JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 11, No. 1, March 2026, Pp.120-134

APLIKASI DATA MINING MENGGUNAKAN METODE APRIORI UNTUK MENGIDENTIFIKASI POLA PEMBELIAN DI TOKO ANANDA BARU

Anindya Fidela*[1)], Chyntia Raras Ajeng Widiawati [2)], Ali Nur Ikhsan[3)]

1. Sistem Informasi, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto, Indonesia
2. Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto, Indonesia
3. Informatika, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto, Indonesia

Article Info

Kata Kunci: Data Mining; Algoritma Apriori; Asosiasi; Toko Ananda Baru; Google Colab; Pyton

Keywords: Data Mining; Apriori Algorithm; Association; New Ananda Store; Google Colab; Python

Article history:
Received 22 April 2025
Revised 5 May 2025
Accepted 7 June 2025
Available online 1 March 2026

DOI :
https://doi.org/10.29100/jipi.v11i1.7700

* Corresponding author.
Corresponding Author
E-mail address: fidelaanindya88@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk mengidentifikasi pola pembelian produk di Toko Ananda Baru dengan memanfaatkan algoritma Apriori, sebagai upaya dalam mendukung pengelolaan stok dan pengembangan strategi pemasaran. Toko Ananda Baru sering menghadapi tantangan dalam pengelolaan persediaan yang efektif, terutama karena kesulitan dalam meramalkan permintaan konsumen. Untuk itu, analisis ini menggunakan data transaksi yang mencakup 7. 851 baris dan 9 kolom informasi, yang meliputi kode produk, nama produk, ID transaksi, jumlah, biaya, harga, diskon, keuntungan, serta persentase keuntungan. Proses analisis dimulai dengan mengimpor dan membersihkan data, termasuk penghapusan nilai-nilai yang hilang (NaN) untuk menjamin akurasi. Setelah proses pembersihan selesai, data transaksi diubah menjadi format biner melalui teknik one-hot encoding, yang memungkinkan penerapan algoritma Apriori untuk menemukan pola asosiasi. Hasil analisis mengungkap beberapa temuan signifikan: produk "ADEM SARI CAN 320ML HERBAL" menunjukkan nilai support sebesar 0,013072 (1,31%), mengindikasikan bahwa produk ini sering dibeli secara terpisah. Sementara kombinasi produk "BAGUS FRESH DISHWASH REF 575ML" dengan produk lainnya menunjukkan nilai support sebesar 0,019608 (1,96%), menandakan adanya asosiasi yang kuat antar produk. Selanjutnya, nilai lift tertinggi yang ditemukan dalam analisis ini mencapai 76,5, menunjukkan hubungan yang sangat kuat antara kombinasi produk yang sering dibeli bersama. Temuan ini memberikan wawasan berharga bagi Toko Ananda Baru untuk merancang strategi pemasaran yang lebih efektif, seperti promosi bundling untuk produk yang sering dibeli bersamaan serta pengaturan penempatan produk yang lebih strategis di dalam toko. Dengan menerapkan hasil temuan ini, diharapkan pengelolaan persediaan dapat ditingkatkan, mengurangi risiko kelebihan atau kekurangan stok, serta meningkatkan kepuasan pelanggan. Hal ini pada gilirannya akan memperkuat daya saing Toko Ananda Baru di pasar ritel.

ABSTRACT

This study aims to identify product purchase patterns at Toko Ananda Baru by utilizing the Apriori algorithm, as an effort to support inventory management and the development of marketing strategies. Toko Ananda Baru frequently faces challenges in effective inventory management, particularly due to difficulties in forecasting consumer demand. Therefore, this analysis uses transaction data consisting of 7,851 rows and 9 columns of information, including product codes, product names, transaction IDs, quantities, costs, prices, discounts, profits, and profit percentages. The analysis process begins with importing and cleaning the data, including removing missing values (NaN) to ensure accuracy. After the cleaning process is complete, the transaction data is transformed into a binary format using one-hot encoding, allowing the application of the Apriori algorithm to find association patterns. The analysis results reveal several significant findings: the product "ADEM SARI CAN 320ML HERBAL" shows a support value of 0.013072 (1.31%), indicating that this product is frequently purchased separately. Meanwhile, the combination of "BAGUS FRESH DISHWASH REF 575ML" with other products shows

a support value of 0.019608 (1.96%), suggesting a strong association between the products. Furthermore, the highest lift value found in this analysis reaches 76.5, indicating a very strong relationship between combinations of products that are frequently purchased together. These findings provide valuable insights for Toko Ananda Baru to design more effective marketing strategies, such as bundling promotions for products that are frequently purchased together and more strategic product placement in the store. By implementing these findings, it is expected that inventory management can be improved, reducing the risk of overstocking or stockouts, and increasing customer satisfaction. This, in turn, will strengthen Toko Ananda Baru's competitiveness in the retail market.

.

I. PENDAHULUAN

DALAM sektor bisnis ritel, memahami perilaku konsumen adalah hal yang penting untuk mempertahankan daya saing dan meningkatkan kinerja operasional. Toko Ananda Baru, yang menawarkan berbagai barang kebutuhan sehari-hari. Meskipun Toko Ananda Baru telah menerapkan sistem kasir untuk mencatat setiap transaksi, tantangan dalam pengelolaan persediaan barang secara efektif masih tetap ada. Saat ini, data transaksi yang tersedia belum dimanfaatkan secara maksimal untuk meramalkan kebutuhan konsumen. Hal ini sering kali berujung pada kelebihan atau kekurangan stok, yang pada gilirannya dapat memengaruhi kepuasan pelanggan dan efisiensi dalam pengelolaan inventaris. Oleh karena itu, sangat penting untuk mengadopsi metode analisis data yang dapat mengidentifikasi pola pembelian konsumen, guna mendukung pengambilan keputusan yang lebih akurat.

Salah satu cara strategis yang bisa diterapkan adalah dengan memastikan tersedia produk yang paling diminati oleh konsumen. Jika barang yang diinginkan konsumen tidak tersedia, mereka akan cenderung mencari alternatif di tempat lain, yang bisa mengakibatkan penurunan pendapatan dari usaha tersebut. Strategi lain mencakup pengorganisasian tata letak barang di dalam toko dengan cara yang paling efisien. Penempatan produk yang baik dapat mendorong konsumen untuk membeli barang tambahan yang letaknya berdekatan. Selain itu, memberikan saran langsung kepada pembeli, berdasarkan riwayat belanja atau kebiasaan membeli sebelumnya, juga bisa meningkatkan keberhasilan penjualan. Ini menunjukkan betapa pentingnya data dan informasi dalam menentukan keberhasilan strategi tersebut.

Sebagai contoh pengaturan susunan produk yang paling efisien memerlukan data transaksi asli (data mentah), informasi mengenai kemungkinan produk yang sering dibeli secara bersama, serta data tambahan lain. Dengan demikian, diperlukan sebuah metode yang dapat mengubah data menjadi informasi yang relevan untuk mendukung pengambilan keputusan. Metode atau proses untuk mengidentifikasi pola atau informasi dari sekumpulan data disebut dengan data *mining*[1][2].

Data *Mining* adalah teknik yang dapat mengekstrak informasi penting dari data yang komplek[3]. Salah satu metode pengembangan data yang paling populer untuk menganalisis hubungan antara produk yang dijual pada waktu yang sama adalah metode Apriori. Algoritma Apriori adalah salah satu metode yang digunakan untuk mencari frekuensi itemset dengan menerapkan teknik aturan asosiasi[4]. Selain itu metode ini mampu mengidentifikasi pola-pola asosiasi antar item dalam suatu transaksi[5]. Algoritma ini juga memberikan informasi mengenai keterkaitan antara item-item data dalam basis data[6][7]. Algoritma Apriori termasuk dalam kategori aturan asosiasi dalam bidang data *mining*[8]. Algoritma Apriori memanfaatkan pengetahuan tentang frekuensi atribut yang telah diketahui sebelumnya untuk mengolah informasi lebih lanjut. Dalam penerapannya, algoritma ini menentukan kandidat yang berpotensi muncul dengan memperhatikan nilai *minimum support dan minimum confidence*[4].

Jika dibandingkan dengan algoritma lain seperti FP-Growth dan ECLAT, algoritma Apriori lebih tepat digunakan untuk data transaksi berskala kecil[9]. Keunggulan utama Apriori terletak pada prosesnya yang sederhana dan mudah diimplementasikan[10], tanpa memerlukan struktur data tambahan yang rumit seperti FP-tree atau TIDSet seperti algoritma FP-Growth. Hal ini menjadikan Apriori pilihan yang lebih praktis dan ekonomis, terutama bagi usaha mikro, kecil, dan menengah (UMKM) seperti Toko Ananda Baru, yang biasanya memiliki keterbatasan dalam manajemen data dan sumber daya.

Namun, mayoritas penelitian yang ada saat ini cenderung berfokus pada penerapan algoritma Apriori di sektor ritel besar atau supermarket, yang memiliki volume transaksi jauh lebih besar dan kompleks. Penelitian-penelitian tersebut sering kali mengasumsikan adanya sumber daya dan infrastruktur yang lebih memadai, seperti sistem

manajemen inventaris yang canggih dan basis data transaksi yang lebih luas. Sementara itu, usaha ritel skala kecil dan menengah (UMKM) seperti Toko Ananda Baru menghadapi tantangan yang berbeda, seperti keterbatasan sumber daya, sistem yang lebih sederhana, serta volume transaksi yang lebih kecil namun beragam. Kekosongan dalam literatur ini menunjukkan perlunya penelitian lebih lanjut mengenai penerapan algoritma Apriori untuk menganalisis pola pembelian di ritel skala kecil dan menengah. Mengingat karakteristik UMKM yang lebih sederhana, tetapi tetap memerlukan pemahaman yang mendalam tentang perilaku konsumen dan pengelolaan inventaris, algoritma Apriori berpotensi menjadi alat yang efektif, asalkan disesuaikan dengan konteks dan kapasitas yang dimiliki oleh UMKM.

Dalam dunia bisnis, mengenali perilaku pelanggan adalah faktor penting untuk pengambilan keputusan yang strategis. Misalnya dengan memahami barang-barang yang biasanya diberi secara bersamaan[11], sebuah toko bisa merancang promosi yang lebih menarik dan terarah. Studi menunjukkan bahwa hubungan antar produk menawarkan informasi berharga dalam memperbaiki efisiensi strategi pemasaran.

Oleh karena itu dengan menerapkan metode Apriori, Toko Baru Ananda dapat menentukan ketertarikan antar produk yang sering dibeli bersamaan dalam satu transaksi. Misalnya, dengan mengidentifikasi produk-produk yang sering dibeli secara bersamaan, toko dapat merancang strategi penempatan produk yang lebih efektif dan menciptakan promosi yang relevan dengan pelanggan. Pengetahuan ini juga dapat membantu toko mengembangkan strategi pemasaran yang lebih efektif dan meningkatkan efisiensi manajemen inventaris. Selain bermanfaat dalam menyusun strategi pemasaran, metode Apriori juga dapat membantu Toko Baru Ananda dalam memprediksi kebutuhan produk. Dengan mengetahui produk apa saja yang akan dijual secara bersamaan, toko dapat memprediksi permintaan barang dan memastikan ketersedian sehingga mengurangi resiko kehabisan stok. Hal ini meningkatan kepuasan pelanggan dan menjaga loyalitas pelanggan.

Penerapan metode Apriori pada UMKM *ritel* tradisional, seperti Toko Ananda Baru, dapat memberikan kontribusi yang signifikan dalam meningkatkan daya saing mereka terhadap *ritel* modern yang telah lebih dulu mengadopsi teknologi analitik canggih. Dengan memanfaatkan data transaksi, UMKM mampu memahami preferensi konsumen dengan lebih akurat. Hal ini memungkinkan mereka untuk meningkatkan efisiensi operasional, merancang strategi promosi yang lebih efektif, serta mengoptimalkan pengelolaan stok barang. Upaya tersebut tidak hanya mendukung kelangsungan usaha kecil di tengah persaingan pasar yang semakin ketat, tetapi juga berperan dalam mendorong pertumbuhan ekonomi lokal melalui penguatan sektor informal dan penciptaan lapangan kerja. Oleh karena itu, penelitian ini sangat relevan dalam mendukung pembangunan ekonomi lokal berbasis teknologi dan data.

Pada penelitian sebelumnya data *mining* menggunakan algoritma Apriori menunjukan bahwa penerapan data mining dalam sektor ritel dapat memberikan dampak yang signifikan dalam meningkatkan efisiensi operasional dan penjualan. Studi yang dilakukan oleh [6] beserta rekan-rekanya dengan judul penerapan algoritma apriori dalam analisis keranjang belanja retail di wilayah jawa barat penelitian ini menghasilkan menghasilkan 6 aturan asosiasi dengan nilai tertinggi *confidence* sebesar 0,64. Selain itu penelitian yang dilakukan oleh Mayland Trifena beserta rekan-rekanya dengan judul implementasi algoritma apriori untuk menentukan paket bundel dalam penjualan toko swalayan XYZ berhasil dengan hasil berdasarkan riset di supermarket XYZ, air mineral dengan dukungan 23,61% paling banyak dibeli oleh pelanggan. Analisis korelasi menemukan 10 aturan pemasaran yang terkait. Sebagian besar konsumen yang membeli makanan juga membeli air mineral, memberikan informasi berharga untuk membuat paket produk yang dapat meningkatkan penjualan. Algoritma apriori memberikan wawasan strategis bagi supermarket XYZ dalam meningkatkan keuntungan dan kepuasan pelanggan[12].

Melalui penerapan algoritma Apriori diharapkan Toko Ananda Baru, tidak hanya memberikan manfaat langsung bagi pengelola toko tetapi juga berkontribusi dalam mempromosikan pendekatan berbasis data dalam bisnis *ritel,* khususnya untuk mengatasi tantangan dalam memahami kebiasan konsumen dan pengelolaan stok.

II. METODE PENELITIAN

Penelitian ini berfokus pada penerapan algoritma Apriori dengan mengikuti langkah-langkah terstruktur. Untuk mempermudah perhitungan dan pencarian informasi, seluruh dataset akan diproses menggunakan *Google Collaboratory* dengan bahasa pemrograman *Python. Platform* ini memunginkan implementasi algoritma Apriori melalui berbagai pustaka *Python*. Berikut adalah rincian tahapan penelitian yang dilakukan.

A. Pengumpulan Data

Pengumpulan data adalah langkah krusial dalam penelitian yang bertujuan untuk mendapatkan informasi yang *valid* dan *reliabel*. Berdasarkan berbagai sumber, teknik pengumpulan data dapat dibedakan menjadi dua

pendekatan utama kuantitatif dan kualitatif[13]. Pada penelitian ini metode yang digunakan adalah metode kualitatif dan metode yang digunakan adalah wawancara dan dokumentasi.

Wawancara merupakan interaksi langsung antara peneliti dan responden untuk menggali pandangan dan pengalaman mereka[1]. Sedangkan dokumentasi merupakan pengumpulan data dari dokumen,arsip, atau sumber tertulis lainnya yang relevan dengan penelitian, ini memberikan konteks historis dan mendukung analisis[1]. Pada penelitian ini peneliti berfokus kepada Toko Ananda Baru, dimana proses pengumpulan data dilakukan dengan mendatangi langsung ke toko. Data yang diperoleh dari sistem kasir yang berada di Toko Ananda Baru. Data yang berhasil dikumpulkan mencakup 7.851 transaksi dalam jangka waktu 1 bulan yaitu pada bulan Juni 2024.

B. Preprocessing

Proses *preprocessing* memegang peran penting karena kesuksesan proses data *mining* ditentukan oleh kualitas data yang diolah. Data yang tidak bersih atau tidak konsisten dapat menghasilkan analisis yang tidak tepat dan kurang akurat. Oleh karena itu, setiap tahap *preprocessing* perlu dilakukan dengan seksama agar dataset siap untuk dianalisis dan memberikan wawasan penting untuk pengambilan keputusan[14]. Proses ini termasuk menghapus data yang tidak lengkap,data duplikat, atau data yang tidak relevan. Data yang bersih dan konsisten penting untuk memastikan akurasi analisis. Ini mencakup pemeriksaan kesalahan cetak dan inkonsistensi dalam data. Dalam studi ini, total entri data yang dipakai sejumlah 7.851. data tersebut tidak memerlukan pembersihan lebih lanjut karena data telah memenuhi syarat dan format yang diperlukan untuk analisis. Ini menunjukkan bahwa data tersebut sudah dianggap relevan dan, sehingga tidak perlu menghapus *entri* yang tidak berkaitan, memperbaiki kesalahan data, aau menangani nilai yang hilang. Dengan kata lain, data tersebut telah siap untuk proses selanjutnya tanpa perlu ada perubahan atau perbaikan, dan bisa langsung digunakan pada *fase* berikutnya dalam penelitian.

C. Analisis Pola Frekuensi Tinggi

Frequent pattern mining adalah proses yang digunakan untuk mengidentifikasi himpunan itemset yang sering muncul bersama dalam transaksi atau dataset. Misalnya, dalam analisis keranjang belanja, teknik ini dapat digunakan untuk menemukan bahwa pelanggan yang membeli produk A juga sering membeli Produk B[15]. Dengan menemukan pola ini, perusahaan dapat meningkatkan strategi pemasaran dan penjualan mereka. Pada algoritma Apriori, analisis pola frekuensi tinggi dilakukan untuk mengidentifikasi semua *frequent itemset* dalam dataset, yaitu sekumpulan item atau atribut yang muncul dengan frekuensi tertentu dalam data[16].

D. Pengumpula*n Association Rule*

Association rule adalah teknik dalam data mining yang digunakan untuk menemukan aturan asosiasi antara kombinasi item dalam dataset secara bersamaan[17]. Teknik ini bertujuan untuk mengidentifikasi hubungan atau pola yang sering muncul secara bersamaan dalam transaksi,seperti yang sering dilakukan dalam analisis keranjang belanja di *supermarket*, yang dikenal sebagai *market basket analysis*. Pada tahap ini asosiasi *rule* dibuat dari frequent itemset. Aturan ini dinyatakan dalam format *IF-THEN,*dan *IF* mencakup premis atau item yang muncul lebih awal, sedangkan *THEN* adalah konsekuensi atau item yang muncul sebagai hasil dari premis itu.

E. Uji Lift Ratio

Lift Ratio mengukur seberapa kuat hubungan antara dua item dalam satu aturan asosiasi dibandingkan dengan kemungkinan munculnya item tersebut secara acak. *Left ratio* merupakan metrik untuk mengukur kekuatan hubungan antara nsekuensi (*item THEN)* dan *premis (item IF)* dalam aturan asosiasi, yang menunjukkan seberapa sering konsekuensi muncul bersama premis dibandingkan muncul secara terpisah[18]. Jika *left ratio* mendekati 1, ini menunjukkan bahwa konsekuensi (*THEN)* sangat bergantung pada *premis(IF),* yang berarti aturan asosiasi tersebut kuat. Sebaliknya, jika ratio mendekati 0, maka hubungan antara premis dan konsekuen sangat lemah atau tidak ada. Nilai l*ift* yang diperlukan untuk menentukan korelasi antara kombinasi item yang valid merupakan hasil dari perhitungan yang dilakukan[19].

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Setelah melakukan proses pengumpulan data secara langsung yang dilakukan dengan cara wawancara dengan karyawan toko dan observasi di lokasi, yang mencatat pola transaksi dan kebiasaan belanja konsumen sebagai dasar analisis selanjutnya.

TABEL I
DATA NAMA BARANG/ITEM

Nama Barang
4 BERSAUDARA EMPING MELINJO BESAR

CANTIKA SNACK JENGKOL LBR
LUCIDO-L HMS TBVS 200ML FOR NORMAL
1PUSH VAPE 7,6ML ORANGE
4 BERSAUDARA BASTIK
4 BERSAUDARA EMPING MELINJO
4 BERSAUDARA USUS AYAM CRISPY
A&G PARFUM
ABC APPLE 250ML
ABC ASAM JAWA
ABC GUAVA 250ML
ABC JUS SIRSAK 250ML

Dapat dilihat Tabel I menampilkan dataset produk yang dijual di toko kelontong, dengan sekitar 15.702 item yang ditawarkan kepada pelanggan.

TABEL II
DATASET ITEM TRANSAKSI

No	Kode	Nama Barang	*TolFak*	*TolQty*	*Cost*	*Price*	*Disk.Item*	*Profit*	*%*
1	1500087	RKK DJARUM 76 K 12B	515	2.427	34.570.784	35.273.850	0	703.066	2,03
2	1340011	KOPI KAPAL API SP MIX 25G	169	1.757	2.317.089	2.462.600	0	145.510	6,27
3	1090087	MASAKO SAPI 10G	118	1.739	648.590	735.775	0	87.184	13,44
4	1900136	SOKLIN LIQ 24ML BIRU	61	1.560	595.980	664.800	0	68.820	11,54
5	1140001	MIE SEDAAP GORENG	253	1.481	3.968.541	4.248.855	0	280.313	7,06
.....									
7851	7851 1013317	DELI GUNTING 6054	2	0	0	0	0	0	0,00

Berdasarkan hasil pengamatan serta pencatatannya, telah didapatkan kumpulan data transaksi yang yang dapat dilihat dalam Tabel II. Kumpulan data ini kemudian digunakan untuk menerapkan algoritma Apriori. Agar perhitungan dan pencarian informasi lebih mudah, seluruh kumpulan data akan dihitung menggunakan *Google Collaboratory* dengan menggunakan bahasa pemrograman *Python,* sebagai berikut:

```
# Import library yang dibutuhkan
import pandas as pd
from mlxtend.frequent_patterns import apriori, association_rules
import matplotlib.pyplot as plt
from google.colab import drive
```

Gambar. 1. Proses *Importing library*

Dapat dilihat pada Gambar 1, terdapat kode *Python* yang bertujuan untuk mengimport *library* yang dibutuhkan dalam analisis dengan algoritma Apriori. *Library 'pandas'* di import menggunakan nama panggilan '*pd*' agar dapat memanipulasi dan menganalisis data dalam format Data *Frame. Library 'mlxtend.frequent_patterns'* digunakan untuk mengimport fungsi 'apriori' yang berguna dalam menemukan itemset dengan frekuensi tinggi, serta *'association rules'* untuk membuat aturan asosiasi. Selain itu, *'matplotlib.pyplot*` diimport sebagai *'plt'* untuk menghasilkan visualisasi data, seperti grafik dan diagram. Terakhir *'google collab'* diimpor untuk memfasilitasi akses ke *Google Drive*, yang penting untuk menyimpan atau mengakses database selama tahap analisis.

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 11, No. 1, March 2026, Pp.120-134

```
# 1. Mount Google Drive
drive.mount('/content/drive')

data = '/content/drive/My Drive/APRIORI/JUNI.csv'

# 3. Memuat data penjualan
df = pd.read_csv(data)
```

Gambar. 2. *Import* data dan Pembacaan Dataset

Dapat dilihat pada Gambar 2, proses pengunggahan dataset dilakukan, di mana dataset dibaca dari file *Excel* yang disimpan dalam format *CSV*. Proses mengakses database dari *Google Drive,* dimana *drive* dipasang menggunakan *drive.mount().* Dataset berformat *CSV* yang tersimpan di jalur tertentu diakses dan dibaca menggunakan fungsi *pd.read_csv(),* lalu dimuat ke dalam Data Frame untuk analisis lebih lanjut.

```
# 4. Membersihkan nama kolom dengan menghapus spasi ekstra
df.columns = df.columns.str.strip()

# 5. Menampilkan nama kolom untuk memverifikasi
print(df.columns)
```

Gambar. 3. *Cleaning* Data dan Hasil Pembacaan Dataset

Sebelum menampilkan hasil dataset, langkah awal yang dilakukan dengan memastikan keberhasilan nama-nama kolom dalam *DataFrame.* Hal ini dilakukan dengan menghapus spasi ekstra di sekitar nama kolom menggunakan perintah *df.columns.str.strip().* Langkah ini penting untuk memastikan tidak ada kesalahan saat mengakses kolom akibat adanya spasi tambahan. Setelah itu, nama kolom yang telah diperbaiki diampunkan menggunakan perintah *print(df.clums)* untuk memverifikasi bahwa proses telah berhasil dilakukan.

```
Dataset yang telah dibaca:
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will not call
  and should_run_async(code)
```

	Kode	Nama Barang	Tot.Fak	Tot.Qty	Cost	Price	Disk.Item	Profit	%
0	1 1500087	RKK DJARUM 76 K 12B	515	2.427	34.570.784	35.273.850	0	703.066	2,03
1	2 1340011	KOPI KAPAL API SP MIX 25G	169	1.757	2.317.089	2.462.600	0	145.510	6,27
2	3 1090087	MASAKO SAPI 10G	118	1.739	648.590	735.775	0	87.184	13,44
3	4 1900136	SOKLIN LIQ 24ML BIRU	61	1.560	595.980	664.800	0	68.820	11,54
4	5 1140001	MIE SEDAAP GORENG	253	1.481	3.968.541	4.248.855	0	280.313	7,06

Gambar. 4. Hasil Pembacaan Dataset

Pada dilihat pada Gambar 4 menampilkan hasil keluaran yang dihasilkan dari *source code* yang telah dijalankan sebelumnya.

B. Preprocessing

```
# Menampilkan data sebelum penghapusan nilai NaN
print("Dataset sebelum penghapusan NaN:")
print(df.head())  # Menampilkan 5 baris pertama dari dataset asli
print("\nJumlah baris dan kolom sebelum penghapusan NaN:")
print(df.shape)
```

Gambar. 5. Validasi Item Dataset

Perintah ini berfungsi untuk menampilkan dataset sebelum proses penghapusan nilai *NaN*. Pertama-tama, tampilkan sebuah pesan yang mengkonfirmasi bahwa data yang ditampilkan adalah data asli. Selanjutnya, *`df.*

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 11, No. 1, March 2026, Pp.120-134

head()` digunakan untuk menampilkan lima baris pertama dari dataset, sementara `*df. shape*` memberikan informasi mengenai jumlah baris dan kolom yang terdapat dalam dataset sebelum data dibersihkan dari nilai *NaN.*

```
Dataset sebelum penghapusan NaN:
        Kode                      Item  TransactionID  Quantity       Cost  \
0  1  1500087       RKK DJARUM 76 K 12B            515     2.427  34.570.784
1  2  1340011  KOPI KAPAL API SP MIX 25G            169     1.757   2.317.089
2  3  1090087           MASAKO SAPI 10G            118     1.739     648.590
3  4  1900136      SOKLIN LIQ 24ML BIRU             61     1.560     595.980
4  5  1140001         MIE SEDAAP GORENG            253     1.481   3.968.541

        Price  Disk.Item   Profit      %
0  35.273.850          0  703.066   2,03
1   2.462.600          0  145.510   6,27
2     735.775          0   87.184  13,44
3     664.800          0   68.820  11,54
4   4.248.855          0  280.313   7,06
```

Gambar. 6. Dataset Transaksi

Dataset yang digunakan dalam penelitian ini pada awalnya terdiri dari 7.851 baris dan 9 kolom, yang mencakup informasi mengenai kode produk, nama produk, *ID* transaksi, jumlah, biaya, harga, diskon, keuntungan dan persentase keuntungan. Namun, terdapat beberapa nilai dalam satase tersebut yang mengandung *NaN (Not a Number)* atau data yang hilang. Untuk memastikan kualitas dan akurasi analisis, langkah pembersihan data dilakukan dengan menghapus baris-baris yang mengandung nilai *NaN.* Dengan demikian, dataset yang digunakan untuk analisis selanjutnya memiliki jumlah baris yang lebih sedikit, namun tetap lengkap tanpa data yang hilang. Penghapusan nilai *NaN* ini sangat penting untuk mencegah distorsi dalam hasil analisis serta model yang dibangun, sehingga hanya data yang *valid* dan terperinci yang digunakan dalam proses penelitian.

```
# Menghapus baris yang memiliki nilai NaN
df_cleaned = df.dropna()
```

Gambar. 7. *Filtering nan* dari dataset

```
Dataset setelah penghapusan NaN:
        Kode                      Item  TransactionID  Quantity       Cost  \
0  1  1500087       RKK DJARUM 76 K 12B            515     2.427  34.570.784
1  2  1340011  KOPI KAPAL API SP MIX 25G            169     1.757   2.317.089
2  3  1090087           MASAKO SAPI 10G            118     1.739     648.590
3  4  1900136      SOKLIN LIQ 24ML BIRU             61     1.560     595.980
4  5  1140001         MIE SEDAAP GORENG            253     1.481   3.968.541

        Price  Disk.Item   Profit      %
0  35.273.850          0  703.066   2,03
1   2.462.600          0  145.510   6,27
2     735.775          0   87.184  13,44
3     664.800          0   68.820  11,54
4   4.248.855          0  280.313   7,06

Jumlah baris dan kolom setelah penghapusan NaN:
(7851, 9)
```

Gambar. 8. Dataset Tanpa *NAN*

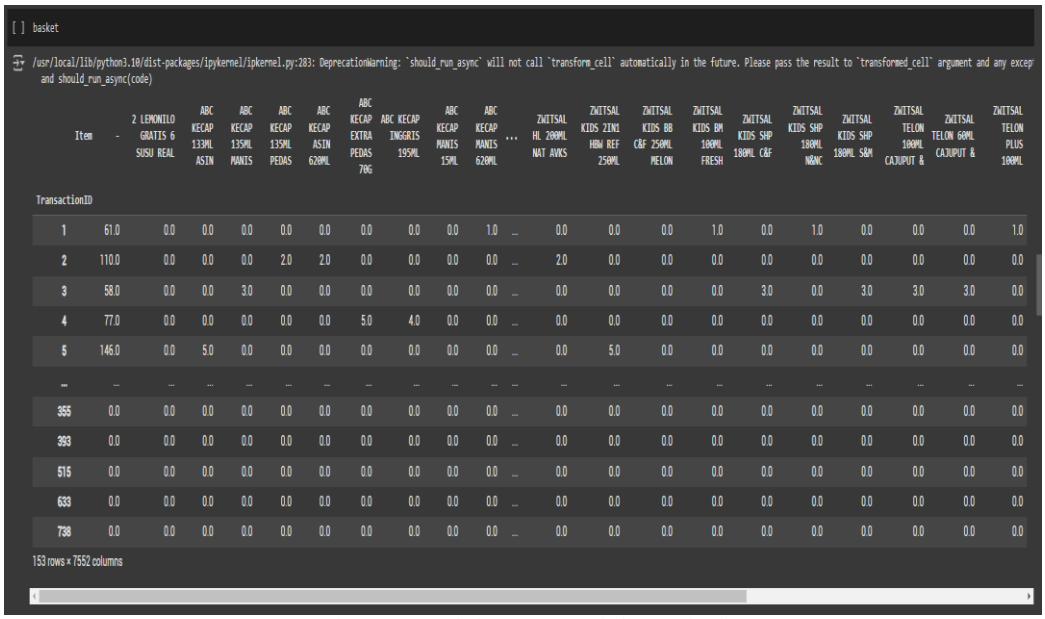
```
#Memeriksa data untuk memastikan kolom yang digunakan ada
df[['TransactionID', 'Item', 'Quantity']].head()

# Preprocessing Data untuk Algoritma Apriori
# membuat data dalam format transaksi (one-hot encoding)
# Mengelompokkan berdasarkan TransactionID dan Item, kemudian mengubah Quantity menjadi 1 jika produk ada dalam transaksi
basket = df.groupby(['TransactionID', 'Item'])['Quantity'].sum().unstack().reset_index().fillna(0).set_index('TransactionID')

basket = basket.fillna(0)
```

```
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will not call `transform_cell` automa
  and should_run_async(code)
```

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 11, No. 1, March 2026, Pp.120-134



```
[ ] basket
```

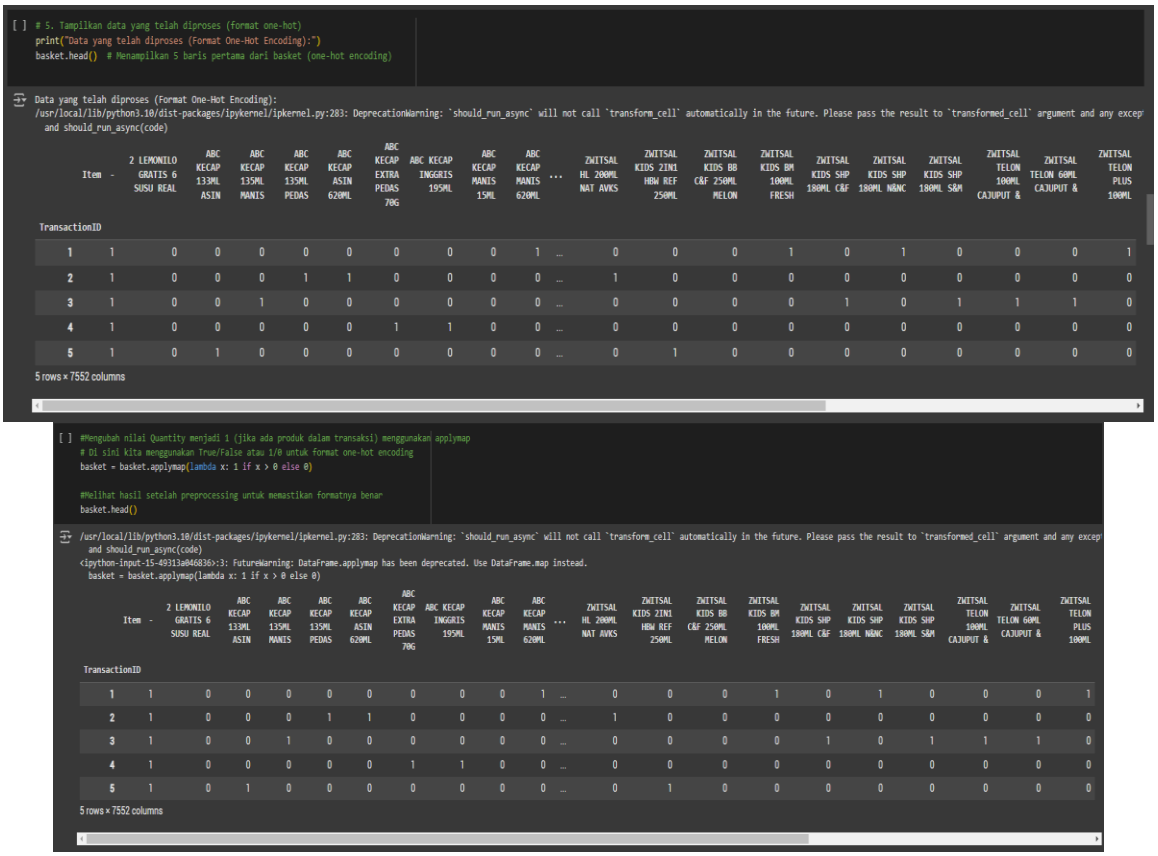
```
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will not call `transform_cell` automatically in the future. Please pass the result to `transformed_cell` argument and any excep
  and should_run_async(code)
```

Item	-	2 LEMONILO GRATIS 6 SUSU REAL	ABC KECAP 133ML ASIN	ABC KECAP 135ML MANIS	ABC KECAP 135ML PEDAS	ABC KECAP ASIN 620ML	ABC KECAP EXTRA PEDAS 70G	ABC KECAP INGGRIS 195ML	ABC KECAP MANIS 15ML	ABC KECAP MANIS 620ML	...	ZWITSAL HL 200ML NAT AVKS	ZWITSAL KIDS 2IN1 HBW REF 250ML	ZWITSAL KIDS BB C&F 250ML MELON	ZWITSAL KIDS BM 100ML FRESH	ZWITSAL KIDS SHP 180ML C&F	ZWITSAL KIDS SHP 180ML N&NC	ZWITSAL KIDS SHP 180ML S&M	ZWITSAL TELON 100ML CAJUPUT &	ZWITSAL TELON 60ML CAJUPUT &	ZWITSAL TELON PLUS 100ML
TransactionID																					
1	61.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	...	0.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	1.0
2	110.0	0.0	0.0	0.0	2.0	2.0	0.0	0.0	0.0	0.0	...	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	58.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	3.0	0.0	3.0	3.0	3.0	0.0
4	77.0	0.0	0.0	0.0	0.0	0.0	5.0	4.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	146.0	0.0	5.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	5.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
355	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
393	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
515	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
633	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
738	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

153 rows × 7552 columns

Gambar. 9. Mengubah Dataset Menjadi Bentuk Biner

Kode tersebut melakukan pemeriksaan data untuk memastikan kolom '*transactionID', 'item',* dan *'Quantity'* tersedia, lalu menampilkan lima baris pertama sebagai langkah verifikasi. Kemudian, dilakukan *preprocessing* untuk mengonversi data transaksi ke dalam format *one-hot-encoding* yang diperlukan oleh Algoritma Apriori. Data dikelompokkan berdasarkan *'TansactionID'* dan 'item', dengan menghitung jumlah kuantitas produk. Selanjutnya data ini diubah menjadi *pivot table*, dimana 'TransactionID' berfungsi sebagai baris 'item' sebagai kolom, dan kualitas sebagai nilai. Untuk mengatasi nilai kosong, digunakan metode *'fillna(0)'* yang menggantinya dengan 0. Selain itu, kualitas diubah menjadi 1 jika produk terdapat dalam transaksi tertentu. Hasil akhirnya adalah data biner, dimana 1 menunjukkan bahwa produk dibeli, sementara 0 menandakan sebaliknya. Format ini sudah siap digunakan untuk analisis asosiasi item.



```
[ ] # 5. Tampilkan data yang telah diproses (format one-hot)
    print("Data yang telah diproses (Format One-Hot Encoding):")
    basket.head()  # Menampilkan 5 baris pertama dari basket (one-hot encoding)
```

```
Data yang telah diproses (Format One-Hot Encoding):
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will not call `transform_cell` automatically in the future. Please pass the result to `transformed_cell` argument and any excep
  and should_run_async(code)
```

Item	-	2 LEMONILO GRATIS 6 SUSU REAL	ABC KECAP 133ML ASIN	ABC KECAP 135ML MANIS	ABC KECAP 135ML PEDAS	ABC KECAP ASIN 620ML	ABC KECAP EXTRA PEDAS 70G	ABC KECAP INGGRIS 195ML	ABC KECAP MANIS 15ML	ABC KECAP MANIS 620ML	...	ZWITSAL HL 200ML NAT AVKS	ZWITSAL KIDS 2IN1 HBW REF 250ML	ZWITSAL KIDS BB C&F 250ML MELON	ZWITSAL KIDS BM 100ML FRESH	ZWITSAL KIDS SHP 180ML C&F	ZWITSAL KIDS SHP 180ML N&NC	ZWITSAL KIDS SHP 180ML S&M	ZWITSAL TELON 100ML CAJUPUT &	ZWITSAL TELON 60ML CAJUPUT &	ZWITSAL TELON PLUS 100ML
TransactionID																					
1	1	0	0	0	0	0	0	0	0	1	...	0	0	0	1	0	1	0	0	0	1
2	1	0	0	0	1	1	0	0	0	0	...	1	0	0	0	0	0	0	0	0	0
3	1	0	0	1	0	0	0	0	0	0	...	0	0	0	0	1	0	1	1	1	0
4	1	0	0	0	0	0	1	1	0	0	...	0	0	0	0	0	0	0	0	0	0
5	1	0	1	0	0	0	0	0	0	0	...	0	1	0	0	0	0	0	0	0	0

5 rows × 7552 columns

```
[ ] ##Mengubah nilai Quantity menjadi 1 (jika ada produk dalam transaksi) menggunakan applymap
    # Di sini kita menggunakan True/False atau 1/0 untuk format one-hot encoding
    basket = basket.applymap(lambda x: 1 if x > 0 else 0)

    ##Melihat hasil setelah preprocessing untuk memastikan formatnya benar
    basket.head()
```

```
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will not call `transform_cell` automatically in the future. Please pass the result to `transformed_cell` argument and any excep
  and should_run_async(code)
<ipython-input-15-49313a046036>:3: FutureWarning: DataFrame.applymap has been deprecated. Use DataFrame.map instead.
  basket = basket.applymap(lambda x: 1 if x > 0 else 0)
```

Item	-	2 LEMONILO GRATIS 6 SUSU REAL	ABC KECAP 133ML ASIN	ABC KECAP 135ML MANIS	ABC KECAP 135ML PEDAS	ABC KECAP ASIN 620ML	ABC KECAP EXTRA PEDAS 70G	ABC KECAP INGGRIS 195ML	ABC KECAP MANIS 15ML	ABC KECAP MANIS 620ML	...	ZWITSAL HL 200ML NAT AVKS	ZWITSAL KIDS 2IN1 HBW REF 250ML	ZWITSAL KIDS BB C&F 250ML MELON	ZWITSAL KIDS BM 100ML FRESH	ZWITSAL KIDS SHP 180ML C&F	ZWITSAL KIDS SHP 180ML N&NC	ZWITSAL KIDS SHP 180ML S&M	ZWITSAL TELON 100ML CAJUPUT &	ZWITSAL TELON 60ML CAJUPUT &	ZWITSAL TELON PLUS 100ML
TransactionID																					
1	1	0	0	0	0	0	0	0	0	1	...	0	0	0	1	0	1	0	0	0	1
2	1	0	0	0	1	1	0	0	0	0	...	1	0	0	0	0	0	0	0	0	0
3	1	0	0	1	0	0	0	0	0	0	...	0	0	0	0	1	0	1	1	1	0
4	1	0	0	0	0	0	1	1	0	0	...	0	0	0	0	0	0	0	0	0	0
5	1	0	1	0	0	0	0	0	0	0	...	0	1	0	0	0	0	0	0	0	0

5 rows × 7552 columns

Gambar. 10. Dataset Transaksi Dalam Bentuk Biner

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 11, No. 1, March 2026, Pp.120-134

C. Analisis Pola Frekuensi Tinggi

```
# Perhitungan nilai support untuk satu item
item_support = frequent_itemsets[frequent_itemsets['itemsets'].apply(lambda x: len(x) == 1)]
print("Support untuk satu item:")
item_support[['itemsets', 'support']].head()  # Menampilkan support untuk satu item
```

Gambar. 11. Nila*i Support* Untuk Perhitungan Satu Item

Proses ini digunakan untuk menghitung nilai support dari setiap item tunggal dalam daftar item yang sering muncul (frequent itemset). Proses ini dilakukan dengan menyaring data agar hanya item yang berdiri sendiri yang dipilih. Dengan demikian, proses ni akan menampilkan nilai support setiap item. Hasil tersebut memberikan wawasan tentang seberapa sering setiap item muncul secara individu dalam transaksi, yang sangat berguna untuk analisis lebih lanjut, seperti mengidentifikasi produk yang paling sering dibeli.

```
Support untuk satu item:
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_ru
  and should_run_async(code)
```

	itemsets	support
0	(-)	0.235294
1	(ADEM SARI CAN 320ML HERBAL)	0.013072
2	(AZARINE MOIST SUNSERUM 100ML)	0.013072
3	(BAGUS FRESH DISHWASH REF 575ML)	0.019608
4	(BLUE BAND MARGARIN 200G)	0.013072

Gambar. 12. Nilai *Support* Satu Item.

Hasil perhitungan *Support* untuk satu item menunjukkan nilai dukungan untuk setiap item yang muncul dalam transaksi, yang mengukur seberapa sering item tersebut muncul dalam keseluruhan transaksi sebagai sebuah proporsi. Sebagai contoh, item "ADEM SARI CAN 320ML HERBAL" memiliki nilai *support* sebesar 0,013072, yang berarti item tersebut muncul dalam sekitar 1,3% dari total transaksi.

```
[ ] # Perhitungan support untuk dua item (dengan threshold tertentu)
    # Menggunakan data yang sudah difilter untuk itemset dengan dua item
    pair_support_filtered = frequent_itemsets[frequent_itemsets['itemsets'].apply(lambda x: len(x) == 2)]
    print("Perhitungan support untuk dua item:")
    pair_support_filtered[['itemsets', 'support']].head()  # Menampilkan support untuk dua item
```

Gambar. 13. Nilai *Support* Dua Item

Kode ini digunakan untuk menyaring itemset yang terdiri dari dua item dalam dataset *`frequent_itemsets`*. Proses ini dilakukan dengan memeriksa panjang setiap itemset melalui fungsi lambda. Setelah berhasil memilih itemset yang mengandung dua item, kode ini akan mencetak pesan yang mengonfirmasi bahwa perhitungan s*upport* untuk pasangan dua item tersebut sedang ditampilkan. Selanjutnya, kolom *`itemsets`* dan *`support`* dari hasil yang telah disaring ditampilkan dengan mengambil lima baris pertama menggunakan fungsi *`. head()`*, sehingga kita dapat melihat dengan singkat pasangan item beserta nilai support-nya.

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 11, No. 1, March 2026, Pp.120-134

```
Perhitungan support untuk dua item:
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will not call
  and should_run_async(code)
```

	itemsets	support
55	(ADEM SARI CAN 320ML HERBAL, -)	0.013072
56	(AZARINE MOIST SUNSERUM 100ML, -)	0.013072
57	(BAGUS FRESH DISHWASH REF 575ML, -)	0.019608
58	(BLUE BAND MARGARIN 200G, -)	0.013072
59	(-, CHARM COOL FRESH 23CM NON WING)	0.013072

Gambar. 14. Nilai *Support* dari Kombinasi Dua Item

```
Support untuk kombinasi dua item (dengan support tertinggi):
     support                          itemsets
57  0.019608    (BAGUS FRESH DISHWASH REF 575ML, -)
86  0.019608       (PADIMAS MOONCAKE 60G KACANG, -)
81  0.019608   (MY BABY KIDS 2IN1 HAIR&BC 100ML, -)
102 0.019608  (ZWITSAL BB HAIR & BODY REF 400ML, -)
70  0.019608     (KAPAS NEW MODERN PINK INOVASI, -)
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will not call `transform_cell` automatically
  and should_run_async(code)
```

Gambar. 15. Nilai *Support* dari Kombinasi Dua Item Tertinggi

Support untuk kombinasi dua item dengan nilai support tertinggi mengindikasikan frekuensi relatif kemunculan pasangan item dalam transaksi. Nilai *support* yang diperoleh adalah 0. 019608, yang berarti kombinasi ini muncul sekitar 1,96% dari total transaksi. Setiap baris pada data menampilkan pasangan item beserta nilai *support*-nya, contohnya kombinasi antara "BAGUS FRESH DISHWASH REF 575ML" dan "-" (item lain yang tidak disebutkan), serta kombinasi lainnya. Dalam penelitian ini, hasil dari *confidence* dan *lift ratio* akan disajikan bersamaan dengan data lainnya untuk membentuk model *association rule* secara menyeluruh. Tujuannya adalah untuk lebih memahami kekuatan dan relevansi hubungan antar item, serta memberikan wawasan mendalam guna mendukung pengambilan keputusan bisnis.

D. Uji Lift Ratio

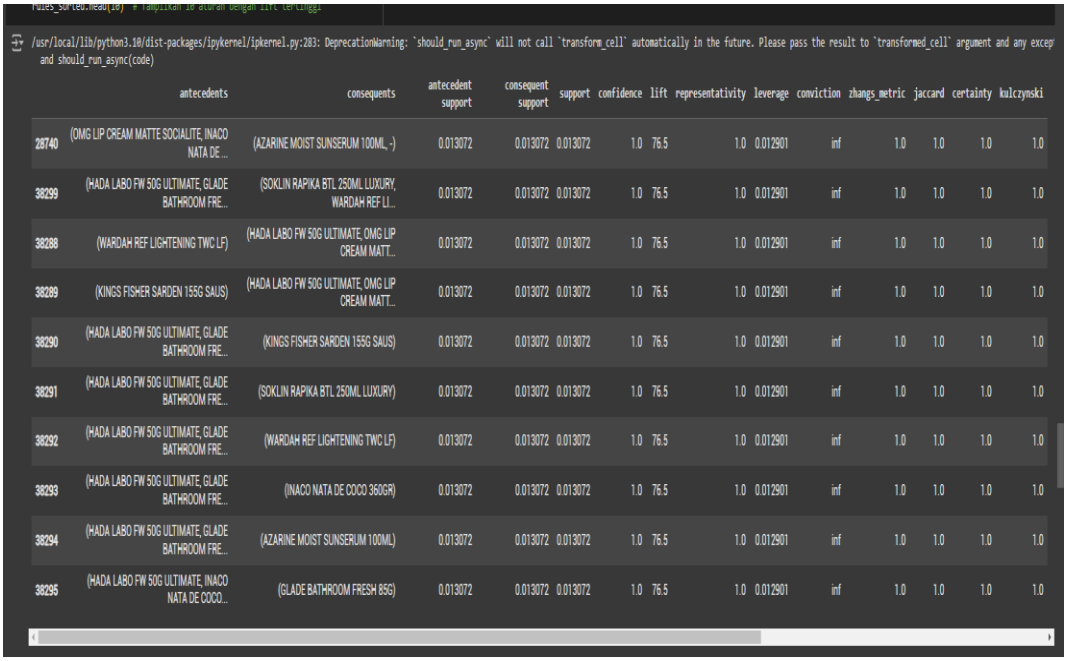


Gambar. 16. *Building* Model Dan *Collecting The Information*

Kode ini berfungsi untuk menyaring dan menampilkan aturan asosiasi berdasarkan nilai *lift* tertinggi, yang menunjukkan kekuatan hubungan antara item-item dalam aturan tersebut. Dengan mengurutkan aturan menurut kolom *lift* secara menurun, kode ini memastikan bahwa aturan yang menunjukkan hubungan terkuat antara item yang sering dibeli bersamaan akan muncul di urutan teratas. Setelah proses pengurutan selesai, kode ini kemudian menampilkan sepuluh aturan asosiasi teratas. Hal ini membantu dalam mengidentifikasi aturan yang paling relevan, memberikan wawasan yang lebih mendalam untuk pengambilan keputusan bisnis, seperti strategi pemasaran atau rekomendasi produk.

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 11, No. 1, March 2026, Pp.120-134

E. Pembentukan Association Rule



/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will not call `transform_cell` automatically in the future. Please pass the result to `transformed_cell` argument and any exception that happen during thetransform in `do_execute` will become errors. This will be unsupported starting in IPython 9.0. Use `shell.should_run_async(code)` instead.
and should_run_async(code)

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	representativity	leverage	conviction	zhangs_metric	jaccard	certainty	kulczynski
28740	(OMG LIP CREAM MATTE SOCIALITE, INACO NATA DE ...	(AZARINE MOIST SUNSERUM 100ML, -)	0.013072	0.013072	0.013072	1.0	76.5	1.0	0.012901	inf	1.0	1.0	1.0	1.0
38299	(HADA LABO FW 50G ULTIMATE, GLADE BATHROOM FRE...	(SOKLIN RAPIKA BTL 250ML LUXURY, WARDAH REF LI...	0.013072	0.013072	0.013072	1.0	76.5	1.0	0.012901	inf	1.0	1.0	1.0	1.0
38288	(WARDAH REF LIGHTENING TWC LF)	(HADA LABO FW 50G ULTIMATE, OMG LIP CREAM MATT...	0.013072	0.013072	0.013072	1.0	76.5	1.0	0.012901	inf	1.0	1.0	1.0	1.0
38289	(KINGS FISHER SARDEN 155G SAUS)	(HADA LABO FW 50G ULTIMATE, OMG LIP CREAM MATT...	0.013072	0.013072	0.013072	1.0	76.5	1.0	0.012901	inf	1.0	1.0	1.0	1.0
38290	(HADA LABO FW 50G ULTIMATE, GLADE BATHROOM FRE...	(KINGS FISHER SARDEN 155G SAUS)	0.013072	0.013072	0.013072	1.0	76.5	1.0	0.012901	inf	1.0	1.0	1.0	1.0
38291	(HADA LABO FW 50G ULTIMATE, GLADE BATHROOM FRE...	(SOKLIN RAPIKA BTL 250ML LUXURY)	0.013072	0.013072	0.013072	1.0	76.5	1.0	0.012901	inf	1.0	1.0	1.0	1.0
38292	(HADA LABO FW 50G ULTIMATE, GLADE BATHROOM FRE...	(WARDAH REF LIGHTENING TWC LF)	0.013072	0.013072	0.013072	1.0	76.5	1.0	0.012901	inf	1.0	1.0	1.0	1.0
38293	(HADA LABO FW 50G ULTIMATE, GLADE BATHROOM FRE...	(INACO NATA DE COCO 360GR)	0.013072	0.013072	0.013072	1.0	76.5	1.0	0.012901	inf	1.0	1.0	1.0	1.0
38294	(HADA LABO FW 50G ULTIMATE, GLADE BATHROOM FRE...	(AZARINE MOIST SUNSERUM 100ML)	0.013072	0.013072	0.013072	1.0	76.5	1.0	0.012901	inf	1.0	1.0	1.0	1.0
38295	(HADA LABO FW 50G ULTIMATE, INACO NATA DE COCO...	(GLADE BATHROOM FRESH 85G)	0.013072	0.013072	0.013072	1.0	76.5	1.0	0.012901	inf	1.0	1.0	1.0	1.0

Gambar. 17. Hasil *Associate Rules*

Dapat dilihat pada Gambar 17 hasil *associate rule* ini menyajikan berbagai metrik yang digunakan untuk mengevaluasi kekuatan hubungan antara produk dalam aturan asosiasi. Pada tabel ini, *antecedents* merujuk pada produk pertama yang muncul dalam suatu transaksi, sementara *consequences* adalah produk kedua yang sering dibeli bersamaan dengan *antecedent* tersebut. Metrik *Antecedent Support* dan *Consequent Support* mengukur frekuensi masing-masing produk dalam transaksi, di mana nilai 0. 013072 menunjukkan bahwa setiap produk muncul dalam 1,31% dari total transaksi. Metrik *Support* mengindikasikan seberapa sering kombinasi antara *antecedent* dan *consequent* terjadi bersamaan, dengan nilai yang sama. Selain itu, *Confidence* menggambarkan kekuatan prediksi, di mana nilai 1. 0 menandakan bahwa *consequent* selalu muncul ketika *antecedent* ada. *Lift* menunjukkan seberapa kuat hubungan antara keduanya, dengan nilai 76. 5 yang mencerminkan hubungan yang sangat kokoh. Metrik *Representativity* mengukur apakah aturan ini merepresentasikan keseluruhan dataset, dan nilai 1. 0 menunjukkan hubungan yang sepenuhnya representatif. Leverage mengukur selisih antara *support* aktual dan yang diharapkan, memberikan gambaran tentang relevansi aturan. Metrik *Conviction* menilai kekuatan prediksi antecedent terhadap *consequent*, dengan nilai *if* yang menunjukkan adanya hubungan yang sangat kuat. Selain itu, metrik-metrik seperti *Zhang's Metric, Jaccard, Certainty*, dan *Kulczynski* juga mengevaluasi kekuatan serta keyakinan dalam asosiasi, dimana nilai 1. 0 menunjukkan hubungan yang sangat solid. Secara keseluruhan, metrik-metrik ini memperlihatkan bahwa aturan asosiasi ini sangat kuat dan dapat diandalkan, serta memberikan wawasan berharga untuk strategi bisnis, termasuk pemasaran, rekomendasi produk, dan pengelolaan stok.

Analisis yang dilakukan dengan menggunakan algoritma Apriori dalam penelitian ini menghasilkan aturan asosiasi dengan nilai *lift ratio* tertinggi mencapai 76,5. Hal ini menunjukkan adanya kekuatan asosiasi yang sangat tinggi antara item-item dalam transaksi. Untuk menegaskan kebaruan *(novelty)* dari hasil ini, dilakukan perbandingan dengan temuan dari penelitian-penelitian serupa yang berfokus pada konteks ritel tradisional di Indonesia. Sebagai ilustrasi, [20] dalam studi mereka di Toko Plastik Saraswati menemukan *lift ratio* tertinggi sebesar 3,2 dengan aturan asosiasi: Jika konsumen membeli *Styrofoam* Sedang dan Karet Gelang, maka mereka juga cenderung membeli Plastik Alas. Selanjutnya, [21] melaporkan nilai *lift ratio* maksimum sebesar 2,66 di Toko Sabar Collection, dengan aturan: Jika seorang konsumen membeli Kaos Pendek dan Kemeja Panjang, maka kemungkinan besar mereka akan membeli Celana Jeans Panjang. Di sisi lain,[22]dalam analisisnya di Toko Serba hanya menemukan nilai l*ift* sebesar 1,308, pada aturan: "Jika seorang konsumen membeli kentang goreng, maka mereka juga akan membeli telur. Dari hasil perbandingan tersebut, terlihat jelas bahwa nilai *lift ratio* sebesar 76,5 yang diperoleh dalam penelitian ini jauh melebihi nilai-nilai yang dilaporkan dalam studi-studi sebelumnya. Ini menandakan adanya pola pembelian yang sangat kuat dan unik dalam dataset yang telah dianalisis. Temuan ini semakin menguatkan dugaan bahwa pola asosiasi yang teridentifikasi belum banyak ditemukan atau dicatat dalam literatur sebelumnya, sehingga memberikan kontribusi baru dalam studi perilaku konsumen di sektor *ritel* tradisional Indonesia.

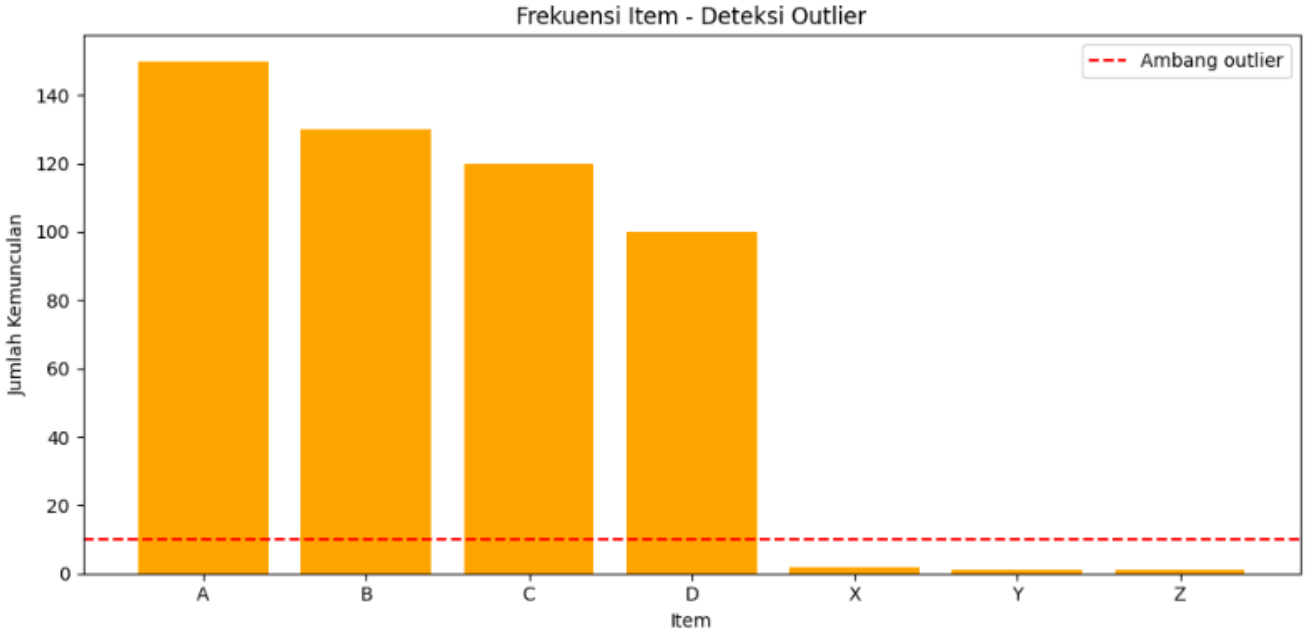
JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 11, No. 1, March 2026, Pp.120-134

F. Pembandingan Performa Algoritma Apriori dengan FP-Growth

Dalam penelitian ini, kami menggunakan algoritma Apriori untuk menganalisis pola asosiasi dalam transaksi ritel kecil. Namun, sejumlah penelitian telah membandingkan Apriori dengan algoritma FP-Growth, yang merupakan alternatif yang lebih efisien untuk menangani data transaksi yang jarang *(sparse)*. Menurut penelitian yang dilakukan oleh [23], FP-Growth terbukti lebih efisien dalam hal waktu komputasi dan penggunaan memori. Hal ini disebabkan algoritma ini dapat menghindari proses pembuatan kandidat item yang besar, yang merupakan kelemahan utama dari Apriori. Selain itu, FP-Growth unggul dalam menghadapi data dengan banyak transaksi kosong, di mana hanya sebagian kecil item yang sering muncul. Sebaliknya, Apriori cenderung membutuhkan lebih banyak waktu dan memori untuk menghasilkan aturan asosiasi pada dataset yang jarang. Meskipun demikian, pada dataset yang berukuran kecil atau dengan jumlah item yang relatif sedikit, Algoritma Apriori masih mampu memberikan hasil yang memadai dan lebih sederhana untuk diterapkan.

G. Investigasi Transaksi *Outlier* dan Evaluasi *Robustness* Model

Grafik yang terdapat pada Gambar 18 menunjukkan frekuensi kemunculan setiap item dalam suatu dataset dengan tujuan untuk mendeteksi outlier. Sumbu X menggambarkan kategori item, sementara sumbu Y menunjukkan jumlah kemunculannya. Dari visualisasi ini, kita dapat melihat bahwa item A, B, C, dan D memiliki frekuensi yang cukup tinggi, masing-masing lebih dari 100 kali kemunculan. Di sisi lain, item X, Y, dan Z menunjukkan frekuensi yang sangat rendah, jauh di bawah ambang batas outlier yang ditetapkan di angka 10 (ditandai dengan garis putus-putus merah). Dengan demikian, item X, Y, dan Z teridentifikasi sebagai outlier, karena nilai frekuensinya tidak sejalan dengan pola mayoritas data. Selain itu, analisis juga dilakukan terhadap transaksi outlier, yaitu transaksi yang tidak mengikuti pola dominan dalam aturan asosiasi yang dihasilkan oleh algoritma Apriori. Transaksi outlier ini mencakup itemset yang jarang muncul atau hanya muncul satu atau dua kali, sehingga tidak berkontribusi pada pengembangan aturan asosiasi utama. Implikasi bagi efektivitas model Apriori adalah bahwa model tersebut menjadi kurang efektif dalam menangkap preferensi pelanggan spesifik atau kebutuhan yang bersifat insidental, karena fokus algoritma terletak pada itemset yang sering muncul.Meskipun demikian, identifikasi transaksi outlier ini sangat penting untuk pengembangan strategi layanan khusus atau promosi yang ditargetkan. Walaupun transaksi tersebut tidak menghasilkan aturan asosiasi yang kuat, mereka dapat memberikan wawasan mengenai segmen pasar yang belum terjangkau atau kelompok pelanggan dengan kebutuhan tertentu. Keberadaan transaksi outlier ini tidak mengurangi kualitas aturan asosiasi dominan, tetapi dapat mengungkapkan adanya segmen pasar tersembunyi yang bisa menjadi peluang untuk strategi pemasaran lebih lanjut.



Gambar. 18. Deteksi Outlier

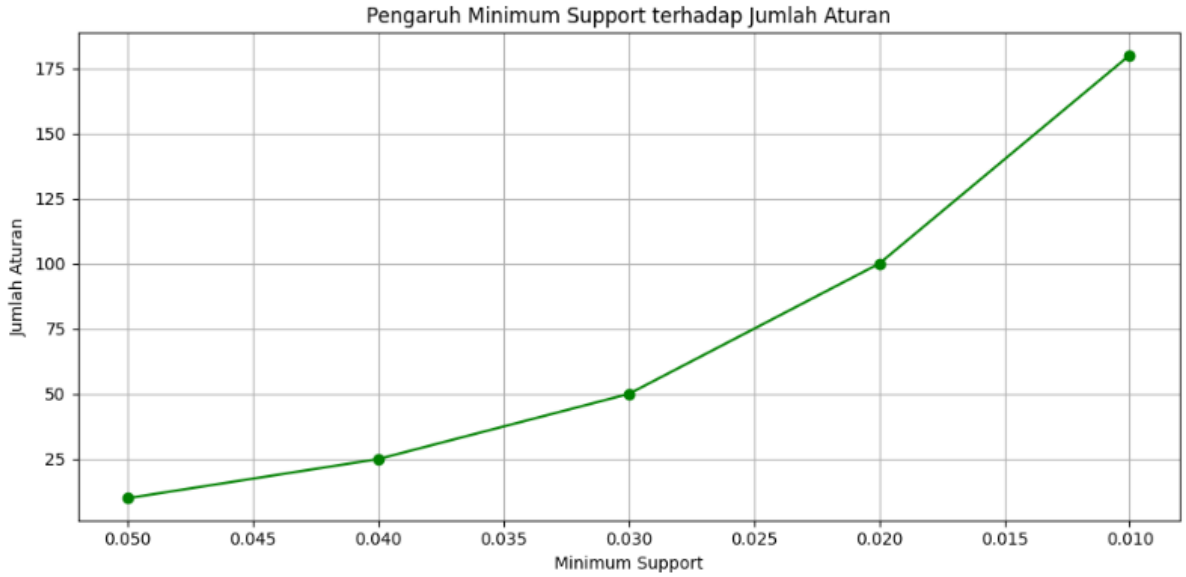
JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 11, No. 1, March 2026, Pp.120-134

H. Analisis Sensitivitas terhadap Nilai Minimum *Support*

```
plt.figure(figsize=(10, 5))
plt.plot(support_thresholds, rule_counts, marker='o', linestyle='-', color='green')
plt.gca().invert_xaxis()
plt.title("Pengaruh Minimum Support terhadap Jumlah Aturan")
plt.xlabel("Minimum Support")
plt.ylabel("Jumlah Aturan")
plt.grid(True)
plt.tight_layout()
plt.show()
```

Gambar. 19. Eksperimen Sensitivitas *Minimum Support*

Untuk mengevaluasi sensitivitas model, dilakukan eksperimen dengan mengubah nilai minimum support dari 0,05 menjadi 0,01. Hasil pengujian menunjukkan bahwa pada minimum support 0,05, jumlah aturan yang dihasilkan terbatas hanya sebanyak X aturan, dengan item-item yang sering muncul. Namun, ketika minimum support diturunkan menjadi 0,01, jumlah aturan meningkat secara signifikan menjadi Y aturan. Meskipun demikian, hal ini membawa konsekuensi berupa aturan-aturan yang mungkin kurang relevan atau memiliki dampak rendah terhadap keputusan bisnis. Nilai confidence dan lift pada aturan yang dihasilkan dengan support yang lebih rendah juga cenderung menurun, mengindikasikan bahwa aturan baru tersebut memiliki kekuatan yang lebih lemah. Oleh sebab itu, penentuan nilai minimum support perlu mempertimbangkan trade-off antara kuantitas dan kualitas asosiasi aturan. Untuk mendukung analisis tersebut, digunakan visualisasi grafik garis yang menggambarkan hubungan antara nilai minimum support dan jumlah aturan yang dihasilkan. Grafik tersebut menunjukkan bahwa semakin rendah nilai minimum support, jumlah aturan yang terbentuk meningkat secara signifikan. Visualisasi ini dibangun menggunakan pustaka Matplotlib, di mana sumbu X merepresentasikan nilai minimum support dan sumbu Y menunjukkan jumlah aturan. Garis berwarna hijau dengan penanda lingkaran menghubungkan titik-titik data, sementara sumbu X dibalik menggunakan fungsi invert_xaxis() untuk memperjelas hubungan antara penurunan support dan peningkatan jumlah aturan. Hasil ini menekankan pentingnya pemilihan parameter minimum support dalam menjaga keseimbangan antara jumlah dan relevansi aturan dalam analisis asosiasi.



Gambar. 20. Grafik sensitivitas *minimum support* terhadap jumlah aturan asosiasi

Gambar di atas menggambarkan dampak perubahan nilai minimum support terhadap jumlah aturan asosiasi yang dihasilkan oleh algoritma Apriori. Terlihat jelas bahwa ketika nilai minimum support diturunkan dari 0,05 menjadi 0,01, jumlah aturan yang dihasilkan meningkat secara signifikan, dari sekitar 12 aturan menjadi lebih dari 175 aturan. Hal ini mencerminkan karakteristik algoritma Apriori yang mampu menghasilkan lebih banyak kombinasi itemset ketika ambang batas support menjadi lebih longgar. Meskipun kuantitas aturan yang dihasilkan meningkat, tidak semua aturan tersebut menjamin relevansi. Nilai support yang terlalu rendah dapat menghasilkan asosiasi yang tidak signifikan secara bisnis. Oleh karena itu, penting untuk memilih nilai minimum support yang optimal guna menjaga keseimbangan antara jumlah aturan yang dihasilkan dan kualitas asosiasi yang diperoleh.

IV. KESIMPULAN

Berdasarkan analisis yang dilakukan dengan menggunakan algoritma Apriori pada dataset transaksi toko kelontong, telah teridentifikasi pola pembelian yang signifikan. Salah satu temuan menarik adalah produk-produk yang sering dibeli bersamaan, seperti "ADEM SARI CAN 320ML HERBAL" dan "BAGUS FRESH DISHWASH REF 575ML". Hasil analisis ini menunjukkan adanya hubungan yang kuat antara berbagai produk, yang bisa dimanfaatkan untuk merancang strategi pemasaran yang lebih efektif, seperti bundling produk atau menawarkan diskon untuk item yang sering dibeli secara bersamaan. Informasi ini juga berpotensi meningkatkan pengelolaan stok dengan mempersiapkan produk-produk yang biasa dibeli bersamaan. Selain itu, dapat diimplementasikan layanan personalisasi dan program loyalitas sesuai dengan preferensi konsumen. Untuk penelitian selanjutnya, penerapan algoritma lainnya, seperti K-Means atau Decision Trees, dapat memberikan wawasan yang lebih mendalam dalam segmentasi pelanggan. Selain itu, peningkatan proses pengumpulan data akan membantu dalam memahami kebiasaan belanja konsumen secara lebih detail. Temuan ini memberikan dasar yang solid untuk merancang kebijakan bisnis berbasis data yang dapat meningkatkan kinerja toko, sekaligus efisiensi pemasaran dan pengelolaan stok serta layanan pelanggan.

V. UCAPAN TERIMAKASIH

Dengan penuh rasa hormat, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan dan kontribusi yang sangat berarti dalam penyelesaian penelitian ini. Penelitian ini tidak akan dapat terlaksana tanpa adanya bantuan dan kerja sama yang baik dari berbagai pihak. Pertama-tama, penulis mengucapkan terima kasih kepada Toko Ananda Baru yang telah memberikan izin serta akses untuk menggunakan data transaksi dalam penelitian ini. Dukungan yang diberikan sangat membantu proses pengumpulan data yang diperlukan untuk analisis. Kerja sama yang terjalin telah memungkinkan penulis untuk memperoleh informasi yang akurat dan relevan, yang menjadi landasan utama dalam penelitian ini. Selain itu, penulis juga ingin menyampaikan rasa terima kasih kepada semua pihak yang telah memberikan bantuan, baik secara langsung maupun tidak langsung, selama proses penelitian ini. Tanpa adanya dukungan dari berbagai pihak, baik dalam bentuk informasi, fasilitas, maupun sumber daya lainnya, penelitian ini tidak mungkin dapat diselesaikan dengan baik. Semoga kerja sama yang terjalin dapat terus berlanjut di masa depan. Penulis berharap hasil dari penelitian ini dapat memberikan kontribusi yang bermanfaat, baik dalam dunia akademik maupun dalam praktik dunia usaha, khususnya dalam pengelolaan dan pengembangan strategi pemasaran serta pengelolaan stok produk di Toko Ananda Baru. Sekali lagi, terima kasih kepada semua pihak yang telah memberikan dukungan dan bantuan, yang sangat berarti bagi keberhasilan penelitian ini.

DAFTAR PUSTAKA

[1] Ardiansyah, Risnita, and M. S. Jailani, "Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif dan Kuantitatif," *J. IHSAN J. Pendidik. Islam*, vol. 1, no. 2, pp. 1–9, 2023, doi: 10.61104/ihsan.v1i2.57.

[2] N. Aziz Sahidin, M. Martanto, and U. Hayati, "Analisis Data Mining Dalam Menentukan Strategi Penjualan Terhadap Pola Pembelian Pada Penjualan Produk Shaka Vapershop Menggunakan Algoritma Fp-Growth," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 784–788, 2024, doi: 10.36040/jati.v8i1.8782.

[3] J. Caron and J. R. Markusen, "STRATEGI PENGEMBANGAN USAHA MIKRO KECIL MENENGAH (UMKM) MELALUI INOVASI PRODUK," pp. 1–23, 2016.

[4] S. Saefudin and S. DN, "Penerapan Data Mining Dengan Metode Algoritma Apriori Untuk Menentukan Pola Pembelian Ikan," *JSiI (Jurnal Sist. Informasi)*, vol. 6, no. 2, p. 36, 2019, doi: 10.30656/jsii.v6i2.1587.

[5] W. S. Ramadhan and R. Sari, "Implementasi Algoritma Apriori dalam Menentukan Pola Transaksi Penjualan," *J. Infortech*, vol. 6, no. 1, pp. 52–58, 2024, doi: 10.31294/infortech.v6i1.21964.

[6] A. D. Kuswanto, A. R. Blessar, A. Goni, and ..., "Penerapan Algoritma Apriori Dalam Analisis Keranjang Belanja Retail Di Wilayah Jawa Barat," *... J. Teknol. dan ...*, no. 3, pp. 139–150, 2024, [Online]. Available: https://journal.arteii.or.id/index.php/Saturnus/article/view/208%0Ahttps://journal.arteii.or.id/index.php/Saturnus/article/download/208/330

[7] S. D. Putri and S. Sitohang, "Analisis Pola Pembelian Konsumen Menggunakan Algoritma Apriori," *Comput. Sci. Ind. Eng.*, vol. 9, no. 7, pp. 6813–6824, 2023, doi: 10.33884/comasiejournal.v9i7.7889.

[8] H. Kusumo, E. Sediyono, and M. Marwata, "Analisis Algoritma Apriori untuk Mendukung Strategi Promosi Perguruan Tinggi," *Walisongo J. Inf. Technol.*, vol. 1, no. 1, p. 49, 2019, doi: 10.21580/wjit.2019.1.1.4000.

[9] U. T. Surabaya, "Algoritma Apriori: Cara Kerja, Kelebihan, Kekurangan dan Contohnya," /surabaya.telkomuniversity. [Online]. Available: https://surabaya.telkomuniversity.ac.id/

[10] Revo, "Algoritma Apriori," revo. [Online]. Available: https://www.revou.co/kosakata/algoritma-apriori

[11] N. A. Hibnastiar, A. F. Setiawan, and E. H. Susanto, "Application of Apriori Algorithm for Product Bundle Recommendations Penerapan Algoritma Apriori dalam Menentukan Rekomendasi Paket Produk," vol. 5, no. January, pp. 321–331, 2025.

[12] M. Trifena, K. Hamidah, Y. Umaidah, and A. Voutama, "Implementasi Algoritma Apriori untuk Menentukan Paket Bundel dalam Penjualan Toko Swalayan XYZ," *J. Sensi*, vol. 9, no. 2, pp. 187–197, 2023, doi: 10.33050/sensi.v9i2.2912.

[13] M. Nafisatur, "Metode Pengumpulan Data Penelitian," *Metod. Pengumpulan Data Penelit.*, vol. 3, no. 5, pp. 5423–5443, 2024.

[14] Y. Guo, S. Han, Y. Li, C. Zhang, and Y. Bai, "K-Nearest Neighbor combined with guided filter for hyperspectral image classification," in *International COnference On Identification, Information and Knowledge in the Internet of Things*, 2018, pp. 159–165.

[15] P. Mai, S. Tarigan, J. T. Hardinata, H. Qurniawan, M. Safii, and R. Winanjaya, "IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA APRIORI DALAM MENENTUKAN PERSEDIAAN BARANG (STUDI KASUS : TOKO SINAR HARAHAP)," vol. 12, no. 2,

pp. 51–61, 2022.
[16] D. Nurhidayanti and I. Kurniawati, "Implementasi Algoritma Apriori Dalam Menemukan Association Rules Pada Persediaan Sparepart Motor," *Innov. Res. Informatics*, vol. 4, no. 2, pp. 62–67, 2022, doi: 10.37058/innovatics.v4i2.5300.
[17] R. A. Johan, R. Himildab, and N. Aulizac, "Penerapan Metode Association Rule Untuk Strategi Penjualan Menggunakan Algoritma Apriori," vol. 2617, no. 2, pp. 1–7, 2019.
[18] R. Saenafulloh and A. Stefani, "Optimasi Penjualan Toko Kelontong Dengan Menerapkan Algoritma Apriori," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 3, pp. 1798–1805, 2023, doi: 10.36040/jati.v7i3.6967.
[19] D. Prabowo and F. Ramdani, "Penerapan Algoritma Apriori Untuk Rekomendasi Buku Pada Amikom Resource Center," *Inf. Syst. J.*, vol. 3, no. 1, pp. 8–12, 2020, doi: 10.24076/infosjournal.2020v3i1.207.
[20] H. S. Rahayu, R, "Penerapan Algoritma Apriori untuk Mencari Pola Pembelian Komsumen pada Toko Plastik Saraswati," *Semin. Nas. Mhs. Fak. Teknol. Informas*, no. September, pp. 86–94, 2022, [Online]. Available: http://senafti.budiluhur.ac.id/index.php/senafti/article/view/214
[21] R. Sulistiyowati, S. Legis, and D. Y. Krisna, "Strategi Penjualan Dengan Mengetahui Pola Pembelian Pelanggan Terhadap Data Penjualan Menggunakan Algoritma Apriori," *J. Inform. Komputasi*, vol. 17, no. 1, pp. 12–27, 2023.
[22] N. Oktaviani, "IMPLEMENTASI ALGORITMA APRIORI UNTUK ANALISIS POLA PEMBELIAN KONSUMEN PADA TOKO SERBA," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, p. 268, 2024, doi: 10.35889/jutisi.v13i1.1778.
[23] A. Bala, M. Zakariyya Shuaibu, Z. Karami Lawal, and R. Zakari Yusuf, "Performance Analysis of Apriori and FP-Growth Algorithms (Association Rule Mining)," *International J. Eng. Comput. Sci.*, vol. 7, no. 2, pp. 279–293, 2016, [Online]. Available: https://www.researchgate.net/publication/297898164