JITEEHA: Journal of Information Technology Application in Education, Economy, Health and Agriculture
Vol. 02, No. 03, October 2025, pp. 80~86, e-ISSN: 3090-3939
https://journal.iteeacademy.org/index.php/jiteeha/

Analysis of Puchase Patterns on Office Stationery Sales Data using Apriori Algorithm

M. Ilham Setyo Wahyudi[1], Firman Nurdiyansyah[2], Dini Kristianti[3]
[1]Department of Informatics Engineering, University of Widyagama Malang, Jl. Borobudur No. 35 Malang, Indonesia
[2]Department of Digital Business, University of Widyagama Malang, Jl. Borobudur No. 35 Malang, Indonesia
[3]Department of Computer Science, Universitas Modern Al Rifa'ie, Jl. Raya Ketawang No. 9 Gondanglegi Malang, Indonesia

Article Info

Article history:

Received August 20, 2025
Revised September 20, 2025
Accepted October 20, 2025

Keywords:

Apriori Algorithm
Association Rules
Frequent Itemsets
Data Mining
Office Stationery.

ABSTRACT

This study analyzes purchasing patterns in office stationery sales using the Apriori algorithm, a data mining method for generating association rules and frequent itemsets. The research examines transaction data to identify combinations of frequently purchased items, aiming to improve inventory management and marketing strategies. The Apriori algorithm calculates metrics such as support, confidence, and lift to determine strong associations between items. Results indicate key purchasing patterns, such as frequent co-purchases of notebooks and pencils, which inform targeted promotions and stock planning. The findings highlight the potential of data-driven decision-making to enhance business efficiency and customer satisfaction in the retail sector.

Corresponding Author:

Firman Nurdiyansyah
Department of Digital Business
Faculty of Science and Information Technology, University of Widyagama Malang
Jl. Borobudur No. 35 Malang, Jawa Timur
Email: firmannurdiyansyah7@gmail.com

1. INTRODUCTION

Office Stationery (ATK) is equipment that is widely used by various groups, such as companies, offices, and students. This equipment has an important role in supporting the smooth running of work and tasks carried out by employees and students. Not only in office and educational environments, ATK is also a very important need for the general public [1][2]. In the retail sector, ATK stores play a role in meeting the needs of customers, both individuals and companies. The volume of items sold in ATK stores often presents challenges in inventory management and the development of effective marketing strategies. Sales transaction data serves as a valuable resource for business decision-making; however, it is often merely archived and utilized for generating sales reports [3]. Without adequate analysis, significant patterns within the transaction data may go unnoticed.

Data mining is the process of extracting important information from a database by identifying existing patterns, which then produces useful insights that are difficult to obtain through manual methods. This process is also known as Knowledge Discovery in Database (KDD) [4]. By utilizing data mining, ATK stores can recognize customer purchasing patterns and trends that emerge in transaction data. This technique allows for more effective data processing, generating insights that can be used to improve business strategies, such as inventory management, product placement, and more targeted promotional offers. One effective method in data mining to analyze transaction data is the Apriori algorithm. This algorithm is used to find relationships between attributes through association rules, often referred to as affinity analysis or market basket analysis. This technique identifies purchasing patterns of items that are often purchased together, for example "If buying a pen, customers tend to also buy notebooks" [5].

Research conducted by Muhammad Aliyul Amri and Syafitri Warda Ningsih (2021) utilized the Apriori algorithm to analyze office stationery sales data. The results showed that items such as Pens (36.36%), Notebooks, and Double Folio Paper (13.64%) were often purchased together, with the association rule that purchases of Notebooks or Double Folio Paper were usually followed by Pens (confidence 100%), which supports stock

management and sales strategies [2]. A study conducted by Kukuh Suryantoro and Shah Khadafi (2023) highlights the effectiveness of the Apriori algorithm in analyzing purchasing patterns within Vertical Stores. The algorithm generates itemset combinations that assist store management in developing marketing strategies, organizing products, and improving sales efficiency [6].

Research by Tumbur Dolok Saribu, Meri Sri Wahyuni, and Syarifah Fadillah Rezky (2024) shows that the Apriori algorithm is effective in identifying stationery sales patterns at CV Ridho Star. This study provides useful information for marketing strategies, helps owners calculate sales accurately, and increases revenue. With this method, owners can find out the combination of products that are often purchased together, so they can plan inventory and promotions better [7]. Research by Tiara Antesia and Sri Eniyati (2020) also succeeded in identifying stationery purchasing patterns at Toko Bagus Adnan Jaya using the Apriori algorithm. This study produces five association rules, with the main rule indicating that if a buyer buys mica_jilid, there is a 91% chance they will also buy lakban_jilid, which can help store owners manage inventory more efficiently and reduce losses [8].

This study aims to apply the Apriori algorithm to analyze purchasing patterns in office stationery sales data. With the Apriori algorithm, it is hoped that stationery stores can gain clearer insights into the combination of items that are often purchased together, so that they can design more effective data-based marketing strategies. In addition, the application of this method can improve inventory management efficiency and provide more appropriate product recommendations to customers, which in turn can improve the shopping experience. The main contribution of this study is the application of the Apriori algorithm to help stationery stores make data-based decisions, which will improve operational performance, accelerate stock management, and optimize sales.

2. METHOD

This research utilizes a data mining approach by implementing the Apriori algorithm to examine sales data for Stationery (ATK). Data mining involves a process designed to uncover valuable insights from datasets, revealing knowledge that would otherwise be inaccessible through manual analysis [9]. The main objectives of data mining include business decision making, outcome prediction, and a deeper understanding of a phenomenon by identifying hidden patterns and relationships in the data. Association rule mining is a data mining method used to identify relationships between item combinations. This approach also serves as the foundation for various other data mining techniques. Notably, one of the key aspects of association analysis that has garnered significant interest from researchers is the development of efficient algorithms for analyzing high-frequency patterns [10].

In the context of Stationery (ATK) sales, the application of the Apriori algorithm and association analysis play an important role in identifying significant purchasing patterns and revealing relationships between goods and consumer preferences. Through sales data analysis, insights are obtained that can be used for strategic decision making in managing the ATK sales business. The stages of the research carried out are as follows:

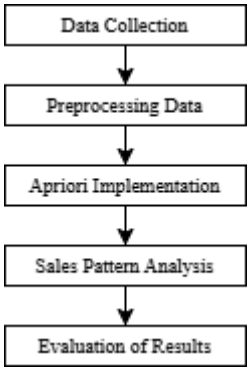

Figure 1. Research stages

In the data collection stage, transaction data is collected from various sources, such as sales recording systems or store databases. The data collected includes information such as transaction dates, lists of items purchased, and quantity of items. Complete and accurate data is very important because it will be the foundation for the next analysis process. After the data is collected, the collected data is then prepared to be suitable for analysis. Transaction data is converted into a form that can be used by the Apriori algorithm, namely a collection of items purchased in each transaction. This stage ensures that the data is in optimal condition for analysis. Apriori algorithm is applied to find frequent itemsets (combinations of items that are often purchased together) with two main metrics: Support and Confidence. Support measures how often a combination of items appears in transaction data, calculated by the formula [11][12]:

$$Support\ (A) = \frac{\sum \quad Jumlah\ transaksi\ mengandung\ A}{Total\ transaksi} * 100\% \quad (1)$$

Confidence in association rules represents a measure of a rule's reliability, indicating the percentage of transactions in the database that include both item A and item B. By utilizing confidence, the strength of the relationship between items within an association rule can be assessed. The formula to calculate the confidence value between two items is as follows [11][12]:

$$Confidence\ (A \rightarrow B) = \frac{Support\ (AB)}{Support\ (A)} * 100\% \quad (2)$$

Lift measures whether two itemsets appear together more often than expected, calculated by the formula [13][14]:

$$Lift\ (A \rightarrow B) = \frac{Confidence\ (AB)}{Support\ (B)} * 100\% \quad (3)$$

Frequent itemsets and association rules generated by the Apriori algorithm are analyzed in more depth to understand purchasing patterns [15]. Combinations of frequently purchased items are evaluated using various metrics, such as lift ratio, to identify significant relationships between items. The results of this analysis can be used to develop more effective business strategies, such as offering promotional packages or better arranging the layout of goods in stores. After the analysis, an evaluation is conducted to test the quality of the association rules and measure their effectiveness, such as sales conversion or increased profits. This evaluation is important to ensure that the research results can be practically applied in business management.

3. RESULTS AND DISCUSSION

Accumulation of sales transaction data in office stationery stores is obtained through weekly sales records. The data is taken from six weekly reports with the highest number of transactions. This information is presented in tabular form to facilitate analysis and identification of purchasing patterns.

Table 1. Sales transaction data

No	Transaction Date	Items Purchased
1	01/09/2024	Pen, Pencil, Eraser, Notebook, Ruler, Marker
2	08/09/2024	Pencil, Sharpener, Pen, Notebook, Highlighter, Marker
3	15/09/2024	Notebook, Pen, Eraser, Pencil, Ruler, Highlighter
4	22/09/2024	Pencil, Sharpener, Eraser, Marker, Highlighter, Notebook
5	29/09/2024	Notebook, Pen, Highlighter, Marker, Pencil, Sharpener
6	06/10/2024	Notebook, Pen, Ruler, Highlighter, Marker, Sharpener
7	13/10/2024	Notebook, Highlighter, Marker, Eraser, Pencil, Sharpener
8	20/10/2024	Pen, Notebook, Ruler, Pencil, Eraser, Highlighter
9	27/10/2024	Notebook, Pen, Highlighter, Marker, Pencil, Sharpener
10	03/11/2024	Notebook, Highlighter, Marker, Eraser, Pencil, Sharpener
11	10/11/2024	Pen, Notebook, Ruler, Marker, Highlighter, Sharpener
12	17/11/2024	Notebook, Pen, Highlighter, Marker, Pencil, Eraser

The tabular format of weekly sales transaction data can be presented in the form of a table that displays information in a structured and neat manner. Presenting data in tabular format facilitates the process of analyzing purchasing patterns, identifying sales frequencies, and managing data more effectively. The following is a table of tabular format transaction data:

Table 2. Tabular format

Week	Pen	Pencil	Eraser	Notebook	Ruler	Marker	Sharpener	Highlighter
1	1	1	1	1	1	1	0	0
2	1	1	1	1	0	1	1	0
3	1	1	1	1	1	0	0	1
4	0	1	1	1	0	1	1	1
5	1	1	0	1	0	1	1	1

6	1	0	1	1	1	1	0	1
7	0	1	1	1	0	1	1	1
8	1	1	1	1	1	0	0	1
9	1	1	0	1	0	1	1	1
10	0	1	1	1	0	1	1	1
11	1	0	1	1	1	1	0	1
12	1	1	0	1	0	1	1	1
Total	9	10	9	12	5	10	7	10

3.1. 1 Itemset (C_1)

Formation of 1 itemset or C_1 with a minimum number of supports = 30%. The following is a calculation table for the formation of 1 itemset:

Table 3. Support 1 itemset

Itemset	Frequency	Support (%)
Pen	9	$\frac{9}{12}.100\% = 75\%$
Pencil	10	$\frac{10}{12}.100\% = 83{,}33\%$
Eraser	9	$\frac{9}{12}.100\% = 75\%$
Notebook	12	$\frac{12}{12}.100\% = 100\%$
Ruler	5	$\frac{5}{12}.100\% = 41{,}67\%$
Marker	10	$\frac{10}{12}.100\% = 83{,}33\%$
Sharpener	7	$\frac{7}{12}.100\% = 83{,}33\%$
Highlighter	10	$\frac{10}{12}.100\% = 83{,}33\%$

3.2. 2 Itemset (C_2)

The next stage is the formation of itemset 2 or C2. At this stage, two items from itemset 1 (C1) are combined that have met the minimum support criteria of 30%. Each pair of items formed is then calculated for the frequency of occurrence in the transaction data. The following table shows the results of the calculation of itemset 2 (C2):

Table 4. Support 2 itemset

Itemset	Frequency	Support (%)
{Pen, Pencil}	7	$\frac{7}{12}.100\% = 53{,}33\%$
{Pen, Eraser}	6	$\frac{6}{12}.100\% = 50\%$
{Pen, Notebook}	9	$\frac{9}{12}.100\% = 75\%$
{Pen, Ruler}	3	$\frac{3}{12}.100\% = 25\%$
{Pen, Marker}	5	$\frac{5}{12}.100\% = 41{,}67\%$
{Pen, Sharpener}	4	$\frac{4}{12}.100\% = 33{,}33\%$
{Pen, Highlighter}	6	$\frac{6}{12}.100\% = 50\%$
{Pencil, Eraser}	8	$\frac{8}{12}.100\% = 66{,}67\%$
{Pencil, Notebook}	10	$\frac{10}{12}.100\% = 83.33\%$
{Pencil, Ruler}	4	$\frac{4}{12}.100\% = 33{,}33\%$
{Pencil, Marker}	6	$\frac{6}{12}.100\% = 50\%$
{Pencil, Sharpener}	5	$\frac{5}{12}.100\% = 41{,}67\%$

{Pencil, Highlighter}	7	$\frac{7}{12}.100\% = 58{,}33\%$
{Eraser, Notebook}	8	$\frac{8}{12}.100\% = 66{,}67\%$
{Notebook, Marker}	9	$\frac{9}{12}.100\% = 75\%$
{Notebook, Highlighter}	7	$\frac{7}{12}.100\% = 58{,}33\%$
{Marker, Highlighter}	8	$\frac{8}{12}.100\% = 66{,}67\%$
{Marker, Sharpener}	5	$\frac{5}{12}.100\% = 41{,}67\%$
{Sharpener, Highlighter}	4	$\frac{4}{12}.100\% = 33{,}33\%$
{Ruler, Notebook}	3	$\frac{3}{12}.100\% = 25\%$
{Ruler, Marker}	3	$\frac{3}{12}.100\% = 25\%$
{Ruler, Sharpener}	3	$\frac{3}{12}.100\% = 25\%$

3.3. Confidence

The subsequent step involves calculating the confidence with a minimum threshold of 70%. This step is intended to assess the strength of the relationships between items in the association rules. A higher confidence value indicates a stronger relationship. Association rules with confidence values meeting or exceeding the minimum threshold are deemed valid. Below is the calculation table for minimum confidence:

Table 5. Confidence value

Itemset	Frequency Itemset	Frequency A	Frequency B	Confidence (A → B)	Confidence (B → A)	Status
{Pen, Pencil}	7	9	10	77,78%	70%	Pass
{Pen, Eraser}	6	9	9	66,67%	66,67%	Fail
{Pen, Notebook}	9	9	12	100%	75%	Pass
{Pen, Ruler}	3	9	5	33,33%	25%	Fail
{Pen, Marker}	5	9	10	55,56%	50%	Fail
{Pen, Sharpener}	4	9	7	57,14%	57,14%	Fail
{Pen, Highlighter}	6	9	10	66,67%	60%	Fail
{Pencil, Eraser}	8	10	9	87,50%	80%	Pass
{Pencil, Notebook}	10	10	12	100%	90%	Pass
{Pencil, Ruler}	4	10	5	50%	40%	Fail
{Pencil, Marker}	6	10	10	75%	60%	Fail
{Pencil, Sharpener}	5	10	7	62,50%	62,50%	Fail
{Pencil, Highlighter}	7	10	10	87,50%	70%	Pass
{Eraser, Notebook}	8	9	12	100%	75%	Pass
{Notebook, Marker}	9	12	10	100%	75%	Pass
{Notebook, Highlighter}	7	12	10	100%	70%	Pass
{Marker, Highlighter}	8	10	10	100%	80%	Pass
{Marker, Sharpener}	5	10	7	62,50%	62,50%	Fail
{Sharpener, Highlighter}	4	7	10	57,14%	57,14%	Fail
{Ruler, Notebook}	3	5	12	33,33%	25%	Fail
{Ruler, Marker}	3	5	10	33,33%	25%	Fail
{Ruler, Sharpener}	3	5	7	33,33%	25%	Fail

The table and diagram below present the 2-itemset or C2 combinations that satisfy the minimum support threshold of 30% and the minimum confidence level of 70%. This step ensures that only itemset combinations

with strong and relevant associations are retained, enabling further analysis to uncover significant purchasing patterns.

Table 6. Confidence that passed

Itemset	Frequency Itemset	Support (%)	Confidence (A → B)	Confidence (B → A)	Lift
{Pen, Pencil}	7	53.33%	77.78%	70%	1.04
{Pen, Notebook}	9	75%	100%	75%	1.33
{Pencil, Eraser}	8	66.67%	87.50%	80%	1.31
{Pencil, Notebook}	10	83.33%	100%	90%	1.20
{Pencil, Stabilo}	7	58.33%	87.50%	70%	1.50
{Penghapus, Buku Tulis}	8	66.67%	100%	75%	1.50
{Buku Tulis, Spidol}	9	75%	75%	90%	1.00
{Buku Tulis, Stabilo}	7	58.33%	100%	70%	1.71
{Spidol, Stabilo}	8	66.67%	80%	80%	1.33

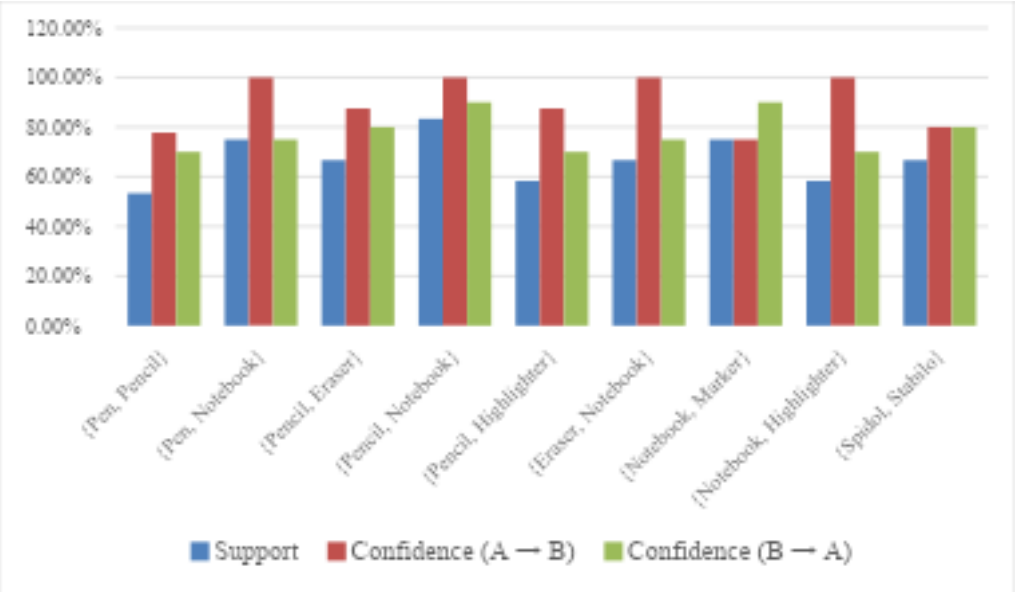

Figure 2. Association Result Diagram

4. CONCLUSION

The results of the study identified the most frequently purchased items, with notebooks having the highest purchase rate of 100%, followed by pencils, markers, sharpeners, and highlighters at 83.33%, erasers at 75%, and rulers at 41.67%. These findings are important for stationery store owners in planning marketing strategies and maintaining stock inventory according to market demand. By utilizing data mining and the Apriori Algorithm, business actors can increase the chances of success and business growth in the retail industry. This study also emphasizes the importance of understanding consumer preferences and market interest in stock management. Adjusting stock according to demand can increase customer satisfaction and optimize profits. Therefore, the results of this study are expected to be a guide for business actors in planning marketing strategies, improving service quality, and encouraging sustainable business growth.

ACKNOWLEDGEMENTS

The author wishes to convey heartfelt appreciation to Widyagama University Malang for offering the facilities and support essential for the successful completion of this research. Deep gratitude is extended to supervisors and colleagues for their insightful guidance and contributions throughout the study. The author is also thankful for the cooperation of the office stationery store (ATK) in facilitating data collection and engaging in meaningful discussions. Lastly, sincere thanks go to family and friends for their unwavering encouragement and support.

REFERENCES

[1] A. D. Purba, W. Saputra, and D. K. Lumbantobing, "Penerapan Data Mining Dalam Menentukan Pola Penjualan Alat Tulis Kantor (Atk) Menggunakan Algoritma Apriori Pada Toko Novacom," *Snastikom*, vol. 2, no. 1, pp. 120–131, 2023.

[2] M. Aliyul Amri and S. Warda Ningsih, "Penerapan Data Mining Penjualan Alat Tulis Kantor Menggunakan Algoritma Apriori Di Tiga Balata," vol. 2, no. 3, pp. 138–146, 2021.

[3] I. Djamaludin and A. Nursikuwagus, "Analisis Pola Pembelian Konsumen Pada Transaksi Penjualan Menggunakan Algoritma Apriori," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 8, no. 2, p. 671, 2017, doi: 10.24176/simet.v8i2.1566.

[4] A. A. Dewayanti, "Penerapan Data Mining Pada Data Nilai Siswa Dengan Menggunakan Algoritma Asosiation Rule Metode Apriori (Studi Kasus di SMP N 36 Semarang)," *Pros. Semin. Nas. Mat. Dan Pendidik. Mat.*, pp. 1–10, 2018.

[5] P. M. S. Tarigan, J. T. Hardinata, H. Qurniawan, M. Safii, and R. Winanjaya, "Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang," *J. Janitra Inform. dan Sist. Inf.*, vol. 2, no. 1, pp. 9–19, 2022, doi: 10.25008/janitra.v2i1.142.

[6] K. Suryantoro and S. Khadafi, "SNESTIK Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika Penerapan Data Mining dengan Metode Algoritma Apriori untuk Menentukan Pola Pembelian Alat Tulis Kantor (Studi Kasus Pada Toko Vertikal)," *SNESTIK Semin. Nas. Tek. Elektro, Sist. Informasi, dan Tek. Inform.*, pp. 326–332, 2023, [Online]. Available: https://ejurnal.itats.ac.id/snestikdanhttps://snestik.itats.ac.id

[7] T. D. Saribu, M. S. Wahyuni, and S. F. Rezky, "Data Mining Dalam Menentukan Pola Penjualan Barang Menggunakan Algoritma Apriori," vol. 3, pp. 183–190, 2024.

[8] T. Antesia[1] and S. Eniyati[2], "Aturan Asosiasi Data Alat Tulis Kantor Menggunakan Algoritma Apriori," *Proceeding SENDIU*, pp. 298–299, 2020.

[9] K. Annisa, B. S. Ginting, and M. A. Syar, "Penerapan Data Mining Pengelompokan Data Pengguna Air Bersih Berdasarkan Keluhannya Menggunakan Metode Clustering Pada Pdam Langkat," *J. Sist. Inf. Kaputama*, vol. 6, no. 2, pp. 165–179, 2022, doi: 10.59697/jsik.v6i2.167.

[10] C. Zhang, "Negative Association Rule," pp. 47–84, 2002, doi: 10.1007/3-540-46027-6_3.

[11] D. Listriani and A. H. Setyaningrum, "Penerapan Metode Asosiasi Menggunakan Algoritma Apriori Pada Aplikasi Pola Belanja Konsumen (Studi Kasus Toko Buku Gramedia Bintaro)," *Int. J. Sci. Eng. Res. (IJ0SER),* vol. 3, no. 4, p. 2, 2015, [Online]. Available: http://journal.uinjkt.ac.id/index.php/ti/article/view/5602/3619

[12] D. Kusumo, Dana Sulistiyo, Bijaksana, Moch. Arief, Darmantoro, "Data Mining Dengan Algoritma Apriori," *Penelit. dan Pengemb. Telekomun.*, pp. 1–5, 2003.

[13] N. F. FAHRUDIN, "Penerapan Algoritma Apriori untuk Market Basket Analysis," *MIND J.*, vol. 1, no. 2, pp. 13–23, 2019, doi: 10.26760/mindjournal.v4i1.13-23.

[14] M. Fauzy, K. R. Saleh W, and I. Asror, "Penerapan Metode Association Rule Menggunakan," *J. Ilm. Teknol. Inf. Terap.*, vol. II, no. 2, pp. 221–227, 2016.

[15] D. Listriani, A. H. Setyaningrum, and F. Eka, "Pengertian Asosiasi Dan Clusterisasi Serta Penerapan Di Toko," *J. Tek. Inform. Vol 9 No. 2, Oktober 2016*, vol. 9, no. 2, pp. 120–127, 2016.