

Penerapan Algoritma Apriori Untuk Menentukan Pola Pembelian Konsumen

Diterima:
19 Juni 2025

Revisi:
19 Juli 2019

Terbit:
25 Desember 2025

^{a*}Nazzel Maulana Mustofa, ^aAhmad Muharram Alfarisi, ^bAbu Tholib

^aUniversitas Nusantara PGRI Kediri

^bUniversitas Nurul Jadid

Abstrak— Latar Belakang: Di era digital, bisnis ritel menghadapi tantangan dalam memahami perilaku konsumen dan menyusun strategi pemasaran yang efektif. Market Basket Analysis (MBA) menjadi pendekatan populer untuk menganalisis pola pembelian konsumen guna mempertahankan daya saing. **Tujuan:** Menemukan pola pembelian pelanggan dan mengidentifikasi aturan asosiasi antar produk yang dapat dimanfaatkan dalam strategi pemasaran seperti penempatan produk, bundling, dan personalisasi. **Metode:** Algoritma Apriori pada dataset transaksi ritel dari Kaggle yang berisi lebih dari 90.000 entri. Data dianalisis setelah melalui tahap pra-pemrosesan dan transformasi dengan teknik one-hot encoding. Algoritma dijalankan dengan parameter minimum support 0,005 dan confidence 0,5. **Hasil:** Hasil menunjukkan bahwa produk “12V U1 L&G 6” memiliki nilai support tertinggi sebesar 2,92%. Pasangan produk “1.5V IND AAA ALK BULK” dan “1.5V IND AA ALK BULK” menunjukkan asosiasi kuat dengan confidence 68,9% dan lift 58,46%. **Kesimpulan:** Penelitian ini berhasil mengidentifikasi pola pembelian konsumen dan menghasilkan aturan asosiasi yang signifikan sebagai dasar strategi pemasaran berbasis data. Penelitian selanjutnya disarankan mengeksplorasi algoritma lain seperti FP-Growth atau Eclat untuk membandingkan efisiensi dan akurasi.

Kata kunci—Market Basket Analysis; Algoritma Apriori; Aturan Asosiasi; Pola Pembelian Konsumen

Abstract— Background Urgency: In today's digital era, retail businesses face significant challenges in understanding customer behavior and formulating effective marketing strategies. Market Basket Analysis (MBA) is a widely used data-driven approach to uncover purchasing patterns and support competitive advantage. **Objective:** The objective of this study is to discover customer purchasing patterns and identify product association rules that can be applied in marketing strategies such as product placement, bundling, and personalization. This study employs a quantitative approach using the Apriori algorithm on a retail transaction dataset sourced from Kaggle, consisting of over 90,000 entries. The data underwent preprocessing and transformation using one-hot encoding before analysis. The algorithm was applied with a minimum support of 0.005 and a confidence threshold of 0.5. **Results:** The results reveal that the product “12V U1 L&G 6” has the highest support value at 2.92%. Additionally, the product pair “1.5V IND AAA ALK BULK” and “1.5V IND AA ALK BULK” demonstrated a strong association with a confidence of 68.9% and a lift of 58.46%. **Conclusion:** Conclusion: This study successfully identifies significant consumer purchasing patterns and association rules, offering valuable insights for developing data-driven marketing strategies. Future research is encouraged to explore alternative algorithms such as FP-Growth or Eclat for comparative performance analysis..

Keywords—Market Basket Analysis; Apriori Algorithm; Association Rules; Consumer Purchase Patterns

This is an open access article under the CC BY-SA License.



Penulis Korespondensi:

Nazzel Maulana Mustofa,
Sistem Informasi,
Universitas Nusantara PGRI Kediri,
Email: nazzelmustofa@gmail.com

I. PENDAHULUAN

Toko ritel melakukan banyak penjualan, sehingga mereka mengumpulkan lebih banyak data transaksi, yang dapat digunakan untuk meningkatkan bisnis mereka. Di era informasi saat ini, proses data mining dapat digunakan untuk mengolah data. Industri ritel elektronik menghadapi persaingan yang semakin ketat dan perilaku konsumen yang berubah-ubah, seperti yang ditunjukkan oleh banyaknya toko ritel yang berdiri sendiri. Toko ritel elektronik harus memahami pola pembelian pelanggan mereka dengan baik jika mereka ingin tetap relevan dan berhasil dalam lingkungan bisnis yang kompetitif ini. Metode analisis pasar adalah salah satu cara efektif untuk mengetahui pola pembelian konsumen (Brighton & Hariyanto, 2024).

Untuk mencapai hal ini, proses eksplorasi data harus dilakukan, yang dapat menghasilkan informasi berguna. Data mining adalah bidang ilmu dan teknik yang sangat cocok untuk masalah seperti ini. Ini adalah proses mengolah kumpulan data untuk menghasilkan informasi atau pengetahuan baru. Diharapkan bahwa penggunaan teknik ini akan membantu dan mempercepat proses pengambilan keputusan. (Rifania et al., 2023). Metode data mining digunakan untuk mengekstrak informasi penting dari kumpulan data besar. (Sari & Khoiriah, 2022). Data mining adalah proses menemukan hubungan pola dan kecenderungan dalam sejumlah besar data yang tersimpan dalam penyimpanan dengan memeriksanya dengan menggunakan teknik pengenalan pola seperti matematika dan statistik. Pengelompokan data mining adalah teknik yang didasarkan pada tugas yang dapat dilakukan (Nur Rahmi & Yosaphat Ananda Mikola, 2021).

Algoritma Apriori merupakan algoritma yang umum digunakan pada Market Basket Analysis (MBA) karena kemampuannya mengidentifikasi aturan asosiasi yang relevan dalam data berukuran besar (Safitry et al., 2024). Dalam data mining, algoritma Apriori termasuk dalam kategori aturan asosiasi. Pada tahun 1994, Agrawal dan Srikant mengusulkannya sebagai algoritma dasar untuk menemukan itemset sering dalam konteks aturan asosiasi boolean. (Oktaviani, 2024). Tujuan dari algoritma apriori ini adalah untuk memberikan informasi mengenai hubungan antar basis data transaksi barang. Proses menganalisis dan memverifikasi data yang disimpan dalam teknik pengenalan pola dikenal dengan data mining. Dalam hal permasalahan dalam penyediaan barang, data mining adalah salah satu bentuk solusi. Data mining adalah metode yang dapat digunakan oleh bisnis untuk mengembangkan strategi bisnis (Fitriana Dewi et al., 2022).

Studi sebelumnya oleh Ghofur et al. menggunakan algoritma apriori untuk menganalisis transaksi penjualan dengan tujuan mengidentifikasi kecenderungan seorang pembeli untuk membeli barang yang memiliki hubungan dalam berbelanja. Dalam penelitian ini, Ghofur et al. mencari kombinasi 2-itemset dan 3-itemset dari 1-itemset yang telah dinyatakan lolos dari

minimum support. Setelah menemukan pola frekuensi tinggi dari kombinasi ini, dibentuk aturan asosiatif untuk nilai minimum keyakinan. Hasil pengujian yang dilakukan oleh peneliti dengan menggunakan algoritma apriori pada sistem yang dibuat menghasilkan aturan asosiatif dua pola pada kombinasi dua itemset dengan nilai minimum dukungan 40% dan nilai minimum keyakinan 30%. Pengujian ini dilakukan pada sepuluh data transaksi penjualan. (Ghofur et al., 2020).

Dalam penelitian lain yang dilakukan oleh Rusdianto et al., mereka menemukan pola belanja konsumen untuk merancang ulang tata letak barang dengan menggunakan data historis transaksi penjualan. Penelitian ini menggunakan data sebanyak 125 baris dengan support dan confidence ditetapkan masing-masing 3% dan 20%, dan aturan asosiasi yang terbentuk ditunjukkan sebagai saran untuk menentukan tata letak barang. Hasil penelitian menunjukkan bahwa tata letak barang harus diubah. (Rusdianto et al., 2020).

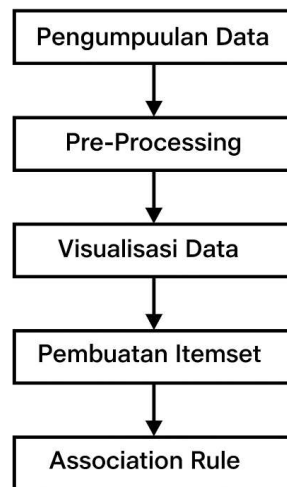
Penelitian lebih lanjut oleh Sibarani et al. menggunakan algoritma apriori untuk mengidentifikasi hubungan antar item dalam dataset, atau sekumpulan data, yang telah ditentukan sebelumnya. Dengan menggunakan teknik aturan hubungan, para peneliti dapat menemukan kombinasi yang mungkin dari dataset item yang sering. Peneliti menggunakan data Apotik Pusaka Arta dengan nilai minimum dukungan 20% dan nilai minimum keyakinan 50%. Peneliti menemukan 7 kombinasi aturan asosiasi. (Sibarani, 2020)

Namun, penelitian-penelitian tersebut masih terbatas pada skala lokal dengan data transaksi yang terbatas, serta belum memanfaatkan dataset besar yang tersedia secara terbuka. Selain itu, penerapannya belum secara eksplisit diarahkan pada strategi bundling dan sistem rekomendasi produk dalam toko ritel elektronik berbasis data besar. Perbedaan penelitian ini dengan penelitian sebelumnya adalah penggunaan dataset transaksi berskala besar dari situs Kaggle, yang terdiri dari lebih dari 90.000 entri, serta fokus penelitian yang tidak hanya mencari asosiasi antar produk, tetapi juga mengeksplorasi penerapannya untuk menyusun strategi bundling dan penempatan produk secara *data-driven* di lingkungan toko ritel elektronik. Penelitian ini berusaha menjawab tantangan pengelolaan produk dan perilaku belanja yang kompleks dengan pendekatan analitis berbasis algoritma Apriori.

Tujuan penelitian ini adalah untuk menganalisis dataset transaksi ritel yang diperoleh dari Kaggle guna menemukan pola pembelian konsumen dan menghasilkan aturan asosiasi antar produk yang dapat dijadikan dasar dalam strategi pemasaran, penataan produk, serta pengembangan sistem rekomendasi yang mendukung peningkatan penjualan pada toko ritel.

II. METODE

Metode yang digunakan untuk mengolah dataset *market basket analysis* yaitu algoritma apriori, dengan dataset yang di peroleh dari *kegg*le. dataset tersebut diolah kemudian digunakan untuk menganalisis pola pembelian konsumen yang nantinya akan digunakan untuk menentukan.item yang sering dibeli, sehingga dapat digunakan untuk meningkatkan strategi penjualan. Ada empat tahapan yang digunakan dalam proses ini, yaitu pengumpulan data, pra-pemrosesan data, transformasi data, dan penafsiran data (Gambar 1). Tahapan penafsiran termasuk analisis pola frekuensi tinggi, pembentukan aturan hubungan, dan pengujian lift (Purwati & Karnila, 2023):



Gambar 1. Alur Penelitian (Fathurrahman et al., 2023)

1. Pengumpulan Data

Proses pengumpulan data dalam penelitian ini dilakukan secara online dengan menggunakan dataset publik yang tersedia di situs Kaggle. Ini dilakukan karena kualitas dan kelengkapan data yang dikumpulkan akan memengaruhi hasil akhir analisis. Pilihan Kaggle sebagai sumber data didasarkan pada fakta bahwa platform ini menyediakan berbagai dataset yang telah digunakan secara luas dalam proyek penelitian akademik dan analisis data, serta telah melalui proses kurasi awal. Dataset yang digunakan berasal dari sektor ritel elektronik, yang mengandung 92.249 catatan transaksi. Lima komponen utama merupakan atribut atau variabel dalam dataset ini, yaitu:

1. Product
2. Product Category
3. Transaction ID
4. Unique Purchases
5. Product Revenue

Data ini dikumpulkan dengan tujuan untuk mempelajari pola pembelian konsumen dan menggunakan algoritma Apriori untuk mempersiapkan informasi yang relevan untuk analisis

lebih lanjut. Sebagaimana dijelaskan oleh (Fajri et al., 2025), untuk dapat secara akurat menunjukkan perilaku pembelian pelanggan, pengumpulan data harus mempertimbangkan kelengkapan fitur transaksi. Oleh karena itu, dataset yang dipilih dianggap sesuai dan representatif untuk tujuan penelitian ini (Gambar 2).

	Product	Product Category (Enhanced Ecommerce)	Transaction ID	Unique Purchases	Product Revenue
0	3.7V 3400mah LIION 12.6WH	Battery/Consumer Rechargeable	EC0043605902	47	\$1,597.53
1	3V PHOTO LITHIUM	Battery/Primary Other	EC0043507670	47	\$1,246.44
2	12V 11.2AH 225CCA AGM 12/0	Battery/Powersports	EC0043504182	41	\$4,714.59
3	12V 12AH 165CCA FLOODED 6/0	Battery/Powersports	EC0043503186	39	\$2,456.61
4	12V 12AH 210CCA AGM 12/0	Battery/Powersports	EC0043406547	34	\$3,570.00
...
92245	ZUS SMART VEHICLE HEALTH MONITOR MINI	Bulb - Controls/Smart Home	EC0043404746	1	\$9.97
92246	ZUS SMART VEHICLE HEALTH MONITOR MINI	Bulb - Controls/Smart Home	EC0043803104	1	\$15.99
92247	ZUS SMART VEHICLE HEALTH MONITOR MINI	Bulb - Controls/Smart Home	EC0043804320	1	\$15.99
92248	ZUS SMART VEHICLE HEALTH MONITOR MINI	Bulb - Controls/Smart Home	EC0044001978	1	\$15.99
92249	ZUS UNIVERSAL CAR AUDIO ADAPTER	Bulb - Controls/Smart Home	EC0043708571	1	\$41.99

Gambar 2. Gambar Tabel Dataset dari Kaggle

2. Pra-pemrosesan Data

Sebelum data dapat dianalisis lebih lanjut dengan algoritma Apriori, tahap pra-pemrosesan data sangat penting. Pada tahap ini, data dibersihkan dan formatnya disesuaikan untuk menjadi konsisten, terorganisir, dan siap untuk digunakan dalam proses analisis asosiasi. Menurut (Seniwati et al., 2024), preprocessing adalah langkah penting dalam siklus pengolahan data karena bertujuan untuk menghilangkan data duplikat, mengatasi nilai kosong, dan memastikan bahwa semua variabel dalam kumpulan data memiliki format dan tipe data yang sesuai dengan algoritma yang akan digunakan. Proses pra-pemrosesan ini bertujuan untuk memastikan bahwa data yang digunakan bersih, bebas dari kesalahan struktural, dan memiliki integritas yang baik. Ini memungkinkan algoritma Apriori untuk meningkatkan akurasi dan keandalan hasil analisis. Tahap ini harus dilewati karena kemungkinan kesalahan dalam proses pembentukan itemsets sering dan aturan asosiasi akan meningkat. Data hasil pemrosesan kemudian diubah menjadi format yang sesuai dengan kebutuhan analisis. Misalnya data diubah sesuai dengan format algoritma apriori sehingga dapat menemukan pola pembelian konsumen dan barang yang sering dibeli.

3. Visualisasi data

Data yang telah dikumpulkan akan divisualisasikan menggunakan grafik atau chart. Tujuan utama dari visualisasi ini adalah untuk mempermudah pemahaman data dan memperkuat interpretasi hasil yang diperoleh dari algoritma yang digunakan. Melalui visualisasi, data disajikan dalam format yang lebih mudah dicerna oleh audiens. Ini memungkinkan pengguna untuk mengidentifikasi pola, tren, dan hubungan yang mungkin sulit atau bahkan tidak terlihat jika data hanya disajikan dalam bentuk mentah atau tidak terstruktur. Dengan demikian,

visualisasi data berperan krusial dalam mengubah data kompleks menjadi wawasan yang jelas dan mudah dipahami.

4. Pembentukan *Itemset*

Dalam analisis data transaksi, tahap pembentukan itemset merupakan tahap penting di mana sistem mengelompokkan kombinasi item yang sering muncul secara bersamaan dalam satu transaksi. Tujuannya adalah untuk menemukan hubungan yang kuat antara barang dan jasa yang dibeli oleh pelanggan.. Dengan menemukan kombinasi item yang sering terjadi, peneliti atau analis dapat mengenali pola perilaku konsumen yang berulang. Informasi ini sangat bermanfaat dalam mendukung pengambilan keputusan strategis, seperti penataan produk di rak, rekomendasi produk, atau penentuan promosi yang efektif. Pembentukan itemset menjadi dasar utama sebelum diterapkannya algoritma untuk menemukan aturan asosiasi dalam data.

5. *Association Rule*

Konsep menarik dalam data mining adalah aturan asosiasi untuk menemukan keterkaitan atau asosiasi antar data (Ulfha & Amin, 2020). Tahapan ini bertujuan untuk mengidentifikasi hubungan atau keterkaitan antar item dalam dataset transaksi. Aturan asosiasi menghasilkan aturan yang menunjukkan bagaimana satu item berhubungan dengan item lainnya berdasarkan seberapa sering mereka bertemu dalam suatu transaksi.. Proses ini membantu mengungkap wawasan tersembunyi dalam data yang dapat digunakan untuk memahami perilaku konsumen serta mendukung pengambilan keputusan yang lebih tepat. Sebagai contoh, aturan asosiasi dapat menunjukkan bahwa pelanggan yang membeli produk A juga cenderung membeli produk B. Dalam penelitian ini, algoritma Apriori digunakan sebagai metode utama untuk menemukan aturan asosiasi tersebut secara sistematis dan efektif. Dengan mempertimbangkan minimum dukungan dan minimum keyakinan, algoritma Apriori menentukan kandidat yang mungkin muncul (Saefudin & DN, 2019).

III. HASIL DAN PEMBAHASAN

Berdasarkan format data transaksi penjualan yang telah diperoleh, dilakukan proses analisis asosiasi menggunakan algoritma Apriori untuk menemukan pola hubungan antar produk dalam satu transaksi. Analisis ini diawali dengan menentukan nilai support, yaitu persentase jumlah transaksi yang memuat kombinasi dua item tertentu, misalnya item A dan B. Dalam konteks ini, apabila konsumen membeli produk A, maka terdapat kemungkinan besar bahwa mereka juga akan membeli produk B. Nilai support tersebut menjadi indikator awal seberapa sering pasangan item tersebut muncul bersamaan dalam keseluruhan data transaksi.

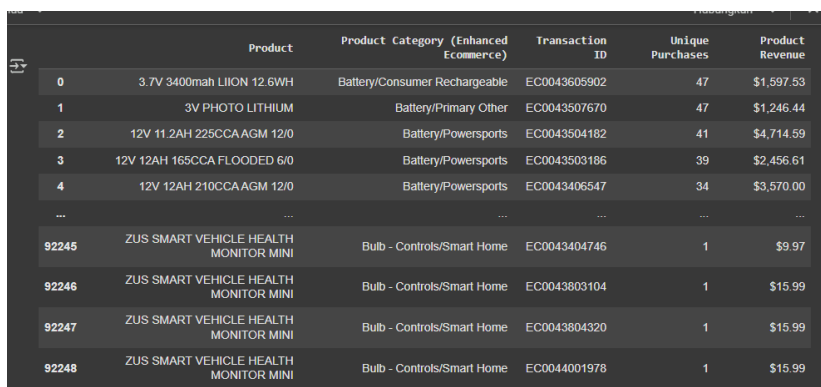
Proses pengolahan data dimulai dengan mengimpor file berformat CSV yang berisi data transaksi ritel sebanyak 99.200 baris. Setelah proses impor selesai, tahap selanjutnya adalah

melakukan encoding untuk merapikan data dan mengubahnya ke dalam format biner (one-hot encoding), sehingga setiap transaksi dapat direpresentasikan dalam bentuk kombinasi item yang siap untuk diproses lebih lanjut (Andy Hermawan et al., 2024). Analisis asosiasi didefinisikan sebagai suatu proses dalam data mining yang bertujuan untuk menemukan semua aturan yang memenuhi batas minimum support dan confidence. Confidence sendiri mengukur seberapa besar kemungkinan item B dibeli ketika item A telah dibeli sebelumnya, atau dengan kata lain, menunjukkan tingkat kepercayaan pada hubungan antara dua produk.

a. Tabel Data

Data transaksi yang digunakan dalam penelitian ini diperoleh dari dataset 99.200 baris data transaksi yang disediakan oleh platform Kaggle.. Setiap baris data merepresentasikan satu transaksi unik dan mencakup lima atribut utama, yaitu:

1. Product – nama produk yang dibeli,
2. Product Category (Enhanced Ecommerce) – kategori produk berdasarkan klasifikasi e-commerce,
3. Transaction ID – nomor unik dari setiap transaksi,
4. Unique Purchases – jumlah pembelian unik per produk, dan
5. Product Revenue – total pendapatan dari penjualan produk dalam transaksi tersebut.



	Product	Product Category (Enhanced Ecommerce)	Transaction ID	Unique Purchases	Product Revenue
0	3.7V 3400mah LIION 12.6WH	Battery/Consumer Rechargeable	EC0043605902	47	\$1,597.53
1	3V PHOTO LITHIUM	Battery/Primary Other	EC0043507670	47	\$1,246.44
2	12V 11.2AH 225CCA AGM 12/0	Battery/Powersports	EC0043504182	41	\$4,714.59
3	12V 12AH 165CCA FLOODED 6/0	Battery/Powersports	EC0043503186	39	\$2,456.61
4	12V 12AH 210CCA AGM 12/0	Battery/Powersports	EC0043406547	34	\$3,570.00
...
92245	ZUS SMART VEHICLE HEALTH MONITOR MINI	Bulb - Controls/Smart Home	EC0043404746	1	\$9.97
92246	ZUS SMART VEHICLE HEALTH MONITOR MINI	Bulb - Controls/Smart Home	EC0043803104	1	\$15.99
92247	ZUS SMART VEHICLE HEALTH MONITOR MINI	Bulb - Controls/Smart Home	EC0043804320	1	\$15.99
92248	ZUS SMART VEHICLE HEALTH MONITOR MINI	Bulb - Controls/Smart Home	EC0044001978	1	\$15.99

Gambar 3. Tabel Data Transaksi (sumber google colab)

Gambar 3 menunjukkan cuplikan sebagian data yang telah berhasil dimuat menggunakan perintah `pd.read_csv()` di Google Colab. Tampak bahwa data memiliki struktur yang baik, dengan format tabel yang rapi dan kolom-kolom yang relevan untuk analisis asosiasi. Namun, sebelum data dapat digunakan dalam proses analisis lebih lanjut, diperlukan proses pembersihan awal. Langkah selanjutnya yang dilakukan adalah pencarian data yang hilang (missing value) dan penghapusan data duplikat guna memastikan kualitas dan validitas data yang akan digunakan dalam algoritma Apriori. Tahapan ini penting untuk mencegah hasil analisis yang bias atau tidak akurat.

b. Data transformation

Untuk memastikan bahwa data dapat digunakan dalam proses analisis, tahap perencanaan data mencakup pembersihan dan transformasi data. Kolom yang tidak digunakan dan data transaksi yang duplikat dihapus. Untuk analisis basket pasar, data transaksi diubah ke format encoding satu-hot, dengan setiap produk diwakili sebagai kolom biner..

```
[ ] df['Product Revenue'] = df['Product Revenue'].replace(['\$'], '', regex=True).astype(float)
print("Missing values per kolom:\n", df.isnull().sum())

# (Optional) Kalau mau hapus baris dengan nilai kosong:
df = df.dropna()
```

Missing values per kolom:

Product	0
Product Category (Enhanced Ecommerce)	0
Transaction ID	0
Unique Purchases	0
Product Revenue	0

dtype: int64

Gambar 4. Hasil Pemeriksaan *Missing Value* (sumber google colab)

Tahap Data Preparation merupakan langkah penting dalam proses data mining karena berfungsi untuk memastikan bahwa data yang akan dianalisis benar-benar bersih, konsisten, dan relevan. Setelah data transaksi berhasil dimuat, dilakukan pengecekan terhadap missing value pada seluruh kolom, seperti Product, Product Category, Transaction ID, Unique Purchases, dan Product Revenue. Hasil pemeriksaan menunjukkan bahwa tidak terdapat nilai kosong pada seluruh kolom (Gambar 4), sehingga tidak diperlukan proses imputasi atau penghapusan baris berdasarkan null value.

```
# Hitung IQR
Q1 = df['Product Revenue'].quantile(0.25)
Q3 = df['Product Revenue'].quantile(0.75)
IQR = Q3 - Q1

# Tentukan batas bawah dan atas
batas_bawah = Q1 - 1.5 * IQR
batas_atas = Q3 + 1.5 * IQR

# Tampilkan data yang termasuk outlier
outliers = df[(df['Product Revenue'] < batas_bawah) | (df['Product Revenue'] > batas_atas)]
print("Outliers:\n", outliers)

# Hapus outlier dari dataframe
df_clean = df[(df['Product Revenue'] >= batas_bawah) & (df['Product Revenue'] <= batas_atas)]
```

Gambar 5. Perhitungan Metode *Interquartile Range (IQR)* (sumber google colab)

Selanjutnya, dilakukan deteksi dan penghapusan outlier pada kolom Product Revenue untuk meningkatkan kualitas data sebelum dianalisis lebih lanjut (Gambar 5). *Interquartile Range (IQR)*, yang menghitung rentang antara kuartil pertama (Q1) dan ketiga (Q3), digunakan sebagai dasar untuk menentukan batas atas dan bawah dari data yang dianggap wajar.. Nilai-nilai yang berada di bawah batas bawah atau di atas batas atas dikategorikan sebagai outlier, karena dianggap menyimpang secara signifikan dari pola umum distribusi data. Outlier-outlier tersebut kemudian dikeluarkan dari dataset agar tidak memengaruhi hasil analisis algoritma Apriori, yang sangat bergantung pada pola frekuensi item dalam transaksi. Proses ini penting dilakukan karena keberadaan outlier dapat menyebabkan hasil analisis menjadi bias dan kurang akurat dalam merepresentasikan perilaku konsumen secara umum.

Outliers:

	Product \
0	3.7V 3400mah LION 12.6WH
1	3V PHOTO LITHIUM
2	12V 11.2AH 225CCA AGM 12/0
3	12V 12AH 165CCA FLOODED 6/0
4	12V 12AH 210CCA AGM 12/0
...	...
92210	YETI 1500X PORTABLE POWER STATION
92212	YETI 3000X PORTABLE POWER STATION
92214	YETI 500X PORTABLE POWER STATION
92215	YETI 500X PORTABLE POWER STATION
92216	YETI 500X PORTABLE POWER STATION

	Product Category (Enhanced Ecommerce)	Transaction ID	Unique Purchases \
0	Battery/Consumer Rechargeable	EC0043605902	47
1	Battery/Primary Other	EC0043507670	47
2	Battery/Powersports	EC0043504182	41
3	Battery/Powersports	EC0043503186	39
4	Battery/Powersports	EC0043406547	34
...
92210	Device/Power Equipment	EC0043708027	1
92212	Device/Power Equipment	EC0043006693	1
92214	Device/Power Equipment	EC0043203604	1
92215	Device/Power Equipment	EC0044001964	1

Gambar 6. Hasil Perhitungan Metode *Interquartile Range (IQR)*

Beberapa produk seperti "3.7V 3400mah LION 12.6WH" dan "YETI 500X PORTABLE POWER STATION" teridentifikasi sebagai outlier karena memiliki nilai Product Revenue yang jauh lebih tinggi dibandingkan dengan mayoritas produk lainnya dalam dataset (Gambar 6). Nilai pendapatan yang ekstrem ini berpotensi memengaruhi hasil analisis pola pembelian karena dapat mendistorsi distribusi data dan memberikan gambaran yang kurang representatif terhadap tren umum konsumen. Produk-produk tersebut mungkin merupakan produk premium dengan harga tinggi dan frekuensi pembelian yang rendah, sehingga tidak mencerminkan kebiasaan belanja mayoritas pelanggan. Oleh karena itu, penghapusan data outlier seperti ini menjadi langkah penting dalam proses data cleaning, agar hasil analisis dengan algoritma Apriori dapat lebih fokus pada item yang memiliki kontribusi signifikan dan relevan secara statistik terhadap pola pembelian yang sering terjadi.

Data setelah cleaning dan tanpa outlier:

	Product \
16	2PK 500/CL/MC
18	CORD BYTES UNICORN/PANDA
22	3.2V 600MAH LIFEPO4 1.5WH
24	OEM replacement battery for baby monitors
30	2PK 9003 POWERSVISION
...	...
92245	ZUS SMART VEHICLE HEALTH MONITOR MINI
92246	ZUS SMART VEHICLE HEALTH MONITOR MINI
92247	ZUS SMART VEHICLE HEALTH MONITOR MINI
92248	ZUS SMART VEHICLE HEALTH MONITOR MINI
92249	ZUS UNIVERSAL CAR AUDIO ADAPTER

	Product Category (Enhanced Ecommerce)	Transaction ID	Unique Purchases \
16	Bulb/Miniature	EC0043701059	15
18	Accessory/Plus Products	EC0042808560	15
22	Battery/Consumer Rechargeable	EC0043400239	13
24	Battery/HHD	EC0042703152	13
30	Bulb/Auto Bulb	EC0043503446	11
...
92245	Bulb - Controls/Smart Home	EC0043404746	1
92246	Bulb - Controls/Smart Home	EC0043803104	1
92247	Bulb - Controls/Smart Home	EC0043804320	1
92248	Bulb - Controls/Smart Home	EC0044001978	1

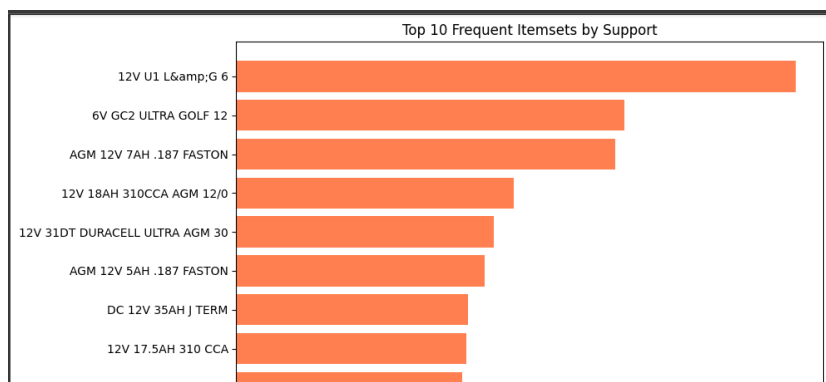
Gambar 7. Data Setelah Dibersihkan dan Bebas dari *Outlier*

Proses transformasi data agar siap digunakan dalam analisis yang berkaitan dengan algoritma Apriori dimulai setelah proses pembersihan dan penghapusan outlier selesai (Gambar 7). Setelah data transaksi telah dibersihkan, data encoding satu panas digunakan. Untuk menunjukkan keberadaannya dalam suatu transaksi, setiap produk diwakili dengan kolom biner berjumlah 0 dan 1. Proses ini memungkinkan sistem untuk mengenali kombinasi item yang sering muncul bersamaan dalam pembelian pelanggan dengan lebih akurat dan efisien. Kolom-kolom yang tidak relevan terhadap analisis asosiasi, seperti Transaction ID dan Product Revenue, dihapus untuk

menyederhanakan struktur data dan memfokuskan analisis pada keterkaitan antar produk. Dengan data yang telah terstruktur dengan baik, tahap pemodelan dapat dilakukan secara optimal untuk menggali pola pembelian konsumen yang signifikan.

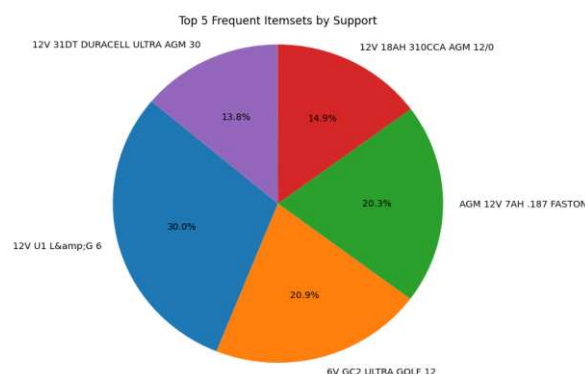
c. Visualisasi Data

Visualisasi data merupakan tahap penting dalam proses analisis karena memungkinkan peneliti dan pemangku kepentingan untuk memahami pola dan kecenderungan dalam data secara lebih intuitif. Dengan memanfaatkan grafik dan diagram, data yang kompleks dapat diterjemahkan menjadi informasi yang mudah dicerna serta membantu dalam proses pengambilan keputusan yang lebih akurat dan cepat. Berikut Gambar 8 adalah hasil dalam beberapa kategori:



Gambar 8. Visualisasi Berdasarkan Item yang Sering Dibeli

Salah satu bentuk visualisasi yang digunakan dalam penelitian ini adalah diagram batang horizontal yang menampilkan Top 10 Frequent Itemsets by Support, seperti yang ditunjukkan pada Gambar 8. Grafik ini memperlihatkan sepuluh produk dengan tingkat frekuensi (support) tertinggi dalam data transaksi penjualan. Produk "12V U1 L&G 6" menempati urutan teratas, diikuti oleh "6V GC2 ULTRA GOLF 12", "AGM 12V 7AH .187 FASTON", dan seterusnya. Gambar ini memperkuat hasil dari proses analisis sebelumnya, yang menunjukkan bahwa beberapa produk memang lebih sering muncul dalam transaksi pembelian dibanding produk lainnya. Hal ini bisa menjadi dasar untuk menyusun strategi penjualan, seperti rekomendasi produk utama, promosi khusus pada item populer, atau strategi penataan produk di toko.

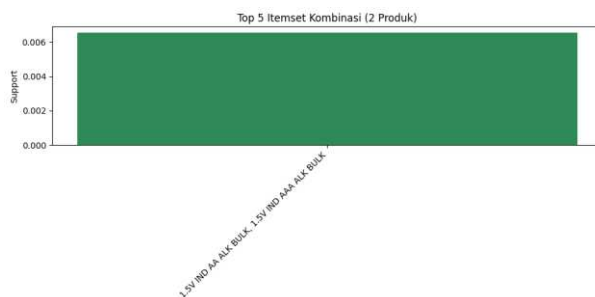


Gambar 9. Lima Item Yang Sering Dibeli

Visualisasi kedua pada Gambar 9 menyajikan Top 5 Frequent Itemsets dalam bentuk diagram lingkaran (pie chart). Visualisasi ini memberikan gambaran proporsional mengenai lima produk teratas yang paling sering muncul dalam transaksi pembelian konsumen. Masing-masing irisan mewakili persentase frekuensi kemunculan item tertentu berdasarkan nilai support yang telah dihitung sebelumnya menggunakan algoritma Apriori.

Dari gambar tersebut terlihat bahwa produk “12V U1 L&G 6” merupakan item dengan tingkat support tertinggi, yakni sebesar 30%, menunjukkan bahwa produk ini muncul pada hampir sepertiga dari seluruh transaksi yang diamati. Di posisi kedua dan ketiga secara berurutan adalah “6V GC2 ULTRA GOLF 12” dengan 20,9%, dan “AGM 12V 7AH .187 FASTON” dengan 20,3%. Sementara dua item lainnya, yaitu “12V 18AH 310CCA AGM 12/0” dan “12V 31DT DURACELL ULTRA AGM 30”, masing-masing mencatatkan persentase 14,9% dan 13,8%.

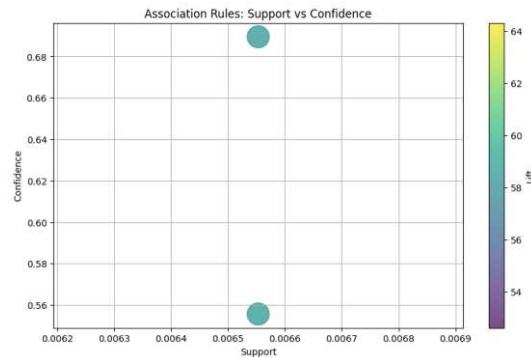
Penggunaan diagram pie ini mempermudah dalam memahami proporsi dominasi setiap produk dalam keseluruhan transaksi. Semakin besar irisan, semakin sering item tersebut muncul dalam keranjang belanja pelanggan. Dengan informasi ini, pelaku bisnis dapat mengarahkan strategi promosi atau pengelompokan produk (product bundling) berdasarkan kecenderungan pembelian nyata di lapangan.



Gambar 10. Kombinasi Produk Berdasarkan Itemset

Gambar 11 memperlihatkan visualisasi kombinasi dua produk (2-itemset) yang paling sering muncul secara bersamaan dalam transaksi pelanggan. Grafik ini menunjukkan itemset kombinasi "1.5V IND AA ALK BULK" dan "1.5V IND AAA ALK BULK" sebagai kombinasi produk dengan support tertinggi, yaitu sekitar 0,0065 atau setara dengan 0,65% dari total keseluruhan transaksi. Meskipun nilai support ini tampak kecil secara persentase, dalam konteks dataset besar seperti yang digunakan dalam penelitian ini (lebih dari 90 ribu transaksi), nilai tersebut tetap signifikan karena menunjukkan pola konsisten dalam pembelian konsumen. Artinya, meskipun tidak terlalu sering terjadi, ketika produk AAA dibeli, produk AA juga sering kali ikut dibeli. Pola ini mengindikasikan adanya preferensi pembelian berpasangan terhadap dua jenis baterai berjenis alkaline bulk tersebut. Dengan kata lain, pelanggan cenderung membeli kedua produk ini secara bersamaan, mungkin karena keduanya digunakan dalam perangkat elektronik yang berbeda tetapi dibutuhkan secara bersamaan.

Visualisasi ini penting dalam mendukung strategi cross-selling, yaitu dengan menempatkan dua produk tersebut secara berdampingan atau mengarahkan promosi bundling yang menasar keduanya. Dalam konteks Market Basket Analysis, pola seperti ini menjadi dasar dalam membentuk aturan asosiasi yang kuat dan relevan untuk implementasi strategi pemasaran yang lebih efektif.



Gambar 11. Peraturan Asosiasi: Support Vs Confidence

Gambar 11 menunjukkan hubungan antara dukungan dan keyakinan dari aturan asosiasi yang terbentuk; warna dan ukuran gelembung menunjukkan nilai lift.. Dalam grafik ini terdapat dua titik utama (bubble), yang masing-masing mewakili satu aturan asosiasi. Sumbu horizontal menunjukkan nilai support (dukungan) dari aturan tersebut, sedangkan sumbu vertikal menunjukkan nilai confidence (tingkat keyakinan). Warna gelembung dan ukurannya menunjukkan besarnya nilai lift, yang mengindikasikan kekuatan asosiasi antara produk dalam aturan tersebut.

- Aturan pertama memiliki nilai support sekitar 0.00655 dan confidence sebesar 68,9%.
- Aturan kedua juga memiliki nilai support yang sama (0.00655) namun dengan confidence yang lebih rendah, yaitu sekitar 55,5%.
- Kedua aturan ini memiliki nilai lift yang tinggi, berada di atas angka 58, menunjukkan bahwa korelasi antara produk dalam aturan tersebut jauh lebih kuat dibandingkan jika pembelian terjadi secara acak.

Interpretasi dari visualisasi ini menunjukkan bahwa:

- Meskipun tingkat kemunculan aturan ini dalam data (support) relatif kecil, aturan tersebut memiliki tingkat kepercayaan yang cukup tinggi, terutama pada aturan pertama.
- Nilai lift yang tinggi memperkuat bukti bahwa ada hubungan yang sangat signifikan antara produk-produk ini, dan layak dijadikan dasar untuk strategi seperti promosi produk bundling, rekomendasi otomatis, atau penempatan produk yang dekat.

Dengan demikian, visualisasi ini membantu peneliti dan analis memahami tidak hanya seberapa sering suatu aturan muncul, tetapi juga seberapa kuat dan relevan aturan tersebut dalam konteks perilaku pembelian pelanggan.

d. Pembuatan Itemset dan Aturan Asosiasi

Tujuan penelitian ini adalah untuk menemukan pola pembelian konsumen dari data transaksi ritel yang telah dibersihkan dan disiapkan sebelumnya. Untuk mencapai tujuan ini, algoritma Apriori digunakan pada tahap pemodelan data. Tahapan ini merupakan bagian penting dari proses analisis saku pasar (MBA) karena memungkinkan untuk menemukan informasi yang tersembunyi dalam kumpulan data yang sangat besar..

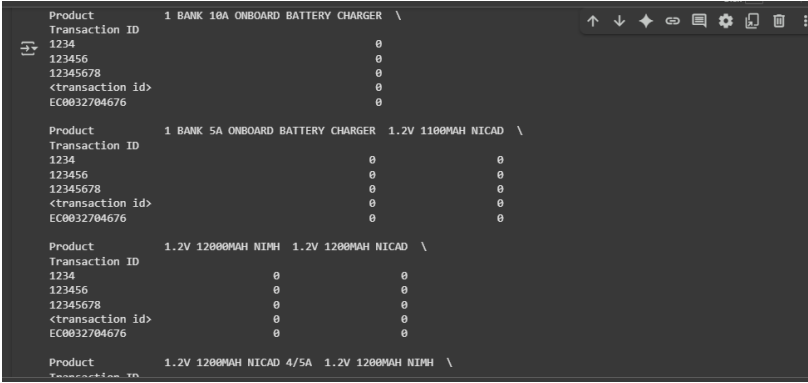
Langkah pertama dalam proses ini adalah menemukan kombinasi item atau produk yang sering muncul bersama dalam suatu transaksi, yang disebut sebagai frequent itemsets. Setelah menemukan frequent itemsets berdasarkan parameter minimum support yang telah ditentukan, proses dilanjutkan dengan membuat peraturan asosiasi berdasarkan nilai kepercayaan dan rasio lift. Aturan asosiasi ini menggambarkan hubungan antar produk, misalnya jika produk A dibeli maka besar kemungkinan produk B juga akan dibeli dalam transaksi yang sama. Proses ini penting untuk merumuskan strategi promosi silang (cross-selling), bundling, dan penataan produk di toko. Sebagai gambaran awal, ditampilkan Tabel 1. Data Sampel Transaksi, yang menyajikan lima produk berbeda yang ditemukan dalam dataset:

Tabel 1. Data Sampel Transaksi

NO ITEM	PRODUCT
1	3.7V 3400mah LIION 12.6WH
2	3V PHOTO LITHIUM
3	12V 11.2AH 225CCA AGM 12/0
4	12V 12AH 165CCA FLOODED 6/0
5	12V 12AH 210CCA AGM 12/0

Tabel 1 berisi daftar produk yang menjadi bagian dari transaksi ritel. Data ini selanjutnya dikonversi ke dalam format biner (binary format) melalui proses one-hot encoding. Setiap baris transaksi diubah menjadi representasi biner, di mana setiap kolom produk bernilai 1 jika produk tersebut ada dalam transaksi dan 0 jika tidak. Transformasi ini memudahkan algoritma Apriori dalam menghitung frekuensi kemunculan kombinasi produk serta membangun aturan asosiasi secara otomatis dan efisien.

Dengan model ini, toko dapat mengidentifikasi kombinasi barang yang sering dibeli bersamaan dan menggunakan data ini untuk membuat keputusan yang lebih strategis berbasis data, seperti membuat rekomendasi otomatis tentang produk atau mengatur rak berdasarkan pola pembelian konsumen.



Product	1 BANK 10A ONBOARD BATTERY CHARGER	\
Transaction ID	1234	0
	123456	0
	12345678	0
	<transaction id>	0
	EC0032704676	0
Product	1 BANK 5A ONBOARD BATTERY CHARGER 1.2V 1180MAH NICAD	\
Transaction ID	1234	0
	123456	0
	12345678	0
	<transaction id>	0
	EC0032704676	0
Product	1.2V 12000MAH NIMH 1.2V 1200MAH NICAD	\
Transaction ID	1234	0
	123456	0
	12345678	0
	<transaction id>	0
	EC0032704676	0
Product	1.2V 1200MAH NICAD 4/5A 1.2V 1200MAH NIMH	\
Transaction ID	1234	0

Gambar 12. Data telah di konversi dalam format biner (sumber google colab)

Gambar 12 menunjukkan hasil dari metode one-hot encoding untuk mengubah data transaksi ke dalam format biner. Setiap baris menunjukkan sebuah transaksi, dan setiap kolom menunjukkan satu produk tertentu. Nilai 1 menunjukkan bahwa produk tersebut ada dalam transaksi tertentu, sedangkan nilai 0 menunjukkan sebaliknya. Representasi ini sangat penting untuk proses analisis pasar kotak (MBA) karena memungkinkan algoritma Apriori untuk menemukan pola itemset secara efisien berdasarkan kemunculan produk-produk secara bersamaan dalam suatu transaksi.

Setelah data berada dalam format biner, tahapan berikutnya adalah menjalankan algoritma Apriori untuk menggali aturan asosiasi. Dalam penelitian ini, digunakan parameter Minimum Support sebesar 20% atau 0.2, yang berarti hanya kombinasi produk yang muncul setidaknya pada 20% dari total transaksi yang akan dipertimbangkan sebagai pola yang signifikan. Selain itu, ditetapkan pula Minimum Confidence sebesar 50% atau 0.5, yang digunakan untuk mengukur tingkat kepercayaan dari hubungan antar produk—misalnya, seberapa besar kemungkinan pelanggan yang membeli produk A juga akan membeli produk B. Pengaturan parameter ini dirancang untuk menyeimbangkan antara jumlah aturan yang ditemukan dengan kualitas dari aturan tersebut, sehingga hanya pola yang paling relevan dan bermakna yang digunakan untuk interpretasi lebih lanjut dan pengambilan keputusan strategis.

Tabel 2. Pembentukan 1 itemset

Product	Support
6V GC2 ULTRA GOLF 12	0.020286
AGM 12V 7AH .187 FASTON	0.019780
12V 18AH 310CCA AGM 12/0	0.014504
12V 31DT DURACELL ULTRA A	0.013468
DC 12V 35AH J TERM	0.012106

Tabel 2 menampilkan hasil perhitungan itemset tunggal yang diperoleh dari penerapan algoritma Apriori pada data transaksi yang telah dibersihkan dan ditransformasikan sebelumnya. Kolom Product menunjukkan nama produk yang teridentifikasi, sementara kolom Support menampilkan nilai dukungan (support) dari masing-masing produk. Nilai support mengindikasikan proporsi kemunculan produk tersebut dalam keseluruhan transaksi. Misalnya, produk “6V GC2 ULTRA GOLF 12” memiliki nilai support sebesar 0.020286, yang berarti produk tersebut muncul dalam sekitar 2,02% dari total transaksi. Produk lain seperti “AGM 12V 7AH .187 FASTON” dan “12V 18AH 310CCA AGM 12/0” juga menunjukkan frekuensi kemunculan yang relatif tinggi, dengan support mendekati 2% dan 1,45%.

Hasil ini memberikan gambaran awal tentang produk-produk yang paling sering muncul dalam transaksi pelanggan. Tahapan ini sangat penting karena hanya itemset dengan nilai support di atas ambang batas minimum (dalam hal ini 20%) yang akan digunakan dalam pembentukan aturan asosiasi (association rules). Selanjutnya, algoritma Apriori akan memproses kombinasi dari itemset ini untuk menghasilkan pasangan atau grup item yang sering dibeli bersamaan (frequent itemsets).

Dengan demikian, tahapan selanjutnya adalah pembentukan pola association rules berdasarkan frequent itemset yang telah ditemukan. Aturan asosiasi ini akan mengidentifikasi relasi antar produk dalam transaksi, misalnya: jika seorang pelanggan membeli produk A, maka kemungkinan besar ia juga akan membeli produk B. Proses ini memperhitungkan nilai confidence dan lift sebagai metrik evaluasi kekuatan dan reliabilitas dari setiap aturan yang terbentuk, sehingga hasil akhirnya dapat dimanfaatkan untuk strategi pemasaran seperti rekomendasi produk, penataan display, atau bundling promosi.

Tabel 3. Pembentukan aturan asosiasi

Antecedent	Consequents	Support	Confidence	lift
1.5V IND AAA	1.5V IND			
ALK	AA ALK	0.006553	68.95%	58.464
BULK	BULK			
1.5V IND AA	1.5V IND			
ALK BULK	AAA ALK	0.006553	55.57%	58.464
	BULK			

Tabel 3 menampilkan hasil pembentukan aturan asosiasi (association rules) yang diperoleh dari kombinasi itemset dengan dukungan kuat. Setiap baris pada tabel merepresentasikan satu aturan asosiasi yang terdiri dari dua bagian, yaitu antecedent (produk yang dibeli terlebih dahulu) dan consequent (produk yang kemungkinan dibeli setelahnya). Metode ini bertujuan untuk

mengidentifikasi keterkaitan pembelian antar produk berdasarkan parameter support, confidence, dan lift.

Salah satu aturan yang ditemukan menunjukkan bahwa apabila seorang konsumen membeli produk “1.5V IND AAA ALK BULK”, maka terdapat kemungkinan sebesar 68,95% (confidence) bahwa mereka juga akan membeli produk “1.5V IND AA ALK BULK”. Nilai support sebesar 0.006553 menunjukkan bahwa kombinasi pembelian kedua produk tersebut muncul pada sekitar 0,66% dari seluruh transaksi dalam dataset. Meskipun nilainya relatif kecil secara proporsi, nilai lift yang sangat tinggi, yaitu 58,464, menandakan adanya korelasi yang sangat kuat antara kedua produk. Lift yang jauh di atas angka 1 menunjukkan bahwa pembelian kedua produk secara bersamaan tidak terjadi secara kebetulan, tetapi memiliki hubungan kuat dan konsisten.

Aturan sebaliknya juga diperoleh, yaitu ketika “1.5V IND AA ALK BULK” dibeli terlebih dahulu, maka “1.5V IND AAA ALK BULK” memiliki peluang 55,57% untuk ikut dibeli. Walaupun nilai confidence sedikit lebih rendah, nilai lift tetap tinggi, mengindikasikan hubungan dua arah yang saling menguatkan antar produk. Hasil ini memiliki implikasi penting dalam strategi pemasaran dan pengelolaan stok. Kombinasi produk dengan asosiasi kuat seperti ini dapat digunakan untuk merancang strategi bundling, penempatan produk berdampingan di etalase, atau promosi silang (cross-selling). Dengan memanfaatkan pola ini, toko dapat meningkatkan nilai rata-rata pembelian pelanggan serta meningkatkan efisiensi dalam menyusun produk yang berpotensi laku tinggi secara bersamaan.

Hasil analisis data transaksi yang dilakukan dengan algoritma Apriori membawa banyak temuan penting yang dapat digunakan untuk membuat strategi pemasaran yang berhasil. Salah satu temuan utama adalah hubungan yang kuat antara produk 1.5V IND AAA ALK BULK dan 1.5V IND AA ALK BULK, dengan nilai keyakinan 68,9% dan lift 58,46. Ini menunjukkan bahwa konsumen yang membeli produk pertama lebih mungkin membeli produk kedua. Oleh karena itu, sangat penting untuk menerapkan strategi cross-selling, seperti menawarkan rekomendasi otomatis melalui platform e-commerce atau menawarkan paket produk yang ditampilkan setelah konsumen melakukan pembelian.

Selain itu, data kombinasi item (itemset) yang sering muncul bersama mengindikasikan peluang besar dalam pembuatan paket bundling. Untuk mendorong lebih banyak pembelian, kombinasi dua produk yang sering dibeli bersamaan dapat dimasukkan ke dalam satu paket promosi. Ini dapat dilakukan dengan memberikan diskon khusus atau menampilkan label "penawaran kombinasi terbaik" di halaman produk.. Produk dengan frekuensi pembelian tertinggi seperti 12V U1 L&G 6, 6V GC2 ULTRA GOLF 12, dan AGM 12V 7AH .187 FASTON juga sebaiknya ditempatkan secara strategis di bagian utama platform penjualan untuk menarik perhatian konsumen dan memicu pembelian lanjutan terhadap produk lain yang berkaitan.

Lebih jauh, strategi pemasaran juga dapat diarahkan pada personalisasi promosi berdasarkan pola pembelian pelanggan. Dengan segmentasi pelanggan berdasarkan produk yang sering dibeli, perusahaan dapat mengirimkan penawaran khusus yang relevan secara personal, seperti promosi baterai AAA kepada pelanggan yang sebelumnya membeli baterai tipe AA. Hal ini dapat meningkatkan efektivitas kampanye promosi dan memperkuat loyalitas pelanggan.

Selain pendekatan personal dan bundling, promosi musiman atau berbasis volume juga menjadi strategi yang potensial. Beberapa produk yang memiliki tingkat pembelian tinggi dapat ditawarkan dalam bentuk diskon volume atau promosi saat event tertentu, seperti awal tahun atau hari raya, untuk mendorong pembelian dalam jumlah besar. Secara keseluruhan, strategi pemasaran berbasis data ini tidak hanya meningkatkan efisiensi penjualan, tetapi juga mampu memaksimalkan pengalaman pelanggan melalui pendekatan yang lebih tepat sasaran dan berbasis bukti nyata dari pola pembelian aktual.

Hasil penelitian menunjukkan bahwa ada hubungan yang sangat kuat antara produk 1.5V IND AAA ALK BULK dan 1.5V IND AA ALK BULK; dengan nilai keyakinan sebesar 68,9% dan lift sebesar 58,46, konsumen sering membeli kedua produk ini bersamaan. Studi sebelumnya oleh Agrawal dan Srikant (1994) menemukan bahwa algoritma Apriori dapat mendeteksi pola asosiasi tersembunyi dalam data transaksi ritel dan dapat digunakan untuk strategi pemasaran seperti bundling dan cross-selling. Selain itu, penelitian oleh Prasetyo et al. (2021) menemukan bahwa menggunakannya pada data ritel dapat meningkatkan efektivitas kampanye promosi melalui rekomposisi.

Berdasarkan hasil analisis tersebut, strategi pemasaran yang dapat diterapkan antara lain adalah promosi bundling pada produk yang memiliki asosiasi tinggi, seperti 1.5V IND AAA ALK BULK dan AA ALK BULK, serta penerapan sistem rekomendasi otomatis pada platform e-commerce untuk mendorong pembelian silang (cross-selling). Produk-produk dengan tingkat pembelian tertinggi seperti 12V U1 L&G 6 dan 6V GC2 ULTRA GOLF 12 juga dapat diprioritaskan untuk ditampilkan dalam kampanye iklan digital atau penawaran khusus karena berpotensi menarik pembeli baru maupun mendorong pembelian ulang. Di samping itu, penyesuaian promosi musiman dan personalisasi penawaran berdasarkan histori pembelian konsumen juga menjadi pendekatan yang potensial untuk meningkatkan loyalitas pelanggan.

Namun, penelitian ini memiliki beberapa keterbatasan. Pertama, analisis hanya dilakukan pada data transaksi dalam satu periode waktu tanpa mempertimbangkan faktor musiman atau tren waktu. Kedua, tidak semua variabel kontekstual seperti demografi pelanggan atau jenis kanal distribusi dianalisis lebih lanjut. Keterbatasan ini dapat mempengaruhi generalisasi hasil, sehingga perlu dilakukan penelitian lanjutan yang melibatkan data dari berbagai periode waktu,

serta memperluas variabel yang dianalisis agar hasil yang diperoleh dapat lebih komprehensif dan aplikatif dalam konteks pemasaran yang lebih luas.

IV. KESIMPULAN

Hasil penelitian menunjukkan bahwa algoritma Apriori dapat digunakan secara efektif untuk mengidentifikasi pola pembelian konsumen berdasarkan data transaksi ritel. Aturan asosiasi yang dihasilkan mampu mengungkap hubungan antar produk yang sering dibeli bersamaan, seperti pada produk “1.5V IND AAA ALK BULK” dan “1.5V IND AA ALK BULK” yang memiliki nilai confidence sebesar 68,9% dan lift 58,46, menunjukkan hubungan asosiasi yang sangat kuat. Temuan ini memberikan dasar yang kuat bagi toko ritel untuk menerapkan strategi penataan produk, bundling, serta promosi silang guna meningkatkan nilai transaksi rata-rata. Namun, penelitian ini memiliki keterbatasan karena hanya menggunakan data transaksi tanpa mempertimbangkan faktor demografis atau perilaku pelanggan yang dapat mempengaruhi pola pembelian.

Oleh karena itu Untuk meningkatkan ketepatan hasil dan mendukung pengambilan keputusan yang lebih personal, disarankan agar metode ini dikombinasikan dengan data demografis dan perilaku pelanggan pada penelitian selanjutnya. Integrasi data tersebut diharapkan dapat memperkaya analisis dan menghasilkan rekomendasi produk yang lebih relevan. Selain itu, pengembangan sistem rekomendasi otomatis berbasis data transaksi real-time juga perlu dipertimbangkan agar hasil analisis dapat langsung diterapkan secara dinamis dalam operasional toko. Dengan demikian, toko ritel dapat merespons kebutuhan pasar secara cepat dan efisien melalui pendekatan analitik berbasis data.

DAFTAR PUSTAKA

- Andy Hermawan, Bayu Wicaksono, Tigfhar Ahmadjayadi, Bagas Surya Prakasa, & Jasico Dacomoro Aruan. (2024). Implementasi Algoritma Apriori pada Market Basket Analysis terhadap Data Penjualan Produk Supermarket. *Algoritma : Jurnal Matematika, Ilmu Pengetahuan Alam, Kebumian Dan Angkasa*, 2(5), 95–105. <https://doi.org/10.62383/algoritma.v2i5.137>
- Brighton, K., & Hariyanto, S. (2024). Penerapan Metode Market Basket Analisis Dengan Algoritma Apriori Pada Toko Ritel Elektronik. *Bit-Tech*, 7(1), 37–46. <https://doi.org/10.32877/bt.v7i1.1417>
- Fajri, I., Faqih, A., & Permana, S. E. (2025). PEMANFAATAN ALGORITMA APRIORI DALAM MENGUNGKAP POLA TRANSAKSI UNTUK DESAIN TATA LETAK

- PRODUK. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(2).
<https://doi.org/10.23960/jitet.v13i2.6181>
- Fathurrahman, M., Pratama, A. R., & Al-Mudzakir, T. (2023). Perbandingan Algoritma Apriori Dan Fp Growth Terhadap Market Basket Analysis Pada Data Penjualan Bakery. *Kesatria : Jurnal Penerapan Sistem Informasi (Komputer Dan Manajemen)*, 4(2).
- Fitriana Dewi, H., Hanny Hikmayanti Handayani, & Jamaludin Indra. (2022). IMPLEMENTASI ALGORITMA APRIORI TERHADAP MARKET BASKET ANALYSIS PADA DATA PENJUALAN RETAIL. *Jurnal Informatika Teknologi Dan Sains*, 4(4), 432–436.
<https://doi.org/10.51401/jinteks.v4i4.2182>
- Ghofur, M., Agus Pranoto, Y., & Ariwibisono, F. . (2020). PENERAPAN ALGORITMA APRIORI UNTUK ANALISIS DATA TRANSAKSI PENJUALAN PADA TOKO BERBASIS WEB (STUDI KASUS KALIBARU MART MALANG). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 4(1). <https://doi.org/10.36040/jati.v4i1.2341>
- Nur Rahmi, A., & Yosaphat Ananda Mikola. (2021). IMPLEMENTASI ALGORITMA APRIORI UNTUK MENENTUKAN POLA PEMBELIAN PADA CUSTOMER (STUDIKASUS : TOKO BAKOEL SEMBAKO). *Information System Journal*, 4(1), 14–19. <https://doi.org/10.24076/infosjournal.2021v4i1.561>
- Oktaviani, N. (2024). IMPLEMENTASI ALGORITMA APRIORI UNTUK ANALISIS POLA PEMBELIAN KONSUMEN PADA TOKO SERBA. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 3706–3711. <https://doi.org/10.36040/jati.v8i3.9624>
- Purwati, N., & Karnila, S. (2023). Strategi Peningkatan Penjualan Produk Menggunakan Market Basket Analysis. *Jurnal Sistem Informasi Bisnis*, 13(2), 96–103.
<https://doi.org/10.21456/vol13iss2pp96-103>
- Rifania, V. S., Saniman, S., & Azlan, A. (2023). Penerapan Algoritma Apriori Dalam Mencari Pola Pembelian Konsumen. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, 2(2), 201. <https://doi.org/10.53513/jursi.v2i2.5750>
- Rusdianto, D., Sutiyono, & Zaelani, L. (2020). IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA APRIORI UNTUK MENGETAHUI POLA PEMINJAMAN BUKU DI PERPUSTAKAAN UNIVERSITAS BALE BANDUNG. *J-SIKA|Jurnal Sistem Informasi Karya Anak Bangsa*, 2(02).
- Saefudin, S., & DN, S. (2019). PENERAPAN DATA MINING DENGAN METODE ALGORITMA APRIORI UNTUK MENENTUKAN POLA PEMBELIAN IKAN. *JSiI (Jurnal Sistem Informasi)*, 6(2). <https://doi.org/10.30656/jsii.v6i2.1587>
- Safitry, D. L., Rosianti, N., Divyaning, E., Zidan, H., Arnecia, Z. J., Paryudi, I., Veritawati, I., & Nursari, S. R. C. (2024). ANALISIS POLA PEMBELIAN KONSUMEN

MENGGUNAKAN ALGORITMA APRIORI UNTUK MENENTUKAN STRATEGI PEMASARAN PRODUK DI TOKO RETAIL X. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(1), 505–511. <https://doi.org/10.36040/jati.v9i1.12429>

Sari, N. D., & Khoiriah, S. (2022). Penerapan Metode Asosiasi Pada Toko Afifa Dengan Algoritma Apriori. *Instink: Inovasi Pendidikan, Teknologi Informasi Dan Komputer*, 1(1), 8–17. <https://doi.org/10.30599/instink.v1i1.1498>

Seniwati, E., Fahrizal Wahyu Syaiful Anam, & Ninik Tri Hartanti. (2024). Implementasi Algoritma Apriori untuk Menentukan Pola Pembelian Konsumen (Studi Kasus : Toko SRC Nanik di Wonogiri – Jawa Tengah). *J-Innovation*, 13(2), 44–51. <https://doi.org/10.55600/jipa.v13i2.291>

Sibarani, A. J. P. (2020). Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan Pola Penjualan Obat. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 7(2). <https://doi.org/10.35957/jatisi.v7i2.195>

Ulfha, N. F., & Amin, R. (2020). IMPLEMENTASI DATA MINING UNTUK MENGETAHUI POLA PEMBELIAN OBAT MENGGUNAKAN ALGORITMA APRIORI. *Komputasi: Jurnal Ilmiah Ilmu Komputer Dan Matematika*, 17(2). <https://doi.org/10.33751/komputasi.v17i2.2156>