



ANALISIS POLA PEMBELIAN PELANGGAN MENGGUNAKAN MARKET BASKET ANALYSIS (ALGORITMA APRIORI) UNTUK MENDUKUNG STRATEGI BUNDLING PRODUK PADA RETAIL

Lutfia Rahmadila^{a*}, Nazwa Puteri Calista^b, Rizka Rahayu^c

^a Fakultas Ilmu Komputer / Sistem Informasi; rahmadilalutfia@gmail.com, Universitas Pamulang; Jl. Raya Puspipetek No.46, Buaran, Kec. Pamulang, Kota Tangerang Selatan, Banten

^b Fakultas Ilmu Komputer / Sistem Informasi; nazwacalista03@gmail.com, Universitas Pamulang; Jl. Raya Puspipetek No.46, Buaran, Kec. Pamulang, Kota Tangerang Selatan, Banten

^c Fakultas Ilmu Komputer / Sistem Informasi; rizkarhyu2005@gmail.com, Universitas Pamulang; Jl. Raya Puspipetek No.46, Buaran, Kec. Pamulang, Kota Tangerang Selatan, Banten

* Penulis Korespondensi: Lutfia Rahmadila

ABSTRACT

This study implements the Apriori algorithm-based Market Basket Analysis (MBA) on the Online Retail II dataset sourced from Kaggle, which contains 1,067,371 transaction records from a UK-based online retail company covering the period 2009–2011. The research aims to identify customer purchasing patterns and discover product combinations frequently bought together to support bundling strategies and cross-selling recommendations. After a data cleaning process that reduced the dataset to 776,872 valid records, a basket matrix was formed comprising 29,527 transactions and 120 top products. The Apriori algorithm was applied with a minimum support threshold of 2%, minimum confidence of 30%, and minimum lift of 1.0, resulting in 148 frequent itemsets and 44 association rules. The analysis revealed several strong product pair associations, including Roses Regency Teacup and Saucer and Green Regency Teacup and Saucer (support: 0.0252, confidence: 0.7036, lift: 22.2670), Spaceboy Lunch Box and Dolly Girl Lunch Box (confidence: 0.6065, lift: 17.9440), and Alarm Clock Bakelike Red and Alarm Clock Bakelike Green (confidence: 0.6136, lift: 17.4052). These findings provide valuable insights for retailers to design effective product bundling, optimize store layouts, and implement targeted promotional campaigns.

Keywords: Market Basket Analysis; Association Rule Learning; Apriori Algorithm; Bundling Strategy; Online Retail

Abstrak

Penelitian ini mengimplementasikan Market Basket Analysis (MBA) berbasis algoritma Apriori pada dataset Online Retail II yang bersumber dari Kaggle, yang memuat 1.067.371 catatan transaksi dari perusahaan retail online berbasis di Inggris periode 2009–2011. Penelitian bertujuan mengidentifikasi pola pembelian pelanggan dan menemukan kombinasi produk yang sering dibeli bersama guna mendukung strategi bundling dan rekomendasi cross-selling. Setelah proses pembersihan data hingga diperoleh 776.872 data valid, dibentuk basket matrix yang terdiri dari 29.527 transaksi dan 120 produk teratas. Algoritma Apriori diterapkan dengan minimum support 2%, minimum confidence 30%, dan minimum lift 1,0, menghasilkan 148 frequent itemset dan 44 aturan asosiasi. Analisis menemukan beberapa pasangan produk dengan asosiasi kuat, di antaranya Roses Regency Teacup and Saucer dan Green Regency Teacup and Saucer (support: 0,0252; confidence: 0,7036; lift: 22,2670), Spaceboy Lunch Box dan Dolly Girl Lunch Box (confidence: 0,6065; lift: 17,9440), dan Alarm Clock Bakelike Red dan Alarm Clock Bakelike Green (confidence: 0,6136; lift: 17,4052). Temuan ini memberikan wawasan bagi peritel untuk merancang bundling produk yang efektif, mengoptimalkan tata letak toko, serta menerapkan kampanye promosi yang tepat sasaran.

Kata Kunci: Market Basket Analysis; Association Rule Learning; Algoritma Apriori; Strategi Bundling; Retail Online

1. PENDAHULUAN

Dalam era persaingan ritel yang semakin ketat, kemampuan memahami perilaku pelanggan secara mendalam menjadi salah satu keunggulan kompetitif yang krusial. Setiap transaksi yang dilakukan pelanggan menyimpan informasi berharga mengenai kombinasi produk yang dibeli, waktu pembelian, nilai transaksi, preferensi pelanggan, serta pola pembelian yang berpotensi mendukung pengambilan keputusan bisnis. Namun, tanpa pendekatan analitik yang tepat, informasi tersebut tidak dapat dimanfaatkan secara optimal untuk mendukung pengambilan keputusan bisnis [1].

Market Basket Analysis (MBA) hadir sebagai salah satu teknik data mining yang paling relevan dalam konteks ini. MBA bekerja dengan menganalisis kumpulan transaksi untuk mengidentifikasi produk-produk yang cenderung dibeli secara bersamaan dalam satu transaksi. Pengetahuan tentang keterkaitan antar produk tersebut dapat dimanfaatkan untuk berbagai keperluan strategis, mulai dari perancangan bundling produk, penempatan produk di rak toko atau halaman katalog online, hingga pengembangan sistem rekomendasi otomatis yang meningkatkan pengalaman berbelanja pelanggan [2].

Salah satu algoritma yang paling banyak digunakan dalam MBA adalah algoritma Apriori, yang pertama kali diperkenalkan oleh Agrawal dan Srikant pada tahun 1994 [3]. Algoritma ini bekerja dengan menemukan himpunan item yang sering muncul bersama (frequent itemsets) berdasarkan ambang batas support minimum, kemudian menghasilkan aturan asosiasi yang dievaluasi menggunakan metrik confidence dan lift. Meskipun terdapat alternatif algoritma yang lebih efisien secara komputasional seperti FP-Growth, algoritma Apriori tetap relevan digunakan karena transparansi prosesnya yang mudah diinterpretasikan [4].

Penelitian ini menggunakan dataset Online Retail II yang tersedia secara publik di Kaggle, mencakup transaksi nyata dari sebuah perusahaan retail online berbasis di Inggris. Dataset ini dipilih karena representativitasnya yang tinggi terhadap konteks e-commerce aktual. Pertanyaan penelitian yang mendasari studi ini adalah: produk apa saja yang sering dibeli secara bersamaan oleh pelanggan dalam satu transaksi, sehingga dapat digunakan sebagai dasar strategi bundling atau rekomendasi produk?

Tujuan bisnis yang ingin dicapai melalui penelitian ini meliputi: (1) mengidentifikasi pola pembelian pelanggan secara kuantitatif, (2) menemukan kombinasi produk yang sering dibeli bersama, (3) meningkatkan penjualan melalui strategi bundling dan cross-selling, serta (4) mengoptimalkan penempatan produk pada tata letak toko maupun katalog online.

2. TINJAUAN PUSTAKA

2.1. Market Basket Analysis

Market Basket Analysis (MBA) adalah teknik yang digunakan dalam data mining untuk mengidentifikasi pola pembelian konsumen dalam sebuah basis data transaksi. MBA membantu peritel memahami produk-produk mana yang cenderung dibeli secara bersamaan, sehingga memungkinkan mereka untuk merancang strategi pemasaran yang lebih efektif, mengoptimalkan tata letak produk, serta mengembangkan sistem rekomendasi yang relevan [5]. Penerapan MBA dalam konteks e-commerce terbukti mampu meningkatkan nilai rata-rata transaksi pelanggan melalui mekanisme cross-selling yang terarah [6].

2.2 Algoritma Apriori

Algoritma Apriori adalah metode klasik dalam Association Rule Mining yang bekerja secara iteratif. Algoritma ini dimulai dengan mencari itemset tunggal yang memenuhi ambang batas minimum support, kemudian secara bertahap menggabungkan itemset tersebut menjadi itemset yang lebih besar, hingga tidak ada lagi itemset baru yang memenuhi kriteria [3]. Proses ini memanfaatkan properti anti-monoton (Apriori property): jika suatu itemset tidak memenuhi minimum support, maka seluruh supersetnya juga tidak akan memenuhinya, sehingga ruang pencarian dapat dipangkas secara efisien [7].

2.3 Metrik Evaluasi Association Rules

Tiga metrik utama yang digunakan untuk mengevaluasi kekuatan aturan asosiasi adalah sebagai berikut [3][8]:

2.3.1. Support

Support mengukur seberapa sering suatu kombinasi item muncul dalam keseluruhan transaksi. Support dihitung sebagai rasio jumlah transaksi yang mengandung kombinasi item tertentu terhadap total jumlah transaksi.

$$\text{Support}(A \rightarrow B) = \frac{|A \cup B|}{N} \quad \text{Persamaan (1)}$$

2.3.2. Confidence

Confidence mengukur probabilitas item B dibeli ketika item A sudah dibeli. Nilai confidence yang tinggi mengindikasikan hubungan yang kuat antara kedua item.

$$\text{Confidence}(A \rightarrow B) = \frac{\text{Support}(A \cup B)}{\text{Support}(A)} \quad \text{Persamaan (2)}$$

2.3.3. Lift

Lift mengukur seberapa jauh hubungan antara A dan B melebihi probabilitas acak. Nilai lift > 1 menunjukkan bahwa pembelian A meningkatkan kecenderungan pembelian B secara signifikan, bukan sekadar kebetulan.

$$\text{Lift}(A \rightarrow B) = \frac{\text{Confidence}(A \rightarrow B)}{\text{Support}(B)} \quad \text{Persamaan (3)}$$

2.4 Penelitian Terkait

Beberapa penelitian sebelumnya telah menerapkan MBA dengan algoritma Apriori pada data transaksi ritel. Umar et al. [9] menerapkan MBA pada data penjualan Mitra Swalayan Salatiga dan berhasil mengidentifikasi aturan asosiasi yang digunakan untuk mengoptimalkan tata letak produk. Hossain et al. [10] membandingkan Apriori dan FP-Growth pada data basket dan menyimpulkan bahwa Apriori menghasilkan aturan yang lebih mudah diinterpretasikan meskipun lebih lambat secara komputasi. Khushi et al. [11] mengaplikasikan Apriori pada data e-commerce dan menemukan pola pembelian yang signifikan untuk mendukung strategi promosi digital.

3. METODOLOGI PENELITIAN

Metode penelitian yang digunakan terdiri dari empat tahapan utama, sebagaimana digambarkan pada alur berikut.

3.1 Pengumpulan Data

Dataset yang digunakan adalah Online Retail II yang tersedia secara publik di Kaggle. Dataset ini merupakan data transaksi nyata dari sebuah perusahaan retail online berbasis di Inggris yang menjual produk-produk hadiah dan dekorasi rumah. Dataset terdiri dari dua lembar kerja: Year 2009–2010 dan Year 2010–2011, yang digabungkan menjadi satu dataframe menggunakan operasi concat. Total data awal setelah penggabungan berjumlah 1.067.371 baris dengan 8 kolom variabel.

3.2 Pemrosesan dan Pembersihan Data

Proses pembersihan data dilakukan menggunakan Python dengan library pandas pada lingkungan Jupyter Notebook. Langkah-langkah yang dilakukan meliputi: (a) penghapusan baris dengan nilai kosong pada kolom kunci Invoice, Description, dan Customer ID; (b) penghapusan baris duplikat; (c) penghapusan transaksi pembatalan yang ditandai dengan kode invoice berawalan huruf 'C'; (d) penghapusan item tidak relevan berdasarkan StockCode non-produk seperti 'POST', 'D', 'M', 'BANK CHARGES', 'PADS', dan 'DOT'; serta (e) penghapusan baris dengan nilai Quantity dan Price yang bernilai nol atau negatif.

Hasil pembersihan data menghasilkan pengurangan dari 1.067.371 baris menjadi 776.872 baris, atau sebanyak 290.499 baris dihapus (sekitar 27,22% dari total data awal). Selanjutnya dilakukan transformasi data dengan menambahkan kolom Revenue sebagai hasil perkalian Quantity dengan Price.

Variabel	Tipe Data	Non-Null	Null	Unik	Keterangan
Invoice	str	776872	0	36645	ID Transaksi
StockCode	object	776872	0	4625	Kode Produk
Description	object	776872	0	5235	Nama Produk
Quantity	int64	776872	0	436	Jumlah Item Dibeli
InvoiceDate	datetime64[us]	776872	0	34308	Tanggal & Waktu Transaksi
Price	float64	776872	0	488	Harga Satuan (£)
Customer ID	str	776872	0	5862	ID Pelanggan
Country	str	776872	0	41	Negara Asal Transaksi
Revenue	float64	776872	0	3765	Hasil Kali Quantity x Price

Tabel 1. Daftar Variabel Awal Setelah Pemrosesan Data
 Sumber : Hasil Analisis Menggunakan Jupyter Notebook dan Python (2025)

3.3 Transformasi dan Pembentukan Basket Matrix

Untuk keperluan analisis MBA, dipilih 120 produk teratas berdasarkan frekuensi kemunculan dalam transaksi unik. Pemilihan 120 produk teratas dilakukan untuk memastikan relevansi analisis sekaligus menjaga efisiensi komputasi. Dari subset tersebut, dibentuk basket matrix menggunakan teknik one-hot encoding: setiap baris merepresentasikan satu transaksi (Invoice) dan setiap kolom merepresentasikan satu produk, dengan nilai 1 jika produk tersebut dibeli dalam transaksi itu dan 0 jika tidak. Basket matrix yang terbentuk memiliki dimensi 29.527 transaksi × 120 produk.

Jumlah 120 produk dipilih sebagai kompromi antara cakupan analisis dan efisiensi komputasi. Pemilihan jumlah yang lebih sedikit berpotensi menghilangkan pola asosiasi penting, sedangkan jumlah yang terlalu besar dapat meningkatkan kompleksitas komputasi tanpa memberikan peningkatan informasi yang signifikan.

Tabel 2. Transformasi

Description	6 RIBBONS RUSTIC CHARM	60 TEATIME FAIRY CAKE CASES	72 SWEETHEART FAIRY CAKE CASES	ALARM CLOCK BAKELIKE GREEN	ALARM CLOCK BAKELIKE RED	ASSORTED COLOUR BIRD ORNAMENT	BAKING SET 9 PIECE RETROSPOT	BAKING SET SPACEBOY DESIGN	BATHROOM METAL SIGN	CHARLOTTE BAG SUKI DESIGN	CHILLI LIGHTS	CHO HOT
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	1	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	1	0	0	0	0	0
4	0	0	0	0	0	0	1	0	0	0	0	0
...
29522	0	0	0	0	0	0	0	0	0	0	0	0
29523	0	0	0	0	0	0	0	0	0	0	0	1
29524	0	0	0	0	0	0	0	0	0	0	0	0
29525	0	0	0	1	1	1	0	0	0	0	0	0
29526	0	0	0	1	1	0	1	0	0	0	0	0

29527 rows × 120 columns

Sumber : Hasil Analisis Menggunakan Jupyter Notebook dan Python (2025)

3.4 Penerapan Algoritma Apriori

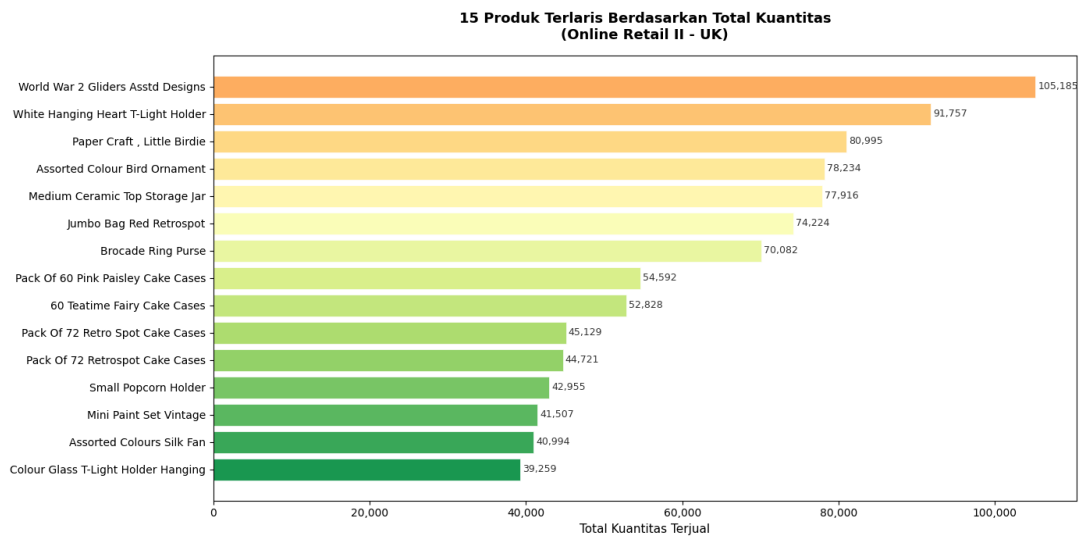
Algoritma Apriori diterapkan menggunakan library mlxtend dengan parameter berikut: (a) Minimum Support: 0,02 (2%); (b) Minimum Confidence: 0,30 (30%); (c) Minimum Lift: 1,0; serta (d) Maximum Itemset Length: 3. Parameter ini dipilih berdasarkan pertimbangan keseimbangan antara jumlah aturan yang dihasilkan dan relevansinya secara bisnis. Nilai minimum support 2% dipilih agar cukup selektif namun tetap menghasilkan itemset yang bermakna pada dataset bervolume besar.

Meskipun panjang itemset maksimum ditetapkan sebesar tiga item, hasil analisis menunjukkan tidak terdapat kombinasi tiga produk yang memenuhi ambang minimum support sebesar 2%, sehingga seluruh frequent itemset yang diperoleh hanya terdiri atas itemset tunggal dan pasangan produk.

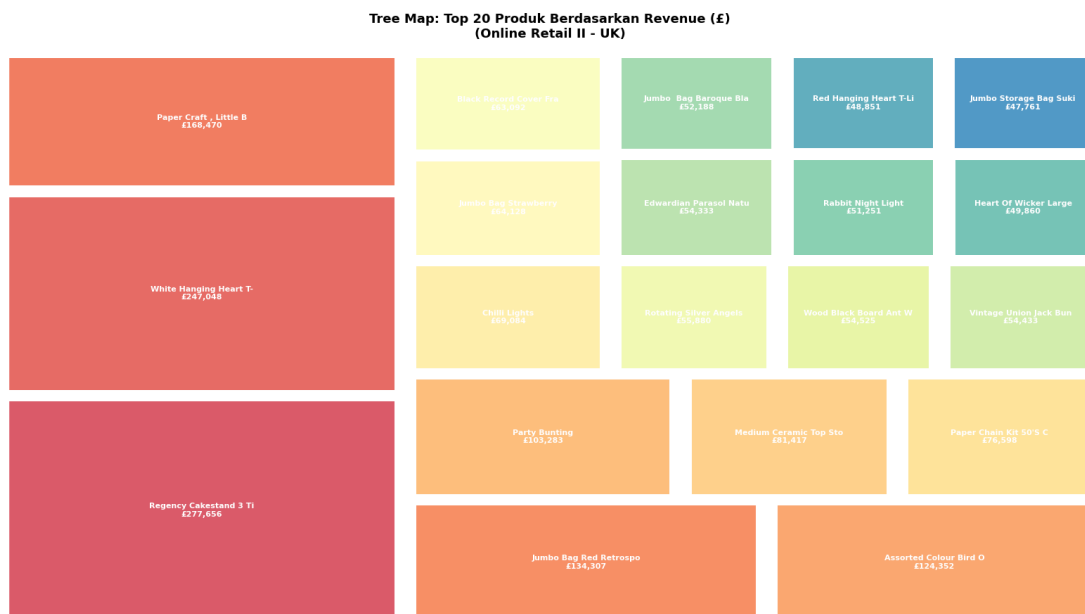
4. HASIL DAN PEMBAHASAN

4.1 Eksplorasi Data: Produk Terlaris

Analisis eksplorasi data dilakukan menggunakan visualisasi barplot horizontal untuk menampilkan 15 produk dengan total kuantitas terjual tertinggi. Hasil visualisasi menunjukkan bahwa White Hanging Heart T-Light Holder menjadi produk dengan frekuensi transaksi tertinggi dengan nilai support sebesar 0,1655 (16,55%), yang dihitung berdasarkan 29.527 transaksi unik pada basket matrix, diikuti oleh Regency Cakestand 3 Tier (support: 0,1123) dan Assorted Colour Bird Ornament (support: 0,0898). Visualisasi tree map berdasarkan total revenue mengkonfirmasi dominasi produk-produk dekorasi rumah dan aksesoris dapur dalam komposisi pendapatan keseluruhan.



Gambar 1. Barplot 15 Produk Terlaris Berdasarkan Total Kuantitas Online Retail II UK
 Sumber : Hasil Analisis Menggunakan Jupyter Notebook dan Python (2025)



Gambar 2. Tree Map Top 20 Produk Berdasarkan Revenue Online Retail II UK
 Sumber : Hasil Analisis Menggunakan Jupyter Notebook dan Python (2025)

4.2 Frequent Itemsets

Penerapan algoritma Apriori menghasilkan total 148 frequent itemsets, yang terdiri dari 121 itemset dengan panjang 1 (single itemsets) dan 27 itemset dengan panjang 2 (pasangan produk). Tidak ditemukan itemset dengan panjang 3 yang memenuhi ambang batas support minimum. Hasil ini menunjukkan bahwa meskipun banyak produk individual yang cukup populer, kombinasi tiga produk sekaligus belum cukup sering terjadi dalam satu transaksi untuk melampaui threshold yang ditetapkan.

Tabel 3. Sampel Frequent Itemsets Hasil Algoritma Apriori (Top 20 berdasarkan Support)

	Itemset	Support	Jumlah Item
0	WHITE HANGING HEART T-LIGHT HOLDER	0.1655	1
1	REGENCY CAKESTAND 3 TIER	0.1123	1
2	ASSORTED COLOUR BIRD ORNAMENT	0.0898	1
3	JUMBO BAG RED RETROSPOT	0.0885	1
4	PARTY BUNTING	0.0703	1
5	LUNCH BAG BLACK SKULL.	0.0676	1
6	LUNCH BAG SPACEBOY DESIGN	0.0635	1
7	REX CASH+CARRY JUMBO SHOPPER	0.0629	1
8	HOME BUILDING BLOCK WORD	0.0620	1
9	STRAWBERRY CERAMIC TRINKET BOX	0.0616	1
10	LUNCH BAG CARS BLUE	0.0609	1
11	HEART OF WICKER SMALL	0.0602	1
12	WOODEN FRAME ANTIQUE WHITE	0.0600	1
13	60 TEATIME FAIRY CAKE CASES	0.0598	1
14	HEART OF WICKER LARGE	0.0590	1
15	BAKING SET 9 PIECE RETROSPOT	0.0590	1
16	JUMBO STORAGE BAG SUKI	0.0580	1
17	PAPER CHAIN KIT 50'S CHRISTMAS	0.0573	1
18	JUMBO SHOPPER VINTAGE RED PAISLEY	0.0569	1
19	WOODEN PICTURE FRAME WHITE FINISH	0.0562	1
20	NATURAL SLATE HEART CHALKBOARD	0.0554	1

Sumber : Hasil Analisis Menggunakan Jupyter Notebook dan Python (2025)

4.3 Association Rules

Dari 148 frequent itemsets, dihasilkan 44 aturan asosiasi yang memenuhi seluruh threshold. Ringkasan statistik aturan asosiasi yang dihasilkan adalah sebagai berikut: rata-rata support 0,0247; rata-rata confidence 0,4829; rata-rata lift 8,7861; nilai lift minimum 2,2527; dan nilai lift maksimum 22,2670. Rata-rata lift sebesar 8,79 menunjukkan bahwa secara keseluruhan, produk-produk dalam aturan yang dihasilkan memiliki kecenderungan untuk dibeli bersama jauh melebihi probabilitas acak.

Secara umum, nilai lift di atas 10 dapat dikategorikan sebagai asosiasi yang sangat kuat, sedangkan nilai lift antara 5 hingga 10 menunjukkan asosiasi yang kuat. Dengan rata-rata lift sebesar 8,79, aturan yang dihasilkan dalam penelitian ini termasuk dalam kategori kuat dan memiliki relevansi bisnis yang tinggi.

Tabel 4. Implementasi Association Rules Terhadap MBA (Top 15 berdasarkan Lift)

	Antecedent (Jika Beli)	Consequent (Maka Beli)	Support	Confidence	Lift	Leverage	Conviction
0	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0252	0.7036	22.2670	0.0240	3.2672
1	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0252	0.7964	22.2670	0.0240	4.7349
2	SPACEBOY LUNCH BOX	DOLLY GIRL LUNCH BOX	0.0228	0.6065	17.9440	0.0215	2.4554
3	DOLLY GIRL LUNCH BOX	SPACEBOY LUNCH BOX	0.0228	0.6733	17.9440	0.0215	2.9465
4	ALARM CLOCK BAKELIKE RED	ALARM CLOCK BAKELIKE GREEN	0.0238	0.6136	17.4052	0.0224	2.4970
5	ALARM CLOCK BAKELIKE GREEN	ALARM CLOCK BAKELIKE RED	0.0238	0.6744	17.4052	0.0224	2.9518
6	SWEETHEART CERAMIC TRINKET BOX	STRAWBERRY CERAMIC TRINKET BOX	0.0277	0.6977	11.3316	0.0252	3.1042
7	STRAWBERRY CERAMIC TRINKET BOX	SWEETHEART CERAMIC TRINKET BOX	0.0277	0.4494	11.3316	0.0252	1.7442
8	LUNCH BAG RED RETROSPOT	LUNCH BAG PINK POLKADOT	0.0211	0.4092	10.9743	0.0191	1.6295
9	LUNCH BAG PINK POLKADOT	LUNCH BAG RED RETROSPOT	0.0211	0.5649	10.9743	0.0191	2.1802
10	WOODEN PICTURE FRAME WHITE FINISH	WOODEN FRAME ANTIQUE WHITE	0.0336	0.5989	9.9741	0.0303	2.3435
11	WOODEN FRAME ANTIQUE WHITE	WOODEN PICTURE FRAME WHITE FINISH	0.0336	0.5601	9.9741	0.0303	2.1454
12	HOT WATER BOTTLE TEA AND SYMPATHY	CHOCOLATE HOT WATER BOTTLE	0.0210	0.4777	9.7469	0.0188	1.8206
13	CHOCOLATE HOT WATER BOTTLE	HOT WATER BOTTLE TEA AND SYMPATHY	0.0210	0.4285	9.7469	0.0188	1.6728
14	PAPER CHAIN KIT VINTAGE CHRISTMAS	PAPER CHAIN KIT 50'S CHRISTMAS	0.0230	0.5507	9.6157	0.0206	2.0982
15	PAPER CHAIN KIT 50'S CHRISTMAS	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.0230	0.4015	9.6157	0.0206	1.6012

Sumber : Hasil Analisis Menggunakan Jupyter Notebook dan Python (2025)

4.4 Interpretasi Hasil dan Implikasi Bisnis

Berdasarkan analisis association rules yang dihasilkan, terdapat beberapa pola pembelian yang dapat diidentifikasi dan dikelompokkan menurut kategori produk.

4.4.1 Kelompok Produk Teacup & Saucer

Pasangan Roses Regency Teacup and Saucer dan Green Regency Teacup and Saucer menghasilkan lift tertinggi sebesar 22,2670 dengan confidence 70,36%. Artinya, pelanggan yang membeli salah satu varian cangkir ini memiliki kecenderungan sangat tinggi untuk juga membeli pasangannya. Nilai lift yang jauh di atas 1 mengkonfirmasi bahwa hubungan ini bukan kebetulan melainkan pola yang sangat konsisten. Rekomendasi bisnis: kedua produk ini sangat ideal untuk dijadikan bundle tetap dengan harga paket yang menarik.

4.4.2 Kelompok Lunch Box Anak

Spaceboy Lunch Box dan Dolly Girl Lunch Box menghasilkan lift 17,9440 dengan confidence 60,65%. Pola ini mengindikasikan kemungkinan pelanggan membeli untuk dua anak sekaligus (desain berbeda gender) atau membeli sebagai hadiah ganda. Rekomendasi: tampilkan keduanya secara berdampingan di halaman produk online disertai penawaran "Beli 2 Lebih Hemat".

4.4.3 Kelompok Jam Weker

Alarm Clock Bakelike Red dan Alarm Clock Bakelike Green memiliki lift 17,4052 dengan confidence di atas 61%. Pola ini menunjukkan pelanggan mengoleksi keduanya karena tertarik pada seri produk yang sama. Rekomendasi: terapkan strategi bundle warna atau "Complete the Set" pada halaman produk.

4.4.4 Kelompok Trinket Box

Sweetheart Ceramic Trinket Box dan Strawberry Ceramic Trinket Box dengan lift 11,3316 dan confidence 69,77% menunjukkan pola koleksi seri. Confidence hampir 70% mengindikasikan bahwa 7 dari 10 pelanggan yang membeli satu varian juga membeli varian lainnya.

4.4.5 Kelompok Lunch Bag Multi-Varian

Terdapat 10 aturan asosiasi yang melibatkan berbagai varian Lunch Bag (Black Skull, Spaceboy Design, Cars Blue, Woodland, Pink Polkadot, Red Retrosport), semuanya dengan lift di atas 6,0. Ini mengindikasikan kebiasaan pelanggan membeli beberapa varian desain sekaligus, kemungkinan sebagai hadiah atau untuk keperluan berbeda. Rekomendasi: buat halaman kategori "Lunch Bag Collection" dengan opsi multi-pilih dan diskon bertingkat.

Salah satu aturan terkuat pada kelompok ini adalah hubungan antara Black Skull Lunch Bag dan Spaceboy Design Lunch Bag dengan nilai lift sebesar 6.1068 dan confidence sebesar 0.4130, yang menunjukkan kecenderungan pelanggan membeli kedua varian desain tersebut secara bersamaan.

4.4.6 Kelompok Jumbo Bag

Pasangan Jumbo Bag Pink Polkadot dan Jumbo Bag Red Retrosport menghasilkan lift 7,1148 dengan confidence 62,94%, menunjukkan pola pembelian koleksi tas serupa. Rekomendasi: tampilkan seluruh seri Jumbo Bag dalam satu halaman koleksi dengan bundle pilihan.

4.4.7 Kelompok Aksesori Rumah

Wooden Picture Frame White Finish dan Wooden Frame Antique White menghasilkan lift 9,9741, mengindikasikan pelanggan membeli bingkai foto dalam set untuk dekorasi. Red Hanging Heart T-Light Holder dan White Hanging Heart T-Light Holder dengan confidence 70,69% dan lift 4,2704 menunjukkan pola pembelian set dekorasi pemegang lampu teh berwarna berbeda.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengimplementasikan algoritma Apriori dalam Market Basket Analysis pada dataset Online Retail II yang mencakup 776.872 transaksi valid setelah proses pembersihan data. Dengan parameter minimum support 2%, minimum confidence 30%, dan minimum lift 1,0, ditemukan 148 frequent itemsets dan 44 aturan asosiasi yang bermakna secara bisnis.

Temuan utama penelitian menunjukkan bahwa kombinasi produk terkuat untuk mendukung strategi bundling adalah: (1) Roses Regency Teacup and Saucer dengan Green Regency Teacup and Saucer (lift: 22,27); (2) Spaceboy Lunch Box dengan Dolly Girl Lunch Box (lift: 17,94); (3) Alarm Clock Bakelike Red dengan Alarm Clock Bakelike Green (lift: 17,41); (4) Sweetheart Ceramic Trinket Box dengan Strawberry Ceramic Trinket Box (lift: 11,33); dan (5) berbagai kombinasi varian Lunch Bag (lift: 6,10–10,97). Secara keseluruhan, rata-rata lift sebesar 8,79 mengindikasikan kualitas aturan asosiasi yang tinggi dan relevan secara bisnis.

Berdasarkan temuan tersebut, disarankan beberapa tindakan strategis berikut. Pertama, desain paket bundling tetap untuk pasangan produk dengan lift tertinggi, terutama kelompok Teacup & Saucer, Lunch Box anak, dan Alarm Clock, dengan harga paket yang memberikan insentif diskon. Kedua, implementasi sistem rekomendasi "Frequently Bought Together" pada platform e-commerce yang secara otomatis menampilkan produk pasangan berdasarkan aturan asosiasi yang ditemukan. Ketiga, penataan ulang layout toko atau urutan tampilan katalog online untuk menempatkan produk-produk dengan asosiasi kuat secara berdekatan. Keempat, pengembangan kampanye pemasaran bertema "Complete the Collection" yang memanfaatkan pola koleksi seri produk yang teridentifikasi.

Untuk penelitian selanjutnya, disarankan untuk membandingkan performa algoritma Apriori dengan FP-Growth pada dataset yang sama, mengeksplorasi segmentasi analisis berdasarkan negara atau periode waktu, serta mengintegrasikan hasil MBA dengan analisis RFM (Recency, Frequency, Monetary) untuk mendapatkan profil pelanggan yang lebih komprehensif.

Penelitian selanjutnya juga dapat menguji sensitivitas hasil dengan menggunakan variasi nilai minimum support dan minimum confidence yang berbeda, misalnya minimum support sebesar 1% atau 3%, untuk mengevaluasi stabilitas pola asosiasi yang dihasilkan.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada dosen pembimbing yang telah memberikan arahan selama proses penelitian, Universitas Pamulang yang telah mendukung pelaksanaan penelitian, serta penyedia dataset Online Retail II melalui platform Kaggle.

DAFTAR PUSTAKA

- [1] R. Dalayah, "Analisis perilaku konsumsi pengguna aplikasi e-money pada mahasiswa Fakultas Ekonomi Universitas Negeri Surabaya," *Jurnal Pendidikan Tata Niaga (JPTN)*, vol. 8, no. 3, pp. 946–952, 2020. <https://doi.org/10.26740/jptn.v8n3.p946-952>

- [2] A. Musalem, L. Aburto, and M. Bosch, "Market basket analysis insights to support category management," *European Journal of Marketing*, vol. 52, no. 7/8, pp. 1550–1573, 2018.
<https://doi.org/10.1108/EJM-06-2017-0367>
- [3] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proc. 20th Int. Conf. Very Large Data Bases (VLDB)*, Santiago, Chile, 1994, pp. 487–499.
Available: <https://www.vldb.org/conf/1994/P487.PDF>
- [4] R. Garg and P. Gulia, "Comparative study of frequent itemset mining algorithms apriori and FP growth," *International Journal of Computer Applications*, vol. 126, no. 4, pp. 8–12, 2015.
<https://doi.org/10.5120/ijca2015906030>
- [5] M. Kaur and S. Kang, "Market basket analysis: Identify the changing trends of market data using association rule mining," *Procedia Computer Science*, vol. 85, pp. 78–85, 2016.
<https://doi.org/10.1016/j.procs.2016.05.180>
- [6] G. Khushi, K. Shah, and A. A. Kadam, "E-commerce market basket analysis using apriori algorithm," *International Journal of Scientific Research in Engineering and Management (IJSREM)*, 2023.
- [7] N. Karimtabar and M. J. Shayegan Fard, "Finding Frequent Items: A Novel Method for Improving the Apriori Algorithm," *Computer Science*, vol. 23, no. 2, pp. 161–177, 2022.
<https://doi.org/10.7494/csci.2022.23.2.3776>
- [8] M. Hossain, A. S. Sattar, and M. K. Paul, "Market basket analysis using apriori and FP-growth algorithm," in *2019 22nd Int. Conf. Computer and Information Technology (ICCIT)*, Dhaka, 2019, pp. 1–6.
<https://doi.org/10.1109/ICCIT48885.2019.9038197>
- [9] E. Umar, D. Manongga, and A. Iriani, "Market basket analysis menggunakan association rule dan algoritma apriori pada produk penjualan mitra swalayan Salatiga," *Jurnal Media Informatika Budidarma*, vol. 6, no. 3, pp. 1367–1377, 2022.
<https://doi.org/10.30865/mib.v6i3.4217>
- [10] A. I. Idris et al., "Comparison of apriori, apriori-TID and FP-growth algorithms in market basket analysis at grocery stores," *The IJICS (International Journal of Informatics and Computer Science)*, vol. 6, no. 2, 2022.
<https://doi.org/10.30865/ijics.v6i2.4535>
- [11] I. Qoniah and A. T. Priandika, "Analisis market basket untuk menentukan asosiasi rule dengan algoritma apriori (Studi Kasus: Tb. Menara)," *Jurnal Teknologi Dan Sistem Informasi*, vol. 1, no. 2, pp. 26–33, 2020.
<https://doi.org/10.33365/jtsi.v1i2.368>

NOMENKLATUR

1. A: Himpunan item antecedent (item yang dibeli terlebih dahulu)
2. B: Himpunan item consequent (item yang dibeli setelahnya)
3. $A \cup B$: Gabungan himpunan item A dan B dalam satu transaksi
4. $|A \cup B|$: Jumlah transaksi yang mengandung item A dan B secara bersamaan
5. N: Total jumlah seluruh transaksi dalam dataset
6. Support ($A \rightarrow B$): Nilai support dari aturan asosiasi A terhadap B
7. Confidence ($A \rightarrow B$): Nilai confidence dari aturan asosiasi A terhadap B
8. Lift ($A \rightarrow B$): Nilai lift dari aturan asosiasi A terhadap B
9. K: Panjang itemset (jumlah item dalam satu himpunan)
10. min_sup: Nilai ambang batas minimum support yang ditetapkan
11. min_conf: Nilai ambang batas minimum confidence yang ditetapkan