

IMPLEMENTATION OF DATA MINING ON MUSLIM WOMEN'S CLOTHING SALES USING THE FP-GROWTH METHOD

Riko Aldinata¹, Raissa Amanda Putri²

^{1,2}Computer Science, Universitas Islam Negeri Sumatera Utara
rikoaldinata.io@gmail.com¹, raissa.ap@uinsu.ac.id²

Abstract

The Muslim women's fashion industry in Indonesia is growing rapidly, leading to intense competition and requiring business owners to optimize their sales strategies and inventory management. This study aims to identify consumer purchasing patterns at TM Collection Store by applying the FP-Growth data mining method. The data used consists of 1,000 sales transactions from January to April 2024. Data collection was conducted through historical data observation, interviews, and literature review, followed by processing using the FP-Growth algorithm in Google Colab. The analysis results reveal strong associations between products, such as the combination of Paris Premium, shirt cuffs XL, and shirt cuffs L, which show high confidence values and significant lift. These patterns provide valuable insights for decision-making related to restocking and promotional strategies. The findings also help improve operational efficiency by more accurately predicting customer demand. Therefore, the implementation of the FP-Growth algorithm proves effective in processing transaction data to generate relevant information and support more targeted business decisions. This data-driven strategy offers an innovative solution to enhance competitiveness in the continuously growing Muslim women's fashion industry.

Keywords: FP-Growth; Purchasing Patterns; Operational Efficiency

Abstrak

Industri fashion Muslim wanita di Indonesia berkembang pesat, memicu persaingan yang ketat dan menuntut pelaku usaha untuk mengoptimalkan strategi penjualan serta manajemen stok. Penelitian ini bertujuan mengidentifikasi pola pembelian konsumen di Toko TM Collection dengan menerapkan metode data mining FP-Growth. Data yang digunakan berupa 1.000 transaksi penjualan dari Januari hingga April 2024. Pengumpulan data dilakukan melalui observasi data historis, wawancara, dan studi literatur, kemudian diproses menggunakan algoritma FP-Growth di Google Colab. Hasil analisis mengungkap adanya asosiasi kuat antar produk, seperti kombinasi Paris Premium, Manset Baju XL, dan Manset Baju L, yang menunjukkan nilai confidence tinggi dan lift signifikan. Pola ini memberikan wawasan penting dalam pengambilan keputusan terkait restocking barang dan strategi promosi. Temuan ini juga membantu meningkatkan efisiensi operasional toko dengan memperkirakan permintaan konsumen secara lebih akurat. Dengan demikian, penerapan algoritma FP-Growth terbukti efektif dalam mengolah data transaksi untuk menghasilkan informasi yang relevan dan mendukung pengambilan keputusan bisnis yang lebih tepat sasaran. Strategi berbasis data ini menjadi solusi inovatif untuk meningkatkan daya saing dalam industri fashion Muslim wanita yang terus berkembang.

Kata kunci: Data Mining; FP-Growth; Pola Pembelian; Efisiensi Operasional

INTRODUCTION

The fashion industry, particularly in the category of Muslim women's clothing, has experienced significant growth in recent years (Srisusilawati et al., 2024). To this day, the Muslim fashion industry has rapidly developed and become a trend among Muslim women, especially in Indonesia (Sudarmi et al., 2024). The increasing interest and awareness of consumers toward Muslim women's fashion have also expanded the

market for this industry. However, this growth has also led to increasingly fierce competition in the Muslim women's clothing sector.

TM Collection, a store that has long operated in the field of Muslim women's clothing sales, has faced several challenges in recent years. Despite having a loyal customer base, TM Collection has experienced a decline in operational efficiency and has not yet been able to fully optimize its sales potential. One of the main causes of this issue is the underutilization of available sales transaction data.

Until now, sales transaction data at TM Collection has only been used for recording and reporting purposes, without any in-depth analysis. As a result, store management has difficulty identifying customer purchasing patterns and tends to restock products based on rough estimates rather than accurate data. For example, some products are overstocked, while other popular items quickly run out of stock (Febiyanto et al., 2024). This situation not only leads to stock inefficiencies but also has the potential to cause revenue losses (Suhada et al., 2020).

In addition, TM Collection struggles to design targeted promotional strategies. Promotions are often not based on actual customer purchasing behavior, making them less effective (Destiawati et al., 2024). With increasingly fierce competition in the Muslim women's clothing market, this situation becomes even more detrimental to TM Collection. To address these issues, the store owner has begun considering the application of data mining techniques to process the available transaction data (Hafizh et al., 2023). Data mining allows the store to uncover valuable insights from previously overlooked data (Alam et al., 2024). One method considered suitable for TM Collection's needs is FP-Growth, a data mining technique effective in identifying frequent purchasing patterns within large datasets (Sugianto & Sukmawati, 2023).

Data mining is a systematic analytical process aimed at uncovering hidden patterns, meaningful information, and new knowledge from large datasets (Jawara et al., 2025). This process not only extracts consistent and relevant information but also supports data-driven decision-making by leveraging statistical techniques, machine learning, and artificial intelligence (Ramdani & Utami, 2022). In the context of research and data-based system development, data mining serves as a vital foundation for discovering insights that conventional analysis methods may fail to identify (Nugroho et al., 2022).

FP-Growth was chosen due to its advantages in data processing efficiency compared to other association methods, such as Apriori (Helyatin et al., 2024). The Apriori method requires repeated iterations to find frequent item combinations (Aktavera et al., 2024), which consumes time and resources, especially with large datasets. In contrast, FP-Growth uses a more efficient structure called the FP-tree, which allows pattern searching to be done in a single data traversal (Atmaja & Rachman, 2025). This makes FP-Growth faster and more capable of handling large volumes of data effectively (Anas, 2020).

FP-Growth is one of the most efficient algorithms for discovering frequent itemsets from large datasets (Andriyanti et al., 2025). It utilizes a special data structure known as the FP-Tree (Frequent Pattern Tree) to represent frequently occurring items without explicitly generating candidate itemsets as in Apriori (Valencia & Atmojo, 2024). This approach makes the pattern search process faster and more memory-efficient, as it stores only relevant information from transactional data (Winarti et al., 2021). In data mining, FP-Growth is widely used in market basket analysis, product recommendation systems, and exploring associations among items (Dwi Insani & Al Fatta, 2023).

By applying FP-Growth to the sales data at TM Collection, the store is expected to optimize product availability (Irawan et al., 2024), design more targeted promotions, and improve operational efficiency and profitability. Additionally, the analysis results will provide better insights into customer behavior, enabling TM Collection to be more responsive to market needs.

The purpose of this study is to provide a practical solution for TM Collection through the application of data mining using the FP-Growth method. By identifying frequent itemsets and analyzing purchasing patterns, this method is expected to help the store optimize its sales strategies, develop more effective marketing campaigns, improve inventory management, and ultimately enhance its competitiveness in the market.

Previous research focused on utilizing the FP-Growth algorithm to predict the availability of essential goods stock in order to support smooth sales operations (Aditiya et al., 2020). In contrast, this study implements the FP-Growth algorithm in a different context, namely in the sales of Muslim women's clothing, with the aim of analyzing consumer purchasing patterns and identifying combinations of products that are frequently bought together. The main differences lie in the research object (essential goods vs. fashion), the purpose of the algorithm (stock prediction vs. sales pattern analysis), and the contribution of the study (inventory management vs. marketing strategy). Therefore, this research fills the gap by presenting the application of FP-Growth in the fashion industry as a strategy to increase sales through an understanding of product association patterns.

RESEARCH METHODS

The stages carried out in this study can be seen in Figure 1.

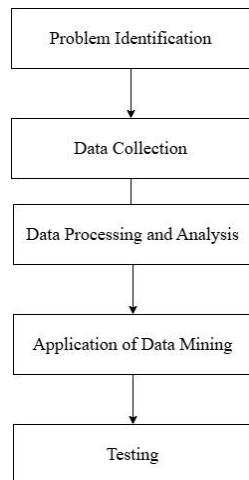


Figure 1 Research Framework

Problem identification

Problem identification is a crucial initial step to understand the context and objectives of the research. The main issues in this context may involve the complexity of sales patterns for Muslim women's clothing, which are difficult to analyze manually. Additionally, there may be challenges in efficiently managing inventory or in deeply understanding customer preferences. This research aims to address these problems by applying data mining techniques, particularly the FP-Growth method, to identify significant purchasing patterns, product associations, and other factors that can provide valuable insights for store owners or producers of Muslim women's clothing.

Data collection

This research employs a Mixed Methods Research approach to obtain relevant data regarding purchasing patterns at TM Collection store. Data is collected from sales transactions using the following techniques:

1. Observation

Data collection was carried out by gathering historical sales data from a Muslim clothing store located in Medan City. The process involved direct observation of sales data at TM Collection, Medan. The observed data included parameters such as the number of daily sales transactions, categories of products sold, restocking frequency, and combinations of products frequently purchased together by customers. Historical sales data was taken from the store's sales system over a 4-month period, from January to April 2024, with a total of 1,000 transactions analyzed. Collected values included the number of units sold per product, time of

purchase, and product combinations that frequently appeared together in a single transaction.

2. Interviews

Data was gathered by conducting interviews with relevant parties at TM Collection, such as the store owner and sales staff. The purpose of the interviews was to gain a deeper understanding of the store's operations, the sales strategies implemented, and the challenges faced—especially in terms of inventory management and promotions. The results of these interviews complement the historical sales data and provide additional insights for deeper analysis.

3. Literature Review

A literature review was conducted to collect data by studying various references that serve as the foundation for the research. In this case, the researcher explored theories related to the study obtained through online sources, including scientific articles and academic textbooks.

Data Processing and Analysis

This stage is one of the key phases that determine the success of the research. In this phase, data processing and analysis will be conducted through a data preprocessing stage. Several steps will be carried out to prepare the data so that it is ready for the application of the FP-Growth algorithm. The following are the data preprocessing steps that will be undertaken:

1. Data Cleaning, by identifying and handling erroneous data, such as missing values or duplicates.
2. Data Transformation, by converting the data into a format suitable for the application of the FP-Growth algorithm.

Application of Data Mining

The next step is to implement the FP-Growth algorithm using the Google Colab tool. The processed dataset will then be used to apply the FP-Growth algorithm in order to identify association patterns within the sales transaction data.

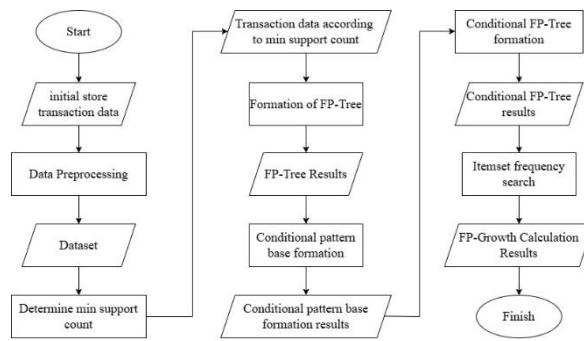


Figure 2 FP-Growth algorithm flowchart

RESULTS AND DISCUSSION

Problem Identification

This study focuses on the application of data mining techniques to analyze sales patterns of Muslim women's clothing using the FP-Growth algorithm. This method is chosen for its ability to efficiently extract item associations without explicitly generating candidate itemsets. The stages carried out in this research include calculating the frequency of item occurrences in transactions, determining support values to identify frequent itemsets, constructing the FP-Tree structure as a compressed representation of the data, and calculating confidence values to generate strong association rules. Through this approach, the study is expected to provide valuable insights for business practitioners in making strategic decisions regarding Muslim fashion product sales.

By clearly identifying the core issues, this study is expected to make a meaningful contribution to enhancing the effectiveness of sales and inventory management. Additionally, it has the potential to increase customer satisfaction by providing a deeper understanding of their needs and preferences. The findings of this research can serve as a foundation for smarter, data-driven decision-making in the Muslim women's clothing industry and support the implementation of more adaptive and responsive marketing strategies that align with evolving market trends.

Data Collection

The data used in this study consists of 1,003 transactions, obtained directly from the Muslim Clothing Store TM Collection. The dataset reflects the purchase history of customers over a specific period and contains information about combinations of Muslim women's clothing products that were purchased together. This substantial amount of data allows for a more representative and relevant analysis of purchasing patterns and provides a clear picture of customer preferences for the products offered by the store. This is illustrated in the following table.

Table 1 Transaction data

No	Date	Name of goods
1	01/01/2024	8 jersey pc gamis, yukensi gamis, L shirt cuff
2	01/01/2024	Paris premium, voal thorsa

3	01/01/2024	Paris premium, XL shirt cuff
4	01/01/2024	Yukensi std shirt cuff
5	01/01/2024	Voal thorsa cuff shirt L
...
1003	30/04/2024	3L shirt cuff, thorsa voal, vergano khimar

Data Processing and Analysis

In the data preprocessing stage, TransactionEncoder is used to convert the transaction data, originally in the form of a list of lists, into a tabular representation as a DataFrame. This process involves identifying all unique items in the dataset and representing each transaction as a row with Boolean values (True/False) to indicate the presence of an item. This transformation is essential to ensure the data aligns with the input format required by association algorithms such as FP-Growth. The transformed data is shown in the following table.

Table 2 Preprocessing data

No	Leg ing 31	Buc kle Scra ves	T- shi rt ca p	Pash mina Cerut y	Pash mina T- shirt	Tho rsa Voa l
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	True
4	True	False	False	False	False	False
...
99	False	False	False	False	False	False
99	False	False	False	False	False	False



99	False						
9	False						
10	False						
00	False						
10	False						
03	False						

Application of Data Mining

The process of identifying itemset frequencies plays a vital role in association rule analysis, as it forms the foundation for uncovering meaningful relationships between items within transaction data. Itemsets refer to groups of items or product characteristics that commonly occur together in a dataset. By analyzing these frequent itemsets, businesses can gain insights into customer purchasing behavior and product bundling opportunities. The results of the itemset frequency analysis, which highlight the most commonly purchased product combinations, are presented in Table 3 below.

Table 3 Frequent item

No	Product Name	Frequency
1	gamis pc 8 jersey	77
2	gamis yukensi	101
3	baju inner L	174
4	premium paris	126
5	voal thorsa	122
6	baju inner XL	168
7	yukensi inner standard	52
8	baju inner 3L	132
9	leggings 3L	76
10	khimar vergano	82
11	paris olive	78
12	yukensi standard inner	81
13	leggings L	109
14	body inner L	46
15	body inner 3L	60
16	buckle scarves	63
17	gamis cotton standard	101
18	pashmina ceruty	83

19	leggings XL	123
20	cotton gamis	75
21	leggings 6L	61
22	jumbo cotton gamis	69
23	knitted ciput	84
24	cotton ciput	144
25	cotton pashmina	68
26	jumbo yukensi inner	75
27	gamins pc 8 jersey	1
28	buckle thorsa	1
29	crinkle gamis	71
30	baju inner	62
31	gaming ceruty	1
32	ceruty gamis	1
33	body inner 2L	2
	Total	2569

Construction of FP-Tree

The following are the results of the FP-Tree structure construction process based on transaction data that has been filtered and sorted according to item frequency.

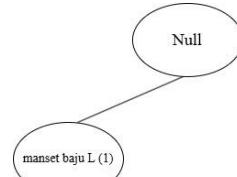


Figure 3 First Transaction FP-Tree

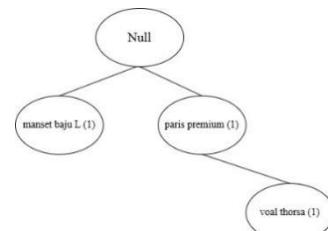


Figure 4 Second Transaction FP-Tree

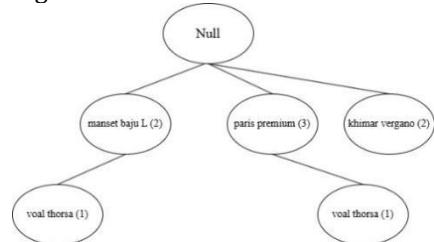


Figure 5 Tenth Transaction FP-Tree

Determining Itemset Frequency

The results of the support value calculations will be presented in a table that shows the frequency of occurrence of each itemset within the transaction data. This table provides a clear overview of how often individual items or combinations of items appear together in transactions. The structured data allows for further analysis to determine which itemsets meet the predefined minimum support threshold and helps identify relevant purchasing patterns to support business decision making.

Table 4 Support Value

Itemset	Support
Voal Thorsa	0.073852
Paris Premium	0.070858
Manset Baju XL	0.067864
Manset Baju Yukensi Std	0.012974
Manset Baju L	0.089820
...	...
Manset Baju, Gamis PC 8 Jersey	0.005988
Gamis Yukensi, Gamis Kringkel	0.007984
Manset Yukensi Std, Gamis Kringkel	0.006986
Manset Yukensi Jumbo, Gamis Kringkel	0.007984
Gamis Kringkel, Phasmina Ceruty	0.005988

Testing

In the testing phase of the Association Rules, an evaluation is conducted to assess the quality and relevance of the association rules generated from the data mining process. The evaluation results are presented in a table using the following parameters: a minimum support of 0.005, a confidence level above 0.5, and a minimum lift value greater than 1.0. A total of four association rules were found to provide insights into sales patterns within the transaction dataset of TM Collection store.

The evaluation results of the Association Rules are presented in the following table.

Atenced ents	Consequ ents	Suppo rt	Confid ence	Lift
Paris Premiu m,	Manset Baju L	0.005 988	0.8571 43	9.542 857

Manset				
Baju XL				
Paris				
Premiu m,	Manset	0.005	0.6666	9.823
Manset	Baju XL	988	67	529
Baju L				
Manset				
Baju XL,	Paris	0.005	0.5000	7.056
Manset	Premiu m	988	00	338
Baju L				
Manset				
Baju XL,	Manset	0.005	0.5000	5.566
Ciput	Baju L	988	00	667
Kaos				

Based on the implementation of the FP-Growth algorithm and the generation of association rules, four association rules were identified that meet the predefined minimum support and confidence thresholds. These four rules reveal combinations of items frequently purchased together by customers and highlight the interrelationships between products. The following is an explanation of the results obtained:

1. The first rule indicates that if customers purchase paris premium and manset baju XL, they are highly likely to also buy manset baju L, with a confidence level of 85.71% and a lift value of 9.54, indicating a very strong association.
2. The second rule states that the combination of manset baju L and manset baju XL tends to be followed by the purchase of paris premium, with a confidence of 50% and a lift of 7.05, suggesting a fairly strong relationship.
3. The third rule suggests that customers who buy paris premium and manset baju L have a 66.67% chance of also purchasing manset baju XL, with a lift value of 9.82, indicating a highly significant association.
4. The fourth rule shows that customers who buy manset baju XL and ciput kaos are also likely to purchase manset baju L, with a confidence of 50% and a lift of 5.57, which also reflects a strong relationship.

These four rules provide valuable insights for TM Collection Store in developing more effective marketing strategies, such as product bundling promotions, stock rearrangement, or personalized purchase recommendation systems for customers.

CONCLUSIONS

Conclusion

Based on the results of the study and the application of the FP-Growth algorithm on TM Collection Store's transaction data, it can be concluded that FP-Growth is effective in identifying consumer purchasing patterns for Muslim women's fashion products. Through stages such as data preprocessing, FP-Tree construction, support calculation, and association rule generation, four association rules were discovered that reveal strong relationships between products frequently bought together. For example, the combination of paris premium and manset baju XL is often followed by the purchase of manset baju L with a confidence of 85.71% and a lift of 9.54. These findings provide valuable insights for sales strategies such as product bundling, purchase recommendations, and more efficient stock management, proving that FP-Growth is an appropriate method for data-driven sales analysis.

Suggestion

It is recommended to expand the data scope and transaction period so that the resulting patterns are more representative and relevant to consumer purchasing behavior in the long term. The author may also explore comparisons with other algorithms such as Apriori or Eclat to evaluate the effectiveness and efficiency.

REFERENCES

Aditiya, R., Defit, S., & Nurcahyo, G. W. (2020). Prediksi Tingkat Ketersediaan Stock Sembako Menggunakan Algoritma FP-Growth dalam Meningkatkan Penjualan. *Jurnal Informatika Ekonomi Bisnis*, 2, 67–73. <https://doi.org/10.37034/infeb.v2i3.44>

Aktavera, B., Oktafia, H., Wijaya, L., Studi, P., Komputer, I., Petulai, U. P., Lebong, R., Studi, P., Informasi, S., Insan, U. B., Apriori, A., & Apriori, A. (2024). Analisis Association Rule Menggunakan Algoritma Apriori dan Algoritma FP Growth. *Jurnal Teknologi Informasi Mura*, 16(1), 54–61.

Alam, F. N., Setyawan, P. A., & Dwianto, S. (2024). *Prediksi ketersediaan stok bulanan kategori susu di toko koperasi menggunakan metode decision tree*. November, 25–34.

Anas, A. (2020). Penerapan Algoritma Fp-Growth Dalam Menentukan Perilaku Konsumen Ghania Mart Muara Bulian. *Jurnal Ilmiah Media Sisfo*, 14(2), 120–129.

Andriyanti, R., Rahaningsih, N., & Ali, I. (2025). *STUDI MODEL ASOSIASI PADA PENJUALAN KOPI BERBASIS ALGORITMA FP-GROWTH*. 9(1), 5113–5119.

Atmaja, G. bayu, & Rachman, R. (2025). *PERBANDINGAN ALGORITMA APRIORI DAN FP-GROWTH PADA ANALISIS PERILAKU KONSUMEN TERHADAP PEMBELIAN DATA ELEKTRONIK*. 298–307.

Destiawati, D., Rahaningsih, N., Bahtiar, A., Ali, I., & Dienwati Nuris, N. (2024). Analisis Pola Penjualan Menggunakan Algoritma Asosiasi Fp-Growth Di Pt Abc. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 3405–3410. <https://doi.org/10.36040/jati.v8i3.9719>

Dwi Insani, F., & Al Fatta, H. (2023). Perbandingan Algoritma Apriori Dan Fp-Growth Untuk Rekomendasi Item Paket Pada Konten Promosi Di Perusahaan Mu-Mart. *Jurnal Bisnis Digital Dan Sistem Informasi*, 4(2), 19–24.

Febiyanto, A., Faqih, A., Herdiyana, R., Dienwati Nuris, N., & Narasati, R. (2024). Penerapan Algoritma Fp-Growth Untuk Menentukan Pola Penjualan Produk Elektronik. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3907–3912. <https://doi.org/10.36040/jati.v7i6.8286>

Hafizh, M., Novita, T., Guswandi, D., Syahputra, H., & Mayola, L. (2023). Implementasi Data Mining Menggunakan Algoritma Fp-Growth Untuk Menganalisa Transaksi Penjualan Eksport Online. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 5(3), 242–249. <https://doi.org/10.47233/jtekstis.v5i3.847>

Helyatin, A. H., Nadzirotul, N., Syarif, F., Khoiri, A., & Fauzen, M. (2024). *JUSTIFY: Jurnal Sistem Informasi Ibrahimy PERANCANGAN SISTEM INFORMASI PRESENSI GURU BERBASIS WEB*. 3(1), 21–29. <https://doi.org/10.35316/justify.v3i1.5565>

Irawan, Y. P., Gema, R. L., Elva, Y., & Komputer, F. I. (2024). *Optimasi persediaan produk penjualan menggunakan metode economic order quantity*. 8(1), 153–162.

Jawara, I. R., Fatah, Z., Ibrahimy, U., Timur, S. J., Ibrahimy, U., & Timur, S. J. (2025). *PENGGUNAAN DATA MINING UNTUK MEMPREDIKSI PENJUALAN PADA TOKO*. 2(1), 52–60.

Nugroho, B. I., Ma'arif, Z., & Arif, Z. (2022). Tinjauan Pustaka Sistematis: Penerapan Data Mining Metode Klasifikasi Untuk Menganalisa Penyalahgunaan Sosial Media. *Jurnal Sistem Informasi Dan ...*, 3(2), 46–51. <http://journal.peradaban.ac.id/index.php/jsti/article/download/1265/860>

Ramdani, F., & Utami, I. Q. (2022). *Pengantar Data Science*.

Srisusilawati, P., Prasetyo, S. N., Nur Hamidah, S. A., Rihhadatull 'Aisy, R. A., & Oktavia, R. (2024).



Tren dan Perkembangan Fashion Syariah Pada Era Modern di Kota Bandung. *Jurnal Ilmiah Ekonomi Islam*, 10(1), 953. <https://doi.org/10.29040/jiei.v10i1.12319>

Sudarmi, Awaluddin, M., Syam, A., & Mahmuddin. (2024). Tren Fashion Syariah dalam Digitalisasi Marketing di Era Milineal dan Gen Z. *Economic Reviews Journal*, 3(3), 2172-2178. <https://doi.org/10.56709/mrj.v3i3.329>

Sugianto, C. A., & Sukmawati, D. (2023). Penerapan Algoritma FP-Growth Untuk Mengetahui Pola Pada Data Transaksi Percetakan (Studi Kasus Java Printing Batuajjar). *Journal of Information Technology*, 5(1), 20-26.

Suhada, S., Ratag, D., Gunawan, G., Wintana, D., & Hidayatulloh, T. (2020). Penerapan algoritma fp-growth untuk menentukan pola pembelian konsumen pada ahass cibadak. *Jurnal Khatulistiwa Informatika*, 8(2), 118-126.

Valencia, S., & Atmojo, W. T. (2024). *Analisis Pola Pembelian pada Data Penjualan CanNgopi menggunakan Algoritma FP-Growth*. 8(2), 214-224.

Winarti, D., Kom, M., Revita, E., & Kom, M. (2021). Penerapan Data Mining untuk Analisa Tingkat Kriminalitas Dengan Algoritma Association Rule Metode FP-Growth. *Jurnal SIMTIKA*, 4(3), 8-22.

