

Analisis Perbandingan Algoritma Apriori dan FP-Growth dalam Asosiasi Data Mining pada Data Transaksi Minimarket FS Barokah Berbasis Google Colab

Bagus Aris Setiawan¹⁾, Yoga Diyas Saputra²⁾, M. Mustahikhul Munir³⁾,
Muhammad Arifin⁴⁾

^{1,2,3,4}Program Studi Sistem Informasi, Universitas Muria Kudus

^{1,2,3,4} Jl. Lkr. Utara, Kayuapu Kulon, Gondangmanis, Kec. Bae, Kabupaten Kudus, Jawa Tengah
59327

Email: ¹202353139@std.umk.ac.id, ²202353151@std.umk.ac.id, ³202353010@std.umk.ac.id,
⁴arifin.m@umk.ac.id

Abstract

Data Mining is a data analysis technique used to discover hidden patterns in transaction data to support business decision-making. This study aims to analyze consumer purchasing patterns and compare the performance of the Apriori and Frequent Pattern Growth (FP-Growth) algorithms using transaction data from FS Barokah Minimarket. The dataset consisted of 1,000 transactions with 27 unique product items and underwent data cleaning, transformation, and one-hot encoding processes. The analysis was conducted using Google Colab with the pandas and mlxtend libraries. The results showed that both Apriori and FP-Growth produced identical outputs, namely 139 frequent itemsets and 224 association rules. The strongest association pattern was found between Air Galon and Chocolate with a lift ratio of 1.31, followed by Biscuits and Cheese with a lift ratio of 1.24. In terms of execution time, Apriori performed faster than FP-Growth on the analyzed dataset. These findings can support cross-selling strategies, product bundling, and more effective product placement in FS Barokah Minimarket.

Keyword: Data Mining, Association Rule, Market Basket Analysis, Apriori, FP-Growth.

Abstrak

*Data Mining merupakan teknik analisis data yang digunakan untuk menemukan pola tersembunyi dalam data transaksi guna mendukung pengambilan keputusan bisnis. Penelitian ini bertujuan menganalisis pola pembelian konsumen serta membandingkan kinerja algoritma Apriori dan Frequent Pattern Growth (FP-Growth) pada data transaksi Minimarket FS Barokah. Dataset yang digunakan terdiri dari 1.000 transaksi dengan 27 item produk unik yang telah melalui proses *data cleaning*, transformasi data, dan *one-hot encoding*. Analisis dilakukan menggunakan Google Colab dengan bantuan pustaka *pandas* dan *mlxtend*. Hasil penelitian menunjukkan bahwa algoritma Apriori dan FP-Growth menghasilkan keluaran yang sama, yaitu 139 *frequent itemsets* dan 224 aturan asosiasi. Pola asosiasi terkuat ditemukan pada pasangan Air Galon dan Coklat dengan nilai *lift ratio* sebesar 1,31, diikuti pasangan Biskuit dan Keju dengan nilai *lift ratio* sebesar 1,24. Dari sisi waktu eksekusi, Apriori lebih cepat dibandingkan FP-Growth pada dataset yang digunakan. Hasil penelitian ini dapat dimanfaatkan untuk mendukung strategi *cross-selling*, *bundling* produk, dan penataan barang yang lebih efektif di Minimarket FS Barokah.*

Kata Kunci: Data Mining, Association Rule, Market Basket Analysis, Apriori, FP-Growth.

1. PENDAHULUAN

Berbagai aspek kehidupan masyarakat, termasuk bisnis, ekonomi, dan pendidikan, telah banyak dipengaruhi oleh pesatnya kemajuan teknologi informasi [1]. Pengusaha, terutama di sektor penjualan, terus mengumpulkan informasi dan data untuk meningkatkan keuntungan dan mengurangi kerugian. Untuk mencapai tujuan tersebut, digunakan salah satu teknik, yaitu Data Mining, yang melibatkan berbagai disiplin ilmu seperti statistika, kecerdasan buatan, dan pengolahan citra, serta berfokus pada ekstraksi pengetahuan baru dari data dalam database untuk meningkatkan manfaat informasi tersebut [2]. *Market Basket Analysis (MBA)* atau Analisis Keranjang Belanja merupakan salah satu aplikasi penting dari data mining di bidang pemasaran [3]. Metode ini digunakan untuk menganalisa pola pembelian di toko seperti minimarket. Banyak transaksi penjualan di minimarket menghasilkan volume data yang besar dengan pola yang berbeda, tetapi seringkali data tersebut dibiarkan menumpuk sebagai arsip digital tanpa dimanfaatkan dengan baik untuk strategi bisnis [4].

Studi ini berfokus pada Minimarket FS Barokah yang berlokasi di Desa Buaran, Kecamatan Mayong, Kabupaten Jepara, dengan tujuan menganalisis perbandingan algoritma asosiasi melalui platform Google Colab. Sebagai penyedia berbagai barang kebutuhan sehari-hari, Minimarket FS Barokah menghadapi tantangan dalam optimalisasi profit. Berdasarkan analisis awal dan hasil wawancara dengan pemilik, diketahui bahwa pendapatan penjualan mengalami tren penurunan setiap bulannya sejak awal tahun 2025. Meskipun data transaksi periode Januari hingga Mei 2025 telah tersedia, belum terdapat sistem otomatis yang mampu menganalisis pola pembelian pelanggan. Akibatnya, proses pengambilan keputusan strategis terkait manajemen stok barang dan perencanaan promosi masih dilakukan secara manual dan konvensional.

Dalam penelitian ini, terdapat beberapa permasalahan utama yang perlu dijawab untuk memahami penerapan algoritma asosiasi dalam menganalisis pola transaksi di Minimarket FS Barokah. Pertama, bagaimana penerapan algoritma *Apriori* dan *FP-Growth* dalam mengidentifikasi pola transaksi pelanggan di minimarket tersebut? Selanjutnya, bagaimana perbandingan performa kedua algoritma tersebut dalam hal waktu pemrosesan dan akurasi aturan asosiasi yang dihasilkan, mengingat pengujian awal menghasilkan asosiasi yang identik? Terakhir, bagaimana hasil analisis pola transaksi dapat digunakan untuk mendukung strategi rekomendasi produk yang lebih efektif bagi pemilik minimarket? Dengan menjawab pertanyaan-pertanyaan ini, penelitian diharapkan dapat memberikan wawasan yang lebih mendalam mengenai efektivitas algoritma asosiasi dalam bisnis ritel serta manfaatnya dalam pengambilan keputusan strategis.

2. METODE PENELITIAN

2.1 Rancangan Penelitian

Penelitian ini menggunakan pendekatan data mining untuk analisis pola transaksi dengan algoritma *Apriori* dan *FP-Growth*. Kedua algoritma ini dipilih karena kemampuannya dalam menemukan aturan asosiasi dari dataset transaksi yang besar. *Apriori* dipilih karena merupakan metode dasar dalam *market basket analysis* yang banyak digunakan, meskipun memiliki keterbatasan dalam efisiensi pemrosesan pada dataset besar. *FP-Growth*, sebagai alternatif, dapat mengatasi keterbatasan tersebut dengan menggunakan struktur *FP-Tree*, yang lebih cepat dalam menghasilkan aturan asosiasi tanpa perlu melakukan iterasi berulang pada dataset. Algoritma *Eclat* tidak digunakan karena lebih cocok untuk dataset kecil dengan skema penyimpanan berbasis *intersecting set*. Dataset yang digunakan berasal dari data transaksi Minimarket FS Barokah, mencakup transaksi dari Januari hingga Mei 2025. Pemilihan dataset ini dilakukan karena minimarket memiliki pola transaksi yang stabil dan representatif terhadap kebiasaan belanja konsumen lokal.

2.2 Preprocessing Data

Proses preprocessing dilakukan untuk memastikan kualitas data sebelum analisis, meliputi:

- a) **Pembersihan Data**
Menghapus data transaksi yang duplikat atau tidak lengkap pada file transaksi Minimarket FS Barokah. Dari total 1.000 data transaksi awal, hasil pengecekan menunjukkan bahwa tidak ditemukan *missing value* maupun data duplikat, sehingga seluruh data dapat digunakan dalam proses analisis tanpa pengurangan jumlah baris.
- b) **Transformasi Data**
Mengonversi format transaksi menjadi representasi yang sesuai untuk algoritma asosiasi, seperti list item dalam setiap transaksi menggunakan *TransactionEncoder* pada platform Google Colab.
- c) **Penentuan *Minimum Support***
Parameter *minimum support* ditetapkan sebesar 0.02 (2%) dan *minimum confidence* sebesar 0.1 (10%) berdasarkan hasil uji coba awal untuk menghindari aturan yang terlalu jarang tetapi tetap relevan dengan pola transaksi Minimarket FS Barokah. Nilai ini dipilih dengan mempertimbangkan karakteristik dataset yang terdiri dari 27 item unik dalam 1.000 transaksi, sehingga *threshold* yang ditetapkan mampu menghasilkan pola yang bermakna tanpa menghasilkan terlalu sedikit maupun terlalu banyak aturan asosiasi.

2.3 Data Mining

Data Mining adalah proses ekstraksi informasi penting dari data yang terdapat dalam database berukuran besar [1]. Definisi lain dari data mining adalah proses mengungkap pola, hubungan yang bermakna, dan kecenderungan dengan menganalisis kumpulan data berukuran besar menggunakan teknik pengenalan pola berbasis mesin [5].

2.4 Asosiasi

Analisis asosiasi, yang juga dikenal sebagai aturan asosiasi, merupakan teknik data mining yang dirancang untuk menemukan aturan hubungan dari berbagai kombinasi produk [1]. Misalnya, kita dapat mempelajari seberapa besar kemungkinan pelanggan membeli Gula dan Telur secara bersamaan di Minimarket FS Barokah. Pemilik minimarket dapat menggunakan data ini untuk mengatur tata letak barang atau membuat taktik pemasaran, seperti memberikan diskon untuk paket produk tertentu. Asosiasi terkenal karena kemampuannya untuk menilai apa yang ada dalam keranjang belanja. Ini biasanya disebut sebagai analisis keranjang belanja atau *market basket analysis*. Nilai *support* dan *confidence* adalah dua komponen utama yang digunakan untuk mengukur akurasi aturan asosiasi. Nilai *support* menunjukkan persentase kemunculan kombinasi item dalam database, sedangkan nilai *confidence* menunjukkan tingkat kekuatan hubungan antara item dalam aturan asosiasi.

2.5 Apriori

Algoritma *Apriori* adalah salah satu algoritma yang sering digunakan dalam aturan asosiasi. Tujuannya adalah untuk mengidentifikasi semua aturan asosiatif yang memenuhi batas minimum nilai support dan confidence [5]. Metodologi utama dalam melakukan analisis asosiasi terbagi menjadi dua tahap, yaitu:

- a) **Analisis Pola Frekuensi Tinggi** Tahap ini bertujuan untuk menemukan kombinasi item yang memenuhi ambang batas *minimum support* pada database. Nilai support untuk setiap item dapat dihitung dengan menggunakan rumus berikut [1]:

$$Support(A) = \frac{\text{Jumlah transaksi mengandung } A}{\text{Total Transaksi}}$$

Selain itu, untuk memperoleh nilai support untuk dua item dapat diperoleh dengan

menggunakan rumus berikut:

$$Support(A, B) = \frac{\text{Jumlah transaksi A dan B}}{\text{Total transaksi}}$$

Frequent itemset mengacu pada kumpulan item yang mempunyai frekuensi kemunculan lebih daripada nilai *minimum support* yang telah ditetapkan.

- b) Pembentukan Aturan Asosiasi Setelah menemukan pola frekuensi itemset, langkah berikutnya adalah membuat aturan asosiasi yang memenuhi ambang batas *minimum confidence*. Rumus berikut digunakan untuk menentukan nilai kepercayaan dari aturan A ke B [1]:

$$ConfidenceP(A|B) = \frac{\text{Jumlah transaksi A dan B}}{\text{Jumlah transaksi A}}$$

2.6 FP-Growth

Salah satu algoritma data mining yang banyak digunakan dalam analisis asosiasi untuk menemukan pola keterkaitan antaritem adalah *Frequent Pattern Growth (FP-Growth)*. Algoritma ini dikembangkan untuk mengatasi kelemahan algoritma Apriori, khususnya dalam proses pembentukan kandidat itemset yang dapat meningkatkan kompleksitas dan waktu komputasi. Berbeda dengan *Apriori*, *FP-Growth* tidak menghasilkan kandidat itemset secara berulang. Sebagai gantinya, algoritma ini memanfaatkan struktur data yang disebut *Frequent Pattern Tree (FP-Tree)* untuk merepresentasikan data transaksi secara lebih ringkas dan efisien [5].

Proses kerja algoritma *FP-Growth* secara umum terdiri dari dua tahapan utama, yaitu sebagai berikut:

- a) Pembangunan FP-Tree (*FP-Tree Construction*)

Pada tahap ini, algoritma terlebih dahulu menghitung frekuensi kemunculan setiap item dalam seluruh transaksi. Item yang tidak memenuhi nilai *minimum support* akan dieliminasi dari proses selanjutnya. Setelah itu, item yang tersisa diurutkan berdasarkan frekuensi kemunculannya dari yang tertinggi hingga terendah. Hasil pengurutan tersebut kemudian digunakan untuk membangun struktur *FP-Tree* yang merepresentasikan hubungan antaritem dalam dataset transaksi.

- b) Pencarian Frequent Itemset (*Frequent Pattern Mining*)

Setelah *FP-Tree* terbentuk, algoritma melakukan proses penelusuran (*traversal*) pada struktur pohon untuk menemukan pola atau kombinasi item yang sering muncul bersama dalam transaksi. Proses ini dilakukan tanpa menghasilkan kandidat itemset secara jelas, sehingga pencarian *frequent itemset* dapat dilakukan dengan lebih cepat dan efisien.

Dibandingkan dengan algoritma *Apriori*, *FP-Growth* memiliki tingkat efisiensi yang lebih tinggi karena mampu mengurangi jumlah proses pemindaian database serta menghilangkan kebutuhan pembentukan kandidat itemset. Oleh karena itu, algoritma ini sangat sesuai digunakan untuk menganalisis dataset transaksi yang berukuran besar dan kompleks.

2.7 Google Colab

Google Colab (*Colaboratory*), atau sering disingkat Colab, merupakan platform berbasis cloud yang dikembangkan oleh Google untuk menulis, menjalankan, menyimpan, dan membagikan program secara daring melalui Google Drive. Platform ini dapat dianggap sebagai *Jupyter Notebook* yang telah terintegrasi dengan layanan Google Drive sehingga memudahkan pengguna dalam pengelolaan dan kolaborasi proyek pemrograman [1]. Pada penelitian ini, Google Colab digunakan sebagai lingkungan pengolahan data karena memiliki kemampuan untuk memproses dataset dalam skala besar tanpa memerlukan spesifikasi perangkat keras lokal yang tinggi. Selain itu, Google Colab mendukung berbagai pustaka *Python* yang dibutuhkan dalam penelitian, seperti *pandas* dan *mlxtend*, yang digunakan untuk melakukan analisis aturan asosiasi. Keunggulan lain dari Google

Colab adalah tersedianya fasilitas komputasi berbasis *GPU* dan *TPU* yang dapat dimanfaatkan apabila proses pengolahan data memerlukan sumber daya yang lebih besar. Melalui platform ini, peneliti juga dapat memanfaatkan berbagai *library Python* yang telah tersedia untuk melakukan eksplorasi, analisis, dan visualisasi data transaksi Minimarket FS Barokah secara lebih efektif dan informatif.

3. HASIL DAN PEMBAHASAN

Pada pembahasan ini analisis difokuskan pada penerapan algoritma *Apriori* dan *FP-Growth* untuk mengidentifikasi pola asosiasi dalam dataset transaksi. Kedua algoritma dipilih karena kemampuannya dalam menggali hubungan antara item-item yang sering muncul bersama dalam suatu transaksi [6]. Proses ini diawali dengan pemahaman terhadap dataset yang digunakan, termasuk jenis data, jumlah transaksi, dan relevansi item yang dianalisis. Selanjutnya, ditentukan parameter *minimum support* dan *confidence* sebagai dasar untuk menghasilkan *frequent itemsets* dan aturan asosiasi. Dengan menetapkan tujuan analisis, diharapkan dapat diperoleh gambaran tentang efektivitas masing-masing algoritma dalam mengolah data dan menghasilkan pola asosiasi yang signifikan.

Langkah-langkah yang dapat dilakukan dalam memproses data algoritma *Apriori* dan *FP-Growth* dengan menggunakan Google Colab meliputi beberapa tahapan [1], yaitu: (1) pengumpulan data yang relevan untuk analisis asosiasi, (2) pembersihan data untuk menghilangkan data yang tidak lengkap atau tidak relevan, (3) pra-pemrosesan data seperti encoding atau transformasi agar dapat digunakan oleh algoritma, (4) penerapan algoritma *Apriori* atau *FP-Growth* untuk memperoleh pola asosiasi, dan (5) evaluasi serta penarikan kesimpulan dari hasil analisis untuk mengidentifikasi aturan asosiasi yang signifikan.

3.1 Pengumpulan Data

Data transaksi penjualan di Minimarket FS Barokah terdiri dari 1.000 entri transaksi yang tercatat dari tanggal 1 Januari 2025 hingga 14 Mei 2025. Data ini diekspor dari aplikasi penjualan dalam bentuk file Excel (.xlsx). Gambar berikut ini memperlihatkan contoh data mentah yang belum diproses yang diperoleh dari sistem.

Table 1. Sampel Data Penjualan Sebelum Proses Preprocessing

Tanggal	No Transaksi	Item	Total	Bayar	Kembali	Kasir
2025-01-01	7000000001	Gula, Detergen, Mie Gelas	57000	62000	5000	Udin
2025-01-01	7000000002	Air Galon, Saus, Beras, Aqua, Kopi Kapal Api	212000	217000	5000	Udin
2025-01-01	7000000003	Roti, Saus, Telur, Aqua	43000	48000	5000	Udin
2025-01-01	7000000004	Aqua, Kopi Kapal Api, Minyak Goreng	70000	72000	2000	Budi

2025-01-01	7000000005	Sirup, Indomie, Beras, Minyak Goreng, Keju	132000	134000	2000	Udin
2025-01-01	7000000006	Keju, Roti, Kerupuk, Kopi Kapal Api	64000	64000	0	Budi
2025-01-02	7000000007	Tisu, Kecap, Permen, Air Galon	83000	93000	10000	Budi
2025-01-02	7000000008	Beras, Detergen, Sirup, Kerupuk	98000	100000	2000	Ijah
2025-01-02	7000000009	Saus, Biskuit, Permen	49000	50000	1000	Budi
2025-01-02	7000000010	Margarin, Air Galon, Biskuit, Kopi Kapal Api, Kerupuk	149000	159000	10000	Ijah

Karena format tabel data penjualan di atas yang diperoleh masih memiliki beberapa kolom yang tidak diperlukan untuk analisis keranjang belanja (*Market Basket Analysis*), langkah pertama yang harus dilakukan sebelum memulai proses data mining adalah normalisasi data. Setelah itu, dilakukan seleksi terhadap field-field yang relevan, terutama kolom "No Transaksi" sebagai identitas unik dan kolom "Item" yang berisi daftar produk dalam satu transaksi. Hasil proses seleksi data ini digambarkan pada gambar berikut.

Table 2. Data Penjualan Setelah Dilakukan Seleksi Data

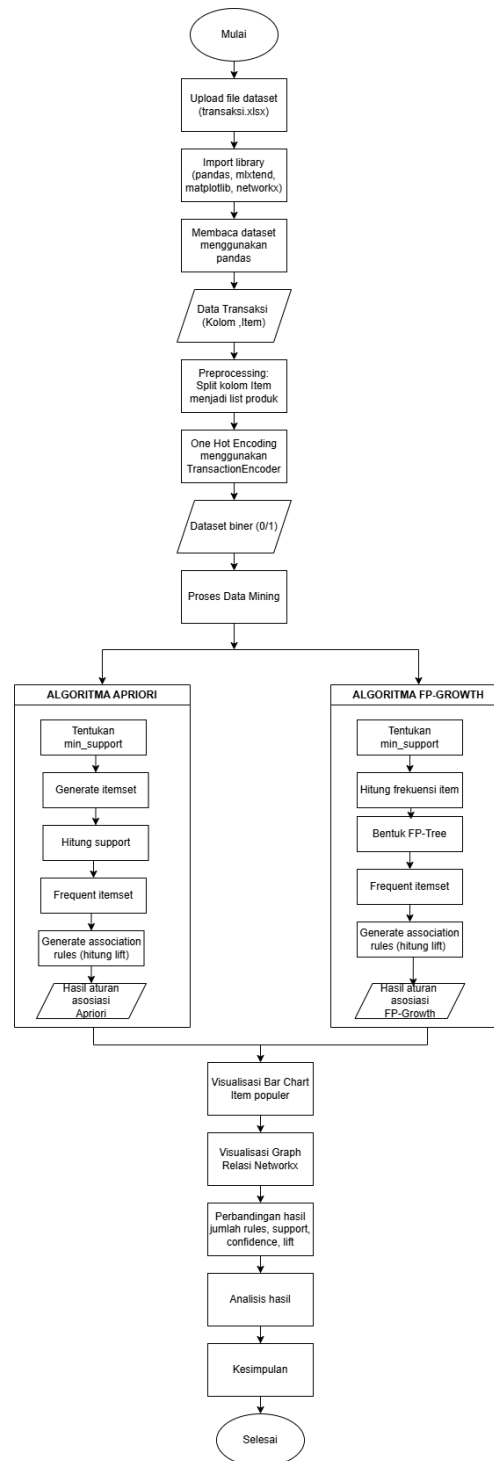
No Transaksi	Item
7000000001	Gula, Detergen, Mie Gelas
7000000002	Air Galon, Saus, Beras, Aqua, Kopi Kapal Api
7000000003	Roti, Saus, Telur, Aqua
7000000004	Aqua, Kopi Kapal Api, Minyak Goreng
7000000005	Sirup, Indomie, Beras, Minyak Goreng, Keju
7000000006	Keju, Roti, Kerupuk, Kopi Kapal Api

7000000007	Tisu, Kecap, Permen, Air Galon
7000000008	Beras, Detergen, Sirup, Kerupuk
7000000009	Saus, Biskuit, Permen
7000000010	Margarin, Air Galon, Biskuit, Kopi Kapal Api, Kerupuk

Berdasarkan data yang ditampilkan di atas, beberapa field dihapus untuk memastikan analisis berjalan lebih efisien dan relevan. Field yang tidak berkaitan langsung dengan pola pembelian seperti "Total", "Bayar", "Kembali", dan "Kasir", dihilangkan karena tidak berkontribusi dalam pembentukan aturan asosiasi. Selain itu, penghapusan field tertentu juga bertujuan untuk mengurangi kompleksitas data sehingga algoritma dapat bekerja lebih cepat dan optimal pada platform Google Colab. Data yang tidak relevan tersebut dihapus untuk menghindari noise yang dapat mengganggu hasil analisis kekuatan hubungan antarproduk. Dengan demikian, langkah ini memungkinkan analisis lebih fokus pada hubungan antarproduk berdasarkan pola transaksi pelanggan yang sebenarnya di Minimarket FS Barokah.

3.2 Preprocessing

Tahap selanjutnya adalah *preprocessing*. *Preprocessing* merupakan tahap perbaikan data yang tujuannya untuk memproses data sehingga siap untuk diproses lebih lanjut dan dapat digunakan secara efektif. Pada tahap ini, dataset disiapkan agar lebih terstruktur dengan fokus pada kolom yang esensial untuk analisis. Dengan demikian, data transaksi kini mencakup nomor transaksi dan nama barang (*item*). Selain itu, proses ini juga mencakup penghapusan data yang tidak relevan serta transformasi data menjadi format biner. Dalam data transaksi ini, field "No Transaksi" digunakan sebagai identitas unik untuk mengelompokkan item yang dibeli dalam satu waktu. Preprocessing data penjualan di Minimarket FS Barokah dengan Google Colab digambarkan dalam flowchart berikut.

Gambar 1. Flowchart *Preprocessing*

- a) Menyiapkan *Library* (Pustaka) Langkah ini bertujuan untuk menyiapkan pustaka (*library*) yang diperlukan untuk visualisasi dan pengolahan data. *Library* seperti *Pandas*, *Matplotlib*, dan *Seaborn* digunakan untuk manipulasi data dan pembuatan grafik guna mempermudah analisis pola transaksi. Selain itu, pustaka *MLxtend* (khususnya fungsi *apriori*, *association rules*, dan *fpgrowth*) disiapkan untuk memproses data dengan algoritma *Apriori* dan *FP-Growth*.
- b) Memuat Dataset Langkah ini melibatkan proses mengunggah (*upload*) dan memuat file excel yang diperoleh dari Minimarket FS Barokah. Dataset dimuat menggunakan fungsi

pd.read_excel() yang memungkinkan data dibaca ke dalam dataframe agar siap diproses lebih lanjut. Adapun hasil dari impor dataset dapat dilihat pada gambar berikut.

Gambar 2. Dataset penjualan FS BAROKAH

	Tanggal	No Transaksi	Item	Total	Bayar	Kembali	Kasir
0	2025-01-01	7000000001	Gula, Detergen, Mie Gelas	57000	62000	5000	Udin
1	2025-01-01	7000000002	Air Galon, Saus, Beras, Aqua, Kopi Kapal Api	212000	217000	5000	Udin
2	2025-01-01	7000000003	Roti, Saus, Telur, Aqua	43000	48000	5000	Udin
3	2025-01-01	7000000004	Aqua, Kopi Kapal Api, Minyak Goreng	70000	72000	2000	Budi
4	2025-01-01	7000000005	Sirup, Indomie, Beras, Minyak Goreng, Keju	132000	134000	2000	Udin

- c) Pembersihan dan Transformasi Data Selanjutnya dilakukan pembersihan data untuk memastikan tidak ada nilai kosong (*NaN*) yang dapat merusak hasil perhitungan. Data transaksi yang awalnya berbentuk daftar teks (*string*) diproses agar setiap item terpisah dengan benar. Untuk kebutuhan algoritma, dilakukan tahap Encoding menggunakan *TransactionEncoder*. Langkah ini mengubah daftar belanja menjadi matriks biner (*True/False*).

Gambar 3. Transformasi Data

Item_List	Air Galon	Aqua	Beras	Biskuit	Coklat	Detergen	Gula	Indomie	Kecap	Keju
No Transaksi										
7000000001	0	0	0	0	0	1	1	0	0	0
7000000002	1	1	1	0	0	0	0	0	0	0
7000000003	0	1	0	0	0	0	0	0	0	0
7000000004	0	1	0	0	0	0	0	0	0	0
7000000005	0	0	1	0	0	0	0	1	0	1

- d) Identifikasi Variasi Produk Berikutnya adalah mengidentifikasi jumlah produk unik yang tersedia dalam dataset. Hal ini dilakukan dengan perintah:

Gambar 4. Frekuensi Item Terlaris

	Item	Frekuensi
0	Gula	166
1	Telur	159
2	Sirup	159
3	Aqua	158
4	Detergen	158
5	Keju	155
6	Tisu	154
7	Permen	153
8	Mie Gelas	153
9	Kopi Kapal Api	151

Langkah ini bertujuan untuk menghitung jumlah produk yang berbeda dalam data transaksi, sehingga memberikan gambaran tentang variasi produk yang dijual di Minimarket FS Barokah.

- e) Visualisasi Produk Populer Menggunakan *library Matplotlib* dan *Seaborn*, dibuat visualisasi untuk menampilkan produk yang paling populer berdasarkan frekuensi transaksinya. Proses ini menghitung seberapa sering setiap produk muncul dalam seluruh transaksi.

3.3 Proses Data Mining

Dalam tahapan ini, data transaksi penjualan di Minimarket FS Barokah yang telah diproses melalui *preprocessing* dan *cleaning* akan diolah untuk memperoleh hasil dari analisis algoritma *Apriori* dan algoritma *FP-Growth*. Hasil analisis data penjualan dari kedua algoritma tersebut adalah sebagai berikut.

a) *Apriori*

Algoritma *Apriori* diterapkan pada data transaksi Minimarket FS Barokah yang terdiri dari 1.000 transaksi dengan 27 item unik. Proses ini menggunakan parameter *minimum support* sebesar 0.02 (2%) dan *minimum confidence* sebesar 0.1 (10%) sebagai ambang batas dalam pembentukan pola asosiasi. [7]. berikut adalah hasil dari pembuatan aturan asosiasi yang diperoleh menggunakan algoritma *Apriori*. Hasil ini mencakup kombinasi produk yang sering muncul pada saat yang sama dalam transaksi, beserta *nilai support*, *confidence*, dan *lift* untuk setiap pola yang dihasilkan.

Gambar 5. Hasil aturan asosiasi Algoritma Apriori

	antecedents	consequents	support	confidence	lift
0	(Air Galon)	(Coklat)	0.027	0.188811	1.311189
1	(Coklat)	(Air Galon)	0.027	0.187500	1.311189
50	(Biskuit)	(Keju)	0.029	0.193333	1.247312
51	(Keju)	(Biskuit)	0.029	0.187097	1.247312
100	(Tisu)	(Detergen)	0.029	0.188312	1.191846
101	(Detergen)	(Tisu)	0.029	0.183544	1.191846
161	(Kerupuk)	(Teh Celup)	0.024	0.161074	1.184366
160	(Teh Celup)	(Kerupuk)	0.024	0.176471	1.184366
218	(Susu)	(Shampoo)	0.024	0.167832	1.173652
219	(Shampoo)	(Susu)	0.024	0.167832	1.173652

Berdasarkan hasil implementasi algoritma *Apriori* pada data transaksi Minimarket FS Barokah, ditemukan hubungan asosiasi yang paling signifikan antara produk Air Galon dan Coklat. Berdasarkan hasil pengolahan data, ditemukan bahwa pelanggan yang membeli Air Galon memiliki peluang (*nilai confidence*) sebesar 18.88% untuk juga membeli Coklat. Hubungan ini diperkuat dengan nilai *lift ratio* sebesar 1.31, yang menunjukkan bahwa kehadiran Air Galon dalam keranjang belanja meningkatkan kemungkinan pembelian Coklat sebanyak 1.31 kali lipat dibandingkan pembelian secara acak. Selain itu, hubungan asosiasi juga ditemukan antara produk Biskuit dan Keju dengan nilai *confidence* sebesar 19.33% dan *lift ratio* 1.24, serta antara Detergen dan Tisu dengan nilai *confidence* 18.35% dan *lift ratio* 1.19. Seluruh aturan asosiasi yang ditemukan memiliki nilai *lift ratio* di atas 1, yang mengindikasikan bahwa seluruh pola pembelian tersebut bukan terjadi secara

kebetulan melainkan memang terdapat kecenderungan nyata konsumen untuk membeli produk-produk tersebut secara bersamaan.

Secara strategis, manajemen Minimarket FS Barokah dapat menata produk Air Galon berdekatan dengan Coklat, serta Biskuit berdekatan dengan Keju untuk memicu pembelian impulsif, atau menawarkan promosi bundling pada produk-produk tersebut. Selain itu, produk Detergen dan Tisu yang juga memiliki hubungan asosiasi kuat dapat dijadikan paket promosi untuk meningkatkan nilai transaksi per pelanggan. Pemeliharaan stok yang sinkron untuk pasangan item-item ini menjadi krusial agar tidak kehilangan potensi penjualan. Kesimpulannya, algoritma *Apriori* berhasil menemukan 224 aturan asosiasi dari 139 *frequent itemsets* dalam waktu eksekusi yang sangat singkat yaitu ± 0.02 detik, membuktikan bahwa algoritma ini memberikan insight yang akurat dan berbasis data dalam memetakan perilaku belanja pelanggan untuk mendukung optimasi profitabilitas Minimarket FS Barokah.

b) *FP-Growth*

Pemrosesan data mining menggunakan algoritma *FP-Growth* dilakukan di platform Google Colab dengan parameter *minimum support* sebesar 0.02 (2%) dan *minimum confidence* sebesar 0.1 (10%). Tujuan dari proses ini adalah untuk menemukan *frequent itemsets* dalam data transaksi Minimarket FS Barokah secara lebih efisien. Berbeda dengan algoritma *Apriori*, *FP-Growth* lebih efisien dalam menemukan pola dengan membangun struktur pohon (*FP-Tree*) tanpa perlu melakukan pembangkitan kandidat itemset secara berulang, sehingga penggunaan memori lebih optimal terutama pada dataset berukuran besar. [8].

berikut adalah hasil dari pembuatan aturan asosiasi yang diperoleh menggunakan algoritma *FP-Growth*. Aturan ini mencakup kombinasi item yang sering muncul pada saat yang sama dalam transaksi, beserta nilai *support* dan *confidence* untuk mengetahui kekuatan hubungan antar item.

Gambar 6. Hasil aturan asosiasi Algoritma *FP-Growth*

	antecedents	consequents	support	confidence	lift
54	(Air Galon)	(Coklat)	0.027	0.188811	1.311189
55	(Coklat)	(Air Galon)	0.027	0.187500	1.311189
132	(Biskuit)	(Keju)	0.029	0.193333	1.247312
133	(Keju)	(Biskuit)	0.029	0.187097	1.247312
109	(Detergen)	(Tisu)	0.029	0.183544	1.191846
108	(Tisu)	(Detergen)	0.029	0.188312	1.191846
153	(Kerupuk)	(Teh Celup)	0.024	0.161074	1.184366
152	(Teh Celup)	(Kerupuk)	0.024	0.176471	1.184366
184	(Susu)	(Shampoo)	0.024	0.167832	1.173652
185	(Shampoo)	(Susu)	0.024	0.167832	1.173652

Berdasarkan hasil implementasi algoritma *FP-Growth* pada data transaksi Minimarket FS Barokah, ditemukan hubungan asosiasi yang paling signifikan antara produk Air Galon dan Coklat. Berdasarkan hasil pengolahan data, ditemukan bahwa pelanggan yang membeli Air Galon memiliki peluang (*nilai confidence*) sebesar 18.88% untuk juga membeli Coklat. Hubungan ini diperkuat dengan nilai *lift ratio* sebesar 1.31, yang menunjukkan bahwa kehadiran Air Galon dalam keranjang belanja meningkatkan

kemungkinan pembelian Coklat sebanyak 1.31 kali lipat dibandingkan pembelian secara acak. Selain itu, hubungan asosiasi juga ditemukan antara produk Biskuit dan Keju dengan nilai *confidence* sebesar 19.33% dan *lift ratio* 1.24, serta antara Tisu dan Detergen dengan nilai *confidence* 18.83% dan *lift ratio* 1.19. Seluruh aturan asosiasi yang ditemukan memiliki nilai *lift ratio* di atas 1, yang mengindikasikan bahwa seluruh pola pembelian tersebut bukan terjadi secara kebetulan melainkan memang terdapat kecenderungan nyata konsumen untuk membeli produk-produk tersebut secara bersamaan.

Algoritma *FP-Growth* bekerja dengan membangun struktur *FP-Tree* (*Frequent Pattern Tree*) dari keseluruhan data transaksi, kemudian melakukan proses mining pola secara langsung dari struktur pohon tersebut tanpa perlu melakukan pembangkitan kandidat itemset secara berulang seperti yang dilakukan oleh algoritma *Apriori*. Pendekatan ini membuat *FP-Growth* lebih efisien dalam hal penggunaan memori pada dataset yang lebih besar. Namun pada dataset Minimarket FS Barokah yang berjumlah 1.000 transaksi, *FP-Growth* membutuhkan waktu eksekusi sebesar ± 0.40 detik, yang relatif lebih lambat dibandingkan *Apriori* pada skala data yang kecil ini.

Hasil yang diperoleh dari *FP-Growth* sepenuhnya identik dengan hasil *Apriori*, yaitu menghasilkan 139 *frequent itemsets* dan 224 aturan asosiasi dengan nilai *support*, *confidence*, dan *lift ratio* yang sama persis. Hal ini membuktikan bahwa kedua algoritma memiliki tingkat akurasi yang setara dalam menemukan pola pembelian konsumen. Secara strategis, manajemen Minimarket FS Barokah dapat memanfaatkan hasil temuan *FP-Growth* ini sebagai dasar penataan produk Air Galon berdekatan dengan Coklat, serta Biskuit berdekatan dengan Keju untuk memicu pembelian impulsif, maupun menawarkan promosi bundling pada pasangan produk tersebut guna meningkatkan nilai transaksi per pelanggan dan mendukung optimasi profitabilitas toko.

3.4 Perbandingan Algoritma *Apriori* dan *FP-Growth*

Hasil proses algoritma *Apriori* dan *FP-Growth* yang telah dilakukan dapat dibandingkan melalui berbagai aspek komputasi dan statistik. Beberapa aspek yang digunakan dalam perbandingan kedua algoritma ini meliputi jumlah itemsets, support tertinggi, jumlah aturan, *confidence*, *lift*, efisiensi waktu, dan kompleksitas dataset [8].

a) Hasil Itemsets

Pada algoritma *Apriori* maupun *FP-Growth*, ditemukan bahwa jumlah *frequent itemsets* yang dihasilkan adalah identik, yaitu sebanyak 139 itemset. Produk dengan nilai support tertinggi ditemukan pada kombinasi Air Galon dan Coklat dengan nilai *support* sebesar 0.027. Meskipun kedua algoritma menghasilkan jumlah itemset yang sama, mekanisme *FP-Growth* dalam memetakan frekuensi pola transaksi terbukti lebih stabil tanpa mengalami redundansi proses, yang menunjukkan keoptimalannya dalam menangkap pola frekuensi pada dataset Minimarket FS Barokah.

b) Aturan Asosiasi

Kedua algoritma menghasilkan aturan asosiasi yang serupa secara kualitas. Pada algoritma *Apriori*, aturan asosiasi dengan nilai *confidence* tertinggi adalah (Biskuit $\rightarrow$ Keju) dengan *confidence* sebesar 0.1933 dan *lift* 1.2473. *FP-Growth* menghasilkan aturan yang sama dengan nilai *confidence* dan *lift* yang identik, dengan total 224 aturan asosiasi yang dihasilkan oleh kedua algoritma. Hal ini menunjukkan bahwa kualitas informasi atau insight bisnis yang diperoleh dari kedua algoritma tidak memiliki perbedaan signifikan untuk pengambilan keputusan.

c) Efisiensi Waktu

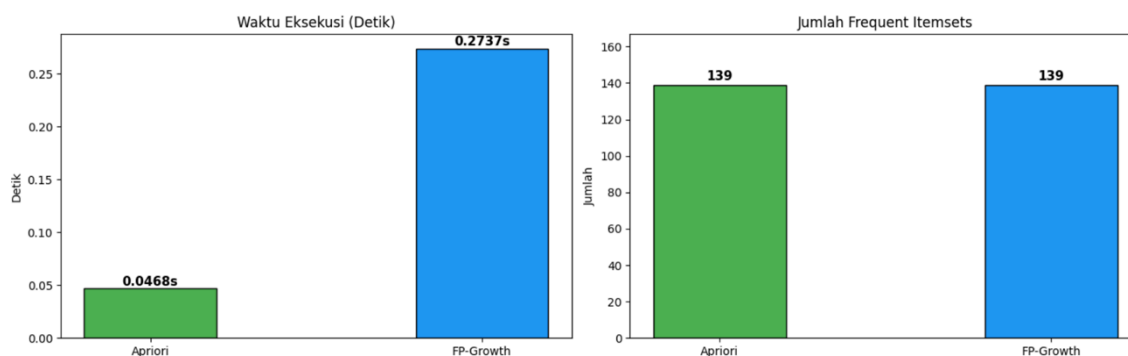
Dalam hal efisiensi waktu, *Apriori* terbukti lebih unggul dibandingkan *FP-Growth* pada dataset berukuran kecil ini. Berdasarkan hasil pengujian di Google Colab, *Apriori* membutuhkan waktu eksekusi sekitar ± 0.02 detik, sedangkan *FP-Growth* membutuhkan

waktu lebih lama yaitu ± 0.40 detik. Hal ini disebabkan karena pada dataset berukuran kecil seperti 1.000 transaksi, overhead pembangunan struktur *FP-Tree* pada *FP-Growth* justru membutuhkan waktu lebih lama dibandingkan proses kandidat itemset pada *Apriori*. Sebaliknya, pada dataset yang lebih besar, *FP-Growth* akan menunjukkan keunggulan karena hanya membutuhkan dua kali pemindaian basis data, sedangkan *Apriori* harus secara berulang menghasilkan kandidat *itemsets* dan melakukan pemeriksaan (*multiple scans*) terhadap data.

d) Kemampuan Menangani Dataset

Dari segi kompleksitas dataset, *Apriori* lebih cocok untuk digunakan pada dataset dengan skala kecil hingga menengah seperti data transaksi Minimarket FS Barokah yang berjumlah 1.000 transaksi dengan 27 item unik. Sebaliknya, *FP-Growth* menunjukkan keunggulan dalam menangani dataset yang lebih besar dan kompleks tanpa mengurangi kualitas hasil yang dihasilkan. Gambar berikut merangkum perbandingan dari kedua algoritma berdasarkan hasil komputasi pada dataset Minimarket FS Barokah.

Gambar 7. Perbandingan Algoritma Apriori dan FP-Growth



Berdasarkan Gambar perbandingan di atas, *Apriori* unggul dalam efisiensi waktu eksekusi dibanding *FP-Growth* pada dataset Minimarket FS Barokah yang berjumlah 1.000 transaksi, yaitu dengan waktu eksekusi ± 0.02 detik dibandingkan *FP-Growth* yang membutuhkan ± 0.40 detik. Meskipun demikian, kualitas itemset dan aturan asosiasi yang dihasilkan kedua algoritma adalah identik, yaitu menghasilkan 139 *frequent itemsets* dan 224 aturan asosiasi dengan nilai *support*, *confidence*, dan *lift ratio* yang sama persis. Oleh karena itu, pemilihan algoritma harus disesuaikan dengan ukuran dataset. *Apriori* lebih tepat untuk dataset kecil hingga menengah seperti data transaksi Minimarket FS Barokah karena strukturnya sederhana dan terbukti lebih cepat, sedangkan *FP-Growth* lebih tepat untuk dataset besar karena memanfaatkan struktur *FP-Tree* yang hanya membutuhkan dua kali pemindaian basis data sehingga prosesnya jauh lebih efisien.

4. KESIMPULAN

Berdasarkan perbandingan algoritma *Apriori* dan *FP-Growth* dalam penelitian ini, kedua metode terbukti sama-sama mampu menghasilkan aturan asosiasi dan *frequent itemsets* yang identik pada data transaksi Minimarket FS Barokah. Hasil pengujian menunjukkan bahwa kedua algoritma berhasil mengekstraksi sebanyak 224 aturan asosiasi dari 139 *frequent itemsets*, dengan pola pembelian yang paling kuat teridentifikasi pada kombinasi produk Air Galon dan Coklat (*support* 0.027, *lift* 1.31), Biskuit dan Keju (*support* 0.029, *lift* 1.24), serta Detergen dan Tisu (*support* 0.029, *lift* 1.19). Namun, dari sisi performa komputasi, *Apriori* menunjukkan efisiensi yang lebih tinggi pada dataset berukuran kecil ini dengan waktu eksekusi sebesar ± 0.02 detik, yang berarti jauh lebih cepat dibandingkan dengan algoritma *FP-Growth* yang memerlukan waktu ± 0.40 detik. Hal ini disebabkan karena pada dataset berjumlah 1.000 transaksi dengan 27 item unik, overhead

pembangunan struktur *FP-Tree* pada *FP-Growth* justru membutuhkan komputasi tambahan yang tidak sebanding dengan ukuran datanya. Keunggulan *FP-Growth* akan lebih terasa pada dataset yang jauh lebih besar, karena struktur *FP-Tree* mampu memproses data tanpa perlu membangkitkan kandidat itemset secara berulang, sehingga jauh lebih efektif untuk menangani dataset berskala besar dibandingkan *Apriori* yang performanya cenderung menurun seiring bertambahnya jumlah data dan item.

Hasil penelitian ini memberikan kontribusi nyata bagi strategi bisnis Minimarket FS Barokah melalui temuan pola asosiasi yang signifikan, seperti hubungan antara Air Galon dengan Coklat (*confidence 18.88%*) serta Biskuit dengan Keju (*confidence 19.33%*). Informasi ini memberikan implikasi strategis bagi pemilik minimarket dalam mengoptimalkan penataan produk (*product placement*) serta perancangan strategi promosi paket hemat (*bundling*) guna meningkatkan nilai transaksi per pelanggan. Meskipun demikian, penelitian ini memiliki keterbatasan karena dataset hanya berasal dari satu lokasi penelitian yaitu Minimarket FS Barokah dan terbatas pada periode Januari hingga Mei 2025, sehingga hasil kurang bisa digeneralisasi secara luas. Penelitian selanjutnya disarankan untuk menguji metode lain atau mengintegrasikan faktor eksternal seperti tren musiman, profil pelanggan, dan kategori produk guna meningkatkan akurasi aturan asosiasi yang dihasilkan. Secara keseluruhan, penelitian ini mempertegas efektivitas implementasi *cloud-based* data mining melalui Google Colab dalam mendukung pengambilan keputusan bisnis yang lebih cerdas dan berbasis data pada skala ritel menengah.

UCAPAN TERIMAKASIH

Penulis menyampaikan terima kasih yang sebesar-besarnya kepada keluarga atas dukungan doa dan motivasi yang tiada henti. Ucapan terima kasih juga ditujukan kepada dosen pembimbing atas arahan dan bimbingan intelektualnya selama penelitian ini. Terima kasih kepada pihak manajemen Minimarket FS Barokah atas penyediaan data transaksi, serta rekan-rekan mahasiswa Sistem Informasi atas diskusi dan bantuan teknis selama proses pengolahan data. Semoga hasil penelitian ini dapat memberikan kontribusi bermanfaat bagi pengembangan strategi bisnis berbasis data.

REFERENCES

- [1] M. Arifin, F. Helmi, and Iddrus, "Analisis Perbandingan Algoritma Asosiasi Data Mining Pada Minimarket Adi Poday Dengan Google Collab," *Jurnal Algoritma*, vol. 22, no. 1, pp. 103–114, May 2025, doi: 10.33364/algoritma/v.22-1.2177.
- [2] K. Alamat, "Desember 2021 Implementasi Data Mining Untuk Analisa Data Penjualan Cat Menggunakan Algoritma Apriori dan Fp Growth (Studi Kasus PT.Sumbermas Unggul Nastari) Widi Setiana 1 , Nabhilah Deviani 2 , Della Andina 3 , Numan Musyaffa 4 Program Studi Sistem Informasi," 2021. [Online]. Available: <http://jurnal.bsi.ac.id/index.php/larik>
- [3] R. Noviana, A. Hermawan, and D. Avianto, "Market Basket Analysis Menggunakan Algoritma Apriori dan FP Growth untuk Menentukan Pola Pembelian Konsumen," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 7, no. 3, p. 1474, Jul. 2023, doi: 10.30865/mib.v7i3.6304.
- [4] I. F. Rahman and D. Riana, "Market Basket Analysis untuk Penjualan Retail: Perbandingan Akurasi Algoritma Apriori dan FP-Growth Berbasis CRISP-DM," *Jurnal Algoritma*, vol. 22, no. 1, pp. 468–479, May 2025, doi: 10.33364/algoritma/v.22-1.2303.
- [5] E. Wahyu Pujiarto and A. Nasiri, "Analisis Perbandingan Kinerja Algoritma Apriori, FP-Growth dan Eclat dalam menemukan Pola Frekuensi pada Dataset INA-CBG'S Comparative Analysis of The Performance of The Apriori, FP-Growth and Eclat Algorithms In Finding Frequency Patterns In The Ina-Cbg's Dataset," *Cogito Smart Journal* |, vol. 9, no. 2, 2023.
- [6] K. Alamat, "Implementasi Data Mining Analisis Terhadap Data Penjualan Produk Herbal Dengan Metode Algoritma Apriori dan Fp Growth Nurlalah Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Nusa Mandiri Nusa Mandiri Tower," 2023. [Online]. Available: <http://jurnal.bsi.ac.id/index.php/larik>
- [7] A. Fergina *et al.*, "Implementasi Algoritma Apriori dan FP-Growth untuk Menganalisis Pola Pembelian Produk Skincare dan Kosmetik."

- [8] C. Tri *et al.*, “Perbandingan Kinerja Algoritma Apriori dan FP-Growth dalam Analisis Pola Pembelian Produk pada Timmy Store”, [Online]. Available: http://ejournal.ust.ac.id/index.php/Jurnal_Means/