uji validitas *Davies Bouldin Index* (DBI) sebesar 0,465. Pencarian aturan asosiasi difokuskan pada tiga cluster yang membutuhkan penyelesaian permasalahan dengan rata-rata permasalahan cluster 0,1, dan 3 adalah perlunya peningkatan penjualan produk. Berdasarkan hasil analisis pada proses clustering sebelumnya ditetapkan syarat minimum support 0,01 dan minimum confidence sebesar 0,5 yang sama pada masing-masing cluster. Meskipun kategori departemen yang paling laku dapat diketahui di masing-masing cluster pada hasil deskripsi statistik, tidak menutup kemungkinan hasil aturan asosiasi yang terbentuk terdapat kombinasi produk yang tidak termasuk dalam kategori salah satu produk pada departemen yang paling laris, tetapi terdapat kombinasi produk baik *premises* ataupun *conclusion* termasuk pada kategori produk paling laku. Dari hasil penelitian dapat disimpulkan, aturan asosiasi yang terbentuk pada masing-masing *cluster* ini menghasilkan sejumlah aturan asosiasi dengan kombinasi produk yang merata dari produk yang paling laris hingga produk yang kurang laris, sehingga kombinasi produk yang didapatkan lebih bervariasi. Perbedaan ini dapat terjadi karena jumlah dataset yang digunakan pada masing-masing cluster, nilai minimum support, dan nilai minimum confidence yang ditetapkan. Semakin besar dataset yang digunakan maka, dibutuhkan nilai minimum support yang rendah. Diharapkan melalui hasil penelitian ini dapat digunakan sebagai bahan pendukung keputusan dalam membentuk *cross selling* sebagai salah satu strategi pemasaran baru bagi Premium Salad.co. Kombinasi produk yang sudah melewati nilai ambang batas (threshold) dari syarat minimum *support* dan *confidence* serta persentase nilai lift ratio yang cukup tinggi memiliki kemungkinan besar untuk dibeli secara bersamaan oleh pelanggan sehingga, dapat

digunakan untuk membentuk sebuah rekomendasi paket menu, paket promo produk *bundling*, dan pemilik bisnis dapat melakukan pengelolaan secara lebih teratur untuk stok bahan produk berdasarkan referensi makanan atau minuman yang memiliki probabilitas paling banyak diminati oleh pelanggan. Dengan demikian, strategi pemasaran yang tepat sesuai dengan kebutuhan pelanggan dan pemilik bisnis dapat diterapkan agar berpotensi dalam peningkatan profit penjualan, pemerataan frekuensi penjualan produk, serta memiliki kekuatan untuk berdaya saing yang dengan bisnis UMKM kuliner lainnya.

UCAPAN TERIMA KASIH

Puji dan syukur peneliti ucapkan utamanya kepada Tuhan Yang Maha Esa yang selalu senantiasa memberikan pertolongan serta rahmat-Nya agar peneliti mampu menyelesaikan tugas akhir mahasiswa sebagai syarat kelulusan dari universitas dalam bentuk jurnal ilmiah ini. Peneliti sangat berterimakasih kepada seluruh pihak yang telah membantu dalam bentuk dukungan dan pikiran sehingga, penelitian ini dapat terselesaikan dengan baik.

DAFTAR PUSTAKA

- [1] A. Setyaningrum and Y. Ramawati, "Peran Dimensi-Dimensi Entrepreneurial Marketing dalam Mendorong Kesuksesan Bisnis UMKM di Industri Ekonomi Kreatif," *J. Ekon. Bisnis dan Kewirausahaan*, vol. 9, no. 2, p. 125, 2020, doi: 10.26418/jebik.v9i2.39818.
- [2] R. N. Fajri, "Pertumbuhan Ekonomi melalui Pengembangan Sistem Keuangan Usaha Mikro, Kecil dan Menengah (UMKM) Makanan dan Minuman di Yogyakarta," *Owner*, vol. 6, no. 2, pp. 1318–1335, 2022, doi: 10.33395/owner.v6i2.723.
- [3] L. Z. Firdausya and D. P. Ompusunggu, "Usaha Mikro Kecil Dan Menengah (Umkm) Di Era Digital Abad 21," *Tali Jagad J.*, vol. 1, no. 1, pp. 16–20, 2023.
- [4] K. K. Rafiah and D. H. Kirana, "Analisis Adopsi Media Sosial Sebagai Sarana Pemasaran Digital Bagi UMKM Makanan dan Minuman di Jatinangor," *Jesya (Jurnal Ekon. Ekon. Syariah)*, vol. 2, no. 1, pp. 188–198, 2019, doi: 10.36778/jesya.v2i1.45.
- [5] D. Widianto, M. Syafrullah, and ..., "Implementasi Algoritma Fp-Growth Dalam Menganalisa Pola Pesanan Berbasis Website Pada Dapur D3," *Pros. Semin. ...*, vol. 2, no. April, pp. 34–44, 2023, [Online]. Available: <http://senafti.budiluhur.ac.id/index.php/senafti/article/view/504%0Ahttp://senafti.budiluhur.ac.id/index.php/senafti/article/download/504/240>
- [6] P. Konsumen, P. Toserba, and Y. Semarang, "Implementasi algoritma apriori untuk analisis pola pembelian konsumen pada toserba Yusuf semarang," vol. 9, no. 2, pp. 1013–1021, 2024.
- [7] Y. D. Krisnawati *et al.*, "Jurnal Ilmiah Pendidikan Ekonomi Dari Relasi Berujung Promosi (Studi Kasus Mata Pencaharian Usaha Rumah Makan Pendahuluan Usaha Mikro Kecil Menengah (UMKM) Merupakan Salah Satu Sektor Ekonomi Yang Saat Ini Sedang Berkembang Dan Dijadikan Cara Alternatif)," *J. Ilm. Pendidik. Ekon.*, vol. 4, no. 2, 2019.
- [8] M. Yetri, S. Yakun, E. Stnik, and T. Dharma, "Analisis Pola Pemilihan Menu Pada Penang Corner Cafe Dan Resto Menggunakan Algoritma Apriori," *J-SISKO Tech*, vol. 114, no. 2, pp. 114–123, 2018.
- [9] F. Nuraeni, D. Tresnawati, Y. Handoko Agustin, and G. Fauzi, "Optimization of Market Basket Analysis Using Centroid-Based Clustering Algorithm and Fp-Growth Algorithm," *J. Tek. Inform.*, vol. 3, no. 6, pp. 1581–1590, 2022, doi: 10.20884/1.jutif.2022.3.6.399.
- [10] N. Rumui, rd Andi Roy, and th Pujo Hari Saputro Informatics Manajemen Politeknik Negeri FakFak FakFak, "Comparison of Apriori Algorithm and FP-Growth in Managing Store Transaction Data 1 st Syukron Anas, 2 nd," *Int. J. Comput. Inf. Syst. Peer Rev. J.*, vol. 03, no. 04, pp. 2745–9659, 2022, [Online]. Available: <https://ijcis.net/index.php/ijcis/index>
- [11] N. Mayasari, D. Y. Niska, and E. R. Samosir, "Implementasi Data Mining untuk Memprediksi Itemset Promosi Penjualan Pada CV. Sumber Segar Utama," *J. Tek. dan Inform.*, vol. 6, no. 1, pp. 31–36, 2019, [Online]. Available: <https://jurnal.pancabudi.ac.id/index.php/Juti>
- [12] Y. P. Sianipar, V. C. Mawardi, and T. Sutrisno, "Penggunaan Apriori Pada Rekomendasi Paket Menu Dan Dilengkapi Fitur Chatbot," *J. Ilmu Komput. dan Sist. Inf.*, vol. 10, no. 1, 2022, doi: 10.24912/jiksi.v10i1.17821.
- [13] S. F. Octavia, I. Permana, and S. Monalisa, "Penerapan Algoritma Association Rules Dalam Penentuan Pola Pembelian Berdasarkan Hasil Clustering," *J. Media Inform. Budidarma*, vol. 7, no. 3, pp. 956–965, 2023, doi: 10.30865/mib.v7i3.6129.
- [14] R. Muliono, "Analisis Efisiensi Algoritma Data Mining," *Semant. (Seminar Nas. Tek. Inform.)*, vol. 1, no. 1, pp. 978–602, 2017, [Online]. Available: [http://fimi.ua.ac.be/data/\[12\]](http://fimi.ua.ac.be/data/[12])
- [15] M. Kaur and S. Kang, "Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining," *Procedia Comput. Sci.*, vol. 85, no. Cms, pp. 78–85, 2016, doi: 10.1016/j.procs.2016.05.180.
- [16] F. R. ARRAFI, "Analisis Pola Perilaku Konsumen Menggunakan Metode Market Basket Analysis (Mba) Pada Toko Sayur Keluarga (Tosaga) Di Masa Pandemi Covid-19," 2021, [Online]. Available: <https://dspace.uii.ac.id/handle/123456789/37774%0Ahttps://dspace.uii.ac.id/bitstream/handle/123456789/37774/16522020.pdf?sequence=1&isAllowed=y>
- [17] N. Mamahit and A. Qoiriah, "Penerapan Algoritma Fp-Growth dan K-Means pada Data Transaksi Minimarket," *J. Informatics Comput. Sci.*, vol. 1, no. 02, pp. 78–83, 2020, doi: 10.26740/jinacs.v1n02.p78-83.
- [18] N. P. Dharshinni, E. Bangun, S. Karunia, and R. Damayanti, "Menu Package Recommendation using Combination of K-Means and FP-Growth Algorithms at Bakery Stores," *J. Mantik*, vol. 4, no. 2, pp. 1272–1277, 2020.
- [19] S. G. Setyorini, E. K. Sari, L. R. Elita, and S. A. Putri, "Analisis Keranjang Pasar Menggunakan Algoritma K-Means dan FP-Growth pada PT. Citra Mustika Pandawa," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 41–46, 2021, doi: 10.57152/malcom.v1i1.62.
- [20] M. Z. F. Imaduddin Syukra, Assad Hidayat, "Implementation of K-Medoids and FP-Growth Algorithms for Grouping and Product Offering Recommendation," *Indonesian J. Artif. Intell. Data Min.*, vol. 2, no. 2, p. 107, 2019.
- [21] G. Triyandana, L. A. Putri, and Y. Umaidah, "Penerapan Data Mining Pengelompokan Menu Makanan dan Minuman Berdasarkan Tingkat Penjualan Menggunakan Metode K-Means," *J. Appl. Informatics Comput.*, vol. 6, no. 1, pp. 40–46, 2022, doi: 10.30871/jaic.v6i1.3824.
- [22] G. Jitri Pabutungan and H. Dwi Purnomo, "Analisa Market Basket Analysis untuk Melihat Pola Transaksi Customer Menggunakan Algoritma Apriori dan FP-Growth," *J. Media Inform. Budidarma*, vol. 7, pp. 966–974, 2023, doi: 10.30865/mib.v7i3.6152.
- [23] S. Ahmad Fadhilah, "ANALISIS PERILAKU PEMBELIAN KONSUMEN MENGGUNAKAN METODE ASSOCIATION RULE - MARKET BASKET ANALYSIS DAN CLUSTERING ANALYSIS (Studi Kasus: JORE COFFEE & EATERY)," *dspace.uii.ac.id*, no. 8.5.2017, pp. 2003–2005, 2022, [Online]. Available: www.aging-us.com
- [24] E. Umar, D. Manongga, and A. Iriani, "Market Basket Analysis Menggunakan Association Rule dan Algoritma Apriori Pada Produk Penjualan Mitra Swalyan Salatiga," *J. Media Inform. Budidarma*, vol. 6, no. 3, p. 1367, 2022, doi: 10.30865/mib.v6i3.4217.
- [25] A. Aquila, C. Pabendon, and H. D. Purnomo, "JURNAL MEDIA INFORMATIKA BUDIDARMA Penerapan Algoritma Apriori dan FP-Growth Untuk Market Basket Analisis Pada Data Transaksi NonPromo," vol. 7, pp. 975–984, 2023, doi: 10.30865/mib.v7i3.6153.
- [26] D. Marini Umi Atmaja, A. Rahman Hakim, D. Haryadi, and N. Suwaryo, "Penerapan Algoritma K-Nearest Neighbor Untuk Prediksi

- Pengelompokan Tingkat Risiko Penyebaran Covid-19 Jawa Barat," *Dewi Mar. Umi Atmaja, SNTEM*, vol. 1, no. November, pp. 1218–1226, 2021.
- [27] R.-Q. Liu, Y.-C. Lee, and H.-L. Mu, "Customer Classification and Market Basket Analysis Using K-Means Clustering and Association Rules: Evidence from Distribution Big Data of Korean Retailing Company," *Knowl. Manag. Rev.*, vol. 19, no. 4, pp. 59–76, 2018, doi: 10.15813/kmr.2018.19.4.004.
- [28] A. W. Oktavia Gama, I. K. Gede Darma Putra, and I. P. Agung Bayupati, "Implementasi Algoritma Apriori Untuk Menemukan Frequent Itemset Dalam Keranjang Belanja," *Maj. Ilm. Teknol. Elektro*, vol. 15, no. 2, pp. 21–26, 2016, doi: 10.24843/mite.1502.04.