



Application of data mining on inventory grouping using clustering method

Suci Ramadani¹, Syukri Hidayat², Ramahdanty^{3*}

¹)Program Studi Komputerisasi Akuntansi, STMIK Kaputama, Binjai, Indonesia

²)Program Studi Pendidikan Geografi, Fakultas Ilmu Sosial, Universitas Negeri Medan, Medan, Indonesia

³)Program Studi Sistem Informasi, STMIK Kaputama, Binjai, Indonesia

Article Info

Article history

Received : Oct 10, 2023

Revised : Nov 15, 2023

Accepted : Nov 20, 2023

Keywords:

Clustering Method;

Data Mining;

Goods Inventory Grouping.

Abstract

Data mining in the business field is considered important, because the inventory system of goods in a store and what types of goods are the top priorities that must be stocked to anticipate the vacancy of goods, so that the store owner can find out the most sold items and the lack of stock items. The existence of daily sales transaction activities at Sahabat Komputer stores will produce a pile of data that is getting bigger and bigger, so that it can cause new problems. If this is allowed, the transaction data will become a pile of data that is detrimental because it requires an increasingly large storage media or database. One way to overcome this is to keep the availability of various types of continuous goods in the warehouse. To find out what items are purchased by consumers, the technique of analyzing the inventory of goods in the warehouse is carried out. Application of Clustering, helps in grouping data of the same characteristics into the same region. And from the whole it can be concluded that in cluster 1 the stock is available on average 1-100 pcs, the number of sales is 1-100 pcs and the sales volume per month is 1-100 units. In cluster 2 there is an average available stock of 101-200 pcs, 101-200 pcs sales quantity of 101-200 units, and monthly sales volume of 101-200 pcs. And in cluster 3 there is an average available stock of 301-400 pcs, the number sold is 401-500 pcs, and the monthly sales volume reaches 301-400 pcs.

Corresponding Author:

Ramahdanty,
Program Studi Sistem Informasi,
STMIK Kaputama,
Jl. Veteran No.4A, Tangsi, Kec. Binjai Kota, Kota Binjai, Sumatera Utara 20714, Indonesia.
Email : suci.ramadani23@gmail.com

This is an open access article under the [CC BY-NC](https://creativecommons.org/licenses/by-nc/4.0/) license.



1. Introduction

Inventory management is a critical aspect of an organization's operations, especially in the context of industry and commerce [1][2]. Efficiently managed inventory can have a positive impact on company performance, including reduced storage costs, improved customer service, and supply chain optimization [3][4]. In the era of information and technology, data mining approaches are becoming increasingly relevant to support decision-making in inventory management [5][6]. One method that can be applied is the clustering method, which allows the grouping of goods based on the similarity of certain characteristics [7][8].

This research aims to apply data mining techniques, particularly clustering methods, to inventory analysis to identify patterns or groups that may not be directly visible [9][10][11][12]. By doing this clustering, it is hoped that useful information can be generated in making decisions related to inventory management. The clustering method was chosen due to its ability to group similar objects into clusters based on similar characteristics, thus facilitating a better understanding of the structure and dynamics of inventory [13][14][15]. This research includes not only the implementation of clustering techniques, but also the evaluation of clustering results, interpretation of emerging patterns, and implementation of inventory management strategies based on the identified clusters. Thus, it is expected that this research can make a positive contribution to the company's operational efficiency and the development of smarter inventory management strategies [14][16][17].

The context of this research is highly relevant given the challenges faced by organizations in managing increasingly complex and dynamic inventory. Intense business competition, fluctuations in market demand, and changing consumer trends require innovative approaches in inventory management [18][19]. In this context, the application of data mining techniques, particularly clustering methods, can be an effective tool to identify hidden patterns in inventory data, which may be difficult or even impossible to discover through manual analysis. The importance of this research also lies in its potential to provide deep insights into the characteristics of inventory, which can form the basis for strategic planning and more informed decision-making [20][21]. By exploring the inventory clusters formed, organizations can optimize resource allocation, reduce the risk of excess or shortage of inventory, and improve the accuracy of meeting customer needs [18][22].

Through this research, it is expected that new insights will emerge regarding the potential application of clustering methods in the context of inventory management. The results of this study can also contribute to the scientific literature by enriching the understanding of the application of data mining techniques in industry and commerce, especially in terms of inventory management. In addition, this research can serve as a basis for further development in the use of technology and data analysis to improve operational efficiency in various sectors. The conclusions and recommendations generated from this research can serve as practical guidelines for inventory managers and related professionals in making better, data-driven decisions.

By summarizing the research focus on the application of data mining with clustering methods in grouping inventory, it is expected that this research can make a significant contribution to the understanding and practice of inventory management in the face of changing market dynamics. The existence of daily sales transaction activities at Sahabat Komputer stores will produce a pile of data that is getting bigger and bigger, so that it can cause new problems. One of the problems that often arises is not knowing the exact number and condition of goods available in the warehouse. This can result in goods being stored in the warehouse for too long, excessive, lacking, or even running out of goods if the goods consist of various types and with a high turnover rate. If this is left unchecked, the transaction data will become a pile of waste that is detrimental because it requires an increasingly large storage media or database [23][24].

One way to overcome this is to keep the availability of various types of continuous goods (continuous) in the warehouse. To find out what items are purchased by consumers, the technique of analyzing the inventory of goods in the warehouse is carried out. The application of Clustering, helps in grouping data of the same characteristics into the same area, the aim is to minimize the objective function which generally tries to minimize the variation in a cluster [25][26][7].

According to Buaton & Relita, 2015 Data mining is a series of processes, data mining can be divided into several stages [27]. These stages are interactive, the user is directly involved or by mediating the knowledge base. Clustering or clustering is a method of grouping data. According to Ikuton (2023) clustering is a process for grouping data into several clusters or groups so that data in one cluster has the maximum level of similarity and data between clusters has minimum similarity [7]. Clustering is the process of partitioning a set of data objects into a set of parts called clusters. Objects in the cluster have similar characteristics between each other and are different from other clusters.

Partitioning is not done manually but with a clustering algorithm. Therefore, clustering is very useful and can find unknown groups in the data.

2. Research Methodolgy

The stages of the research methodology carried out in solving the problem are as follows :

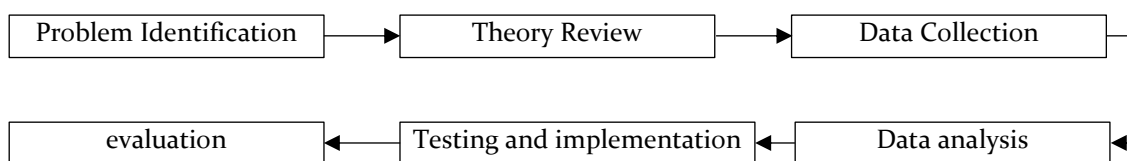


Figure 1: Research Workflow

a. Problem Identification

At this stage, identifying what are the problems that exist in the object of research, namely Toko Sahabat Komputer, which when identifying problems in the store, there are several problems that can be solved using data mining techniques, these problems are in the form of stock inventory that cannot be processed effectively and efficiently by the shop owner.

b. Theory Study

At this stage, searching for the theoretical basis obtained from various sources such as books, scientific journals, and also other references related to the research title in order to complement the research both conceptually and theoretically so that it has a good and relevant reference.

c. Data collection

The observation carried out is by making direct observations of the process of activities related to the Data Mining System in the inventory of goods at Sahabat Komputer. This method is used to obtain primary data. Primary data itself is in the form of information obtained from direct interviews with Sahabat Komputer Store Owners. In this study collecting data.

d. Data analysis

This stage is the stage of managing and analyzing the data that has been obtained so that the data can be grouped according to the predetermined variables. At this stage, the data processing that will be processed is taken from the data on the number of sales transactions from December 2018 - December 2019, this data processing is carried out based on predetermined stages such as data cleaning, data merging, data transformation, k-means clustering calculation, evaluation and analysis of results.

e. Testing and implementation

This stage is a stage that tests the validation and implementation of previously analyzed data and program preparation. This stage consists of analysis, coding, and testing.

f. Evaluation

This stage is the stage of taking conclusions and suggestions that can be done in this research. With the conclusion, the results of the whole will be known and it is hoped that with suggestions there will be improvements and benefits for others.

3. Results and Discussion

a. Data Initialization

From the existing data, data initialization can be done according to the variable needs as follows :

Table 1.
Inventory of Goods

No	Product Type	Brand	Quantity Available
1	Fanproc	Intel	300
2	Fanproc	AMD	130
3	Motherboard	Intel	550
4	Motherboard	AMD	500
5	HDD PC	Seagate	410
6	HDD PC	WD	356
7	HDD PC	Toshiba	312
8	Flashdisk	Sandisk	620
9	Flashdisk	Toshiba	450
10	Flashdisk	HP	450
11	Cartridge	Canon	925
12	Cartridge	Brother	600
13	Cartridge	Epson	511
14	Cartridge	HP	278
15	Tinta	Epson	850
16	Tinta	Canon	365
17	Kertas A4	Natural	400
18	Kertas A4	PaperOne	50
19	Toner	Aiflo	45
20	Toner	Miracle	45

b. Data Transformation

Tabel 2.
Transformasi Data dalam Algoritma K-Means Cluster

Product Type	Brand	Quantity Available
1	1	3
1	2	1
2	1	2
2	2	2
3	3	4
3	4	4
3	5	3
4	6	6
4	5	5
4	7	5
5	8	6
5	9	6
5	10	5
5	7	3
6	10	6
6	8	2
7	11	3
7	12	1
8	13	1
8	14	1

$$\text{Euclidean}(X, Y) = \sqrt{\sum_i (x_i - c_i)^2} \quad (1)$$

Iterations I

Centroid 1 = (1, 2, 1) taken from the data to -2

Centroid 2 = (5, 8, 6) taken from the data to -11

Centroid 3 = (7, 11, 3) taken from the data to -17

Description: The initial Centroid value is taken randomly from the transformed data.

Section Code "1" (1, 1, 3)

$$\text{Distance from } C_1(X) = \sqrt{(1-1)^2 + (1-2)^2 + (3-1)^2} = 2.2$$

$$\text{Distance from } C_2(Y) = \sqrt{(1-5)^2 + (1-8)^2 + (3-6)^2} = 8.6$$

$$\text{Distance from } C_3(Z) = \sqrt{(1-7)^2 + (1-11)^2 + (3-3)^2} = 11.6$$

Section Code "2" (1, 2, 1)

$$\text{Distance from } C_1(X) = \sqrt{(1-1)^2 + (2-2)^2 + (1-1)^2} = 0$$

$$\text{Distance from } C_2(Y) = \sqrt{(1-5)^2 + (2-8)^2 + (1-6)^2} = 8.7$$

$$\text{Distance from } C_3(Z) = \sqrt{(1-7)^2 + (2-11)^2 + (1-3)^2} = 11$$

Section Code "3" (2, 1, 2)

$$\text{Distance from } C_1(X) = \sqrt{(2-1)^2 + (1-2)^2 + (2-1)^2} = 1.7$$

$$\text{Distance from } C_2(Y) = \sqrt{(2-5)^2 + (1-8)^2 + (2-6)^2} = 8.6$$

$$\text{Distance from } C_3(Z) = \sqrt{(2-7)^2 + (1-11)^2 + (2-3)^2} = 11.2$$

Section Code "4" (2, 2, 2)

$$\text{Distance from } C_1(X) = \sqrt{(2-1)^2 + (2-2)^2 + (2-1)^2} = 1.4$$

$$\text{Distance from } C_2(Y) = \sqrt{(2-5)^2 + (2-8)^2 + (2-6)^2} = 7.8$$

$$\text{Distance from } C_3(Z) = \sqrt{(2-7)^2 + (2-11)^2 + (2-3)^2} = 10.3$$

Section Code "5" (3, 3, 4)

$$\text{Distance from } C_1(X) = \sqrt{(3-1)^2 + (3-2)^2 + (4-1)^2} = 3.7$$

$$\text{Distance from } C_2(Y) = \sqrt{(3-5)^2 + (3-8)^2 + (4-6)^2} = 5.7$$

$$\text{Distance from } C_3(Z) = \sqrt{(3-7)^2 + (3-11)^2 + (4-3)^2} = 9$$

Section Code "6" (3, 4, 4)

$$\text{Distance from } C_1(X) = \sqrt{(3-1)^2 + (4-2)^2 + (4-1)^2} = 4.1$$

$$\text{Distance from } C_2(Y) = \sqrt{(3-5)^2 + (4-8)^2 + (4-6)^2} = 4.8$$

$$\text{Distance from } C_3(Z) = \sqrt{(3-7)^2 + (4-11)^2 + (4-3)^2} = 8.1$$

Section Code "7" (3, 5, 3)

$$\text{Distance from } C_1(X) = \sqrt{(3-1)^2 + (5-2)^2 + (3-1)^2} = 4.1$$

$$\text{Distance from } C_2(Y) = \sqrt{(3-5)^2 + (5-8)^2 + (3-6)^2} = 4.6$$

$$\text{Distance from } C_3(Z) = \sqrt{(3-7)^2 + (5-11)^2 + (3-3)^2} = 7.2$$

Section Code "8" (4, 6, 6)

$$\text{Distance from } C_1(X) = \sqrt{(4-1)^2 + (6-2)^2 + (6-1)^2} = 7.1$$

$$\text{Distance from } C_2(Y) = \sqrt{(4-5)^2 + (6-8)^2 + (6-6)^2} = 2.2$$

$$\text{Distance from } C_3(Z) = \sqrt{(4-7)^2 + (6-11)^2 + (6-3)^2} = 6.5$$

Section Code "9" (4, 5, 5)

$$\text{Distance from } C_1(X) = \sqrt{(4-1)^2 + (5-2)^2 + (5-1)^2} = 5.8$$

$$\text{Distance from } C_2(Y) = \sqrt{(4-5)^2 + (5-8)^2 + (5-6)^2} = 3.3$$

$$\text{Distance from } C_3(Z) = \sqrt{(4-7)^2 + (5-11)^2 + (5-3)^2} = 7$$

Section Code "10" (4, 7, 5)

$$\text{Distance from } C_1(X) = \sqrt{(4-1)^2 + (7-2)^2 + (5-1)^2} = 7.1$$

$$\text{Distance from } C_2(Y) = \sqrt{(4-5)^2 + (7-8)^2 + (5-6)^2} = 1.7$$

$$\text{Distance from } C_3(Z) = \sqrt{(4-7)^2 + (7-11)^2 + (5-3)^2} = 5.3$$

Section Code "11" (5, 8, 6)

$$\text{Distance from } C_1(X) = \sqrt{(5-1)^2 + (8-2)^2 + (6-1)^2} = 8.7$$

$$\text{Distance from } C_2(Y) = \sqrt{(5-5)^2 + (8-8)^2 + (6-6)^2} = 0$$

$$\text{Distance from } C_3(Z) = \sqrt{(5-7)^2 + (8-11)^2 + (6-3)^2} = 4.6$$

Section Code "12" (5, 9, 6)

$$\text{Distance from } C_1(X) = \sqrt{(5-1)^2 + (9-2)^2 + (6-1)^2} = 9.9$$

$$\text{Distance from } C_2(Y) = \sqrt{(5-5)^2 + (9-8)^2 + (6-6)^2} = 1$$

$$\text{Distance from } C_3(Z) = \sqrt{(5-7)^2 + (9-11)^2 + (6-3)^2} = 4.1$$

Section Code "13" (5, 10, 5)

$$\text{Distance from } C_1(X) = \sqrt{(5-1)^2 + (10-2)^2 + (5-1)^2} = 9.7$$

$$\text{Distance from } C_2(Y) = \sqrt{(5-5)^2 + (10-8)^2 + (5-6)^2} = 2.2$$

$$\text{Distance from } C_3(Z) = \sqrt{(5-7)^2 + (10-11)^2 + (5-3)^2} = 3$$

Section Code "14" (5, 7, 3)

$$\text{Distance from } C_1(X) = \sqrt{(5-1)^2 + (7-2)^2 + (3-1)^2} = 6.7$$

$$\text{Distance from } C_2(Y) = \sqrt{(5-5)^2 + (7-8)^2 + (3-6)^2} = 3.1$$

$$\text{Distance from } C_3(Z) = \sqrt{(5-7)^2 + (7-11)^2 + (3-3)^2} = 4.4$$

Section Code "15" (6, 10, 6)

$$\text{Distance from } C_1(X) = \sqrt{(6-1)^2 + (10-2)^2 + (6-1)^2} = 10.6$$

$$\text{Distance from } C_2(Y) = \sqrt{(6-5)^2 + (10-8)^2 + (6-6)^2} = 2.2$$

$$\text{Distance from } C_3(Z) = \sqrt{(6-7)^2 + (10-11)^2 + (6-3)^2} = 3.3$$

Section Code "16" (6, 8, 2)

$$\text{Distance from } C_1(X) = \sqrt{(6-1)^2 + (8-2)^2 + (2-1)^2} = 7.8$$

$$\text{Distance from } C_2(Y) = \sqrt{(6-5)^2 + (8-8)^2 + (2-6)^2} = 4.1$$

$$\text{Distance from } C_3(Z) = \sqrt{(6-7)^2 + (8-11)^2 + (2-3)^2} = 3.3$$

Section Code "17" (7, 11, 3)

$$\text{Distance from } C_1(X) = \sqrt{(7-1)^2 + (11-2)^2 + (3-1)^2} = 11$$

$$\text{Distance from } C_2(Y) = \sqrt{(7-5)^2 + (11-8)^2 + (3-6)^2} = 4.6$$

$$\text{Distance from } C_3(Z) = \sqrt{(7-7)^2 + (11-11)^2 + (3-3)^2} = 0$$

Section Code "18" (7, 12, 1)

$$\text{Distance from } C_1(X) = \sqrt{(7-1)^2 + (12-2)^2 + (1-1)^2} = 11.6$$

$$\text{Distance from } C_2(Y) = \sqrt{(7-5)^2 + (12-8)^2 + (1-6)^2} = 6.7$$

$$\text{Distance from } C_3(Z) = \sqrt{(7-7)^2 + (12-11)^2 + (1-3)^2} = 2.3$$

Section Code "19" (8, 13, 1)

$$\text{Distance from } C_1(X) = \sqrt{(8-1)^2 + (13-2)^2 + (1-1)^2} = 13$$

$$\text{Distance from } C_2(Y) = \sqrt{(8-5)^2 + (13-8)^2 + (1-6)^2} = 7.6$$

$$\text{Distance from } C_3(Z) = \sqrt{(8-7)^2 + (13-11)^2 + (1-3)^2} = 3$$

Section Code "20" (8, 14, 1)

$$\text{Distance from } C_1(X) = \sqrt{(8-1)^2 + (14-2)^2 + (1-1)^2} = 13.8$$

$$\text{Distance from } C_2(Y) = \sqrt{(8-5)^2 + (14-8)^2 + (1-6)^2} = 8.3$$

$$\text{Distance from } C_3(Z) = \sqrt{(8 - 7)^2 + (14 - 11)^2 + (1 - 3)^2} = 3.7$$

Table 3.
Iteration I Calculation Results

Centroid Distance			Closest Distance	Followed Cluster
1	2	3		
2.2	8.6	11.6	2.2	1
0	8.7	11	0	1
1.7	8.6	11.2	1.7	1
1.4	7.8	10.3	1.4	1
3.7	5.7	9	3.7	1
4.1	4.8	8.1	4.1	1
4.1	4.6	7.2	4.1	1
7.1	2.2	6.5	2.2	2
5.8	3.3	7	3.3	2
7.1	1.7	5.3	1.7	2
8.7	0	4.6	0	2
9.9	1	4.1	1	2
9.7	2.2	3	2.2	2
6.7	3.1	4.4	3.1	2
10.6	2.2	3.3	2.2	2
7.8	4.1	3.3	3.3	3
11	4.6	0	0	3
11.6	6.7	2.2	2.2	3
13	7.6	3	3	3
13.8	8.3	3.7	3.7	3

Table 4.
Iteration II Calculation Results

Jarak Centroid			Jarak Terdekat	Cluster Yang Diikuti
1	2	3		
2.5	8.1	12.3	2.5	1
2.5	8.3	11.4	2.5	1
2.2	8.1	11.8	2.2	1
1.4	7.3	10.9	1.4	1
1.3	5.3	9.8	1.3	1
1.6	4.4	9	1.6	1
1.9	4.2	7.9	1.9	1
4.6	1.9	7.8	1.9	2
3.3	2.9	8	2.9	2
4.7	1.2	6.5	1.2	2
6.3	0.4	6	0.4	2
8.3	1.2	5.5	1.2	2
7.6	2.3	4.3	2.3	2
4.6	2.8	5.2	2.8	2
8.3	2.5	4.8	2.5	2
6	3.9	3.8	3.8	3
9	4.7	1.5	1.5	3
10.1	6.7	0.7	0.7	3
11.4	7.7	1.7	1.7	3
12.3	8.4	2.6	2.6	3

Group based on the minimum distance to the nearest centroid i.e.:

If the shortest distance is at C_1 then the data is included in group 1

If the shortest distance is at C_2 then the data is included in group 2

If the shortest distance is at C_3 then the data is included in group 3

Older Group = (1,1,1,1,1,1,2,2,2,2,2,2,1,2,3,3,3,3)

New Groups = (1,1,1,1,1,1,2,2,2,2,2,2,1,2,3,3,3,3)

From the calculation results of iteration 2, the same group is obtained as the calculation results in iteration 1. Because there is no group change between iteration 2 and iteration 3, the calculation is stopped.

From 20 data obtained 3 groups, it can be concluded that:

1. Cluster 1

2.5 3.1 2.75

It can be seen that in cluster 1 there are 8 data with the type of AMD motherboard product with the number available 101-200pcs..

2. Cluster 2

4.7 7.8 5.7

It can be seen that in cluster 2 there are 7 types of HP flashdisk products with the number available 401-500pcs..

3. Cluster 3

7.2 11.6 1.6

It can be seen that in cluster 3 there are 5 data with the type of A4 natural paper product with the number available 101-200pcs.

Programming Code

a. Display of Running Program (run)

In this display, Computer Friends must run this program (run) to be able to proceed to the next stage.

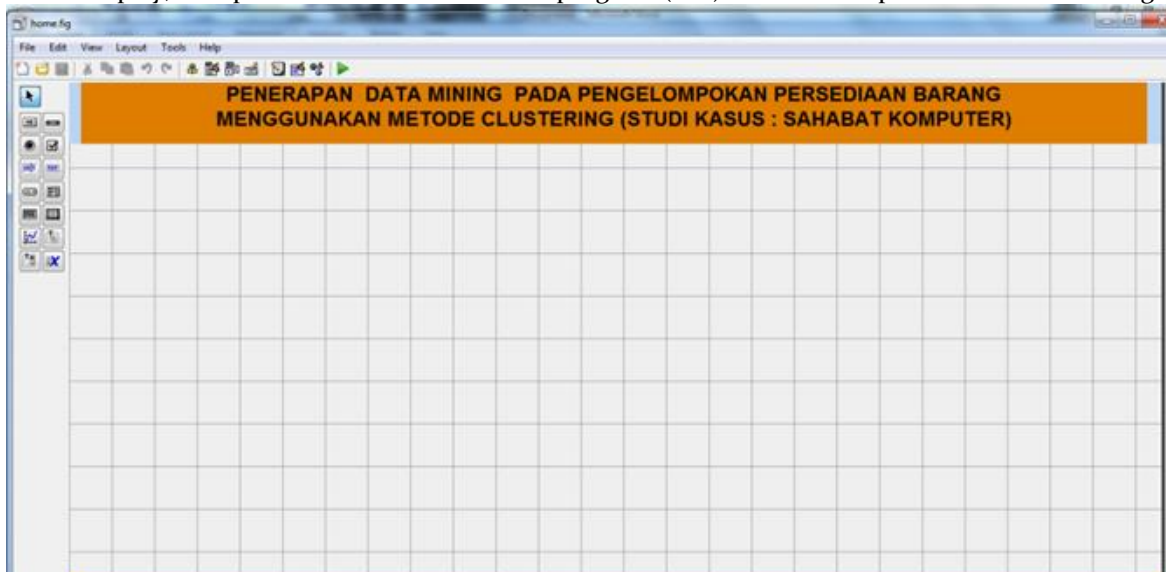


Figure 1. Display of Running Program (run)

b. Display of Change Directory Options

In this display it is only recommended to select the change directory option to proceed to the main page of the program.

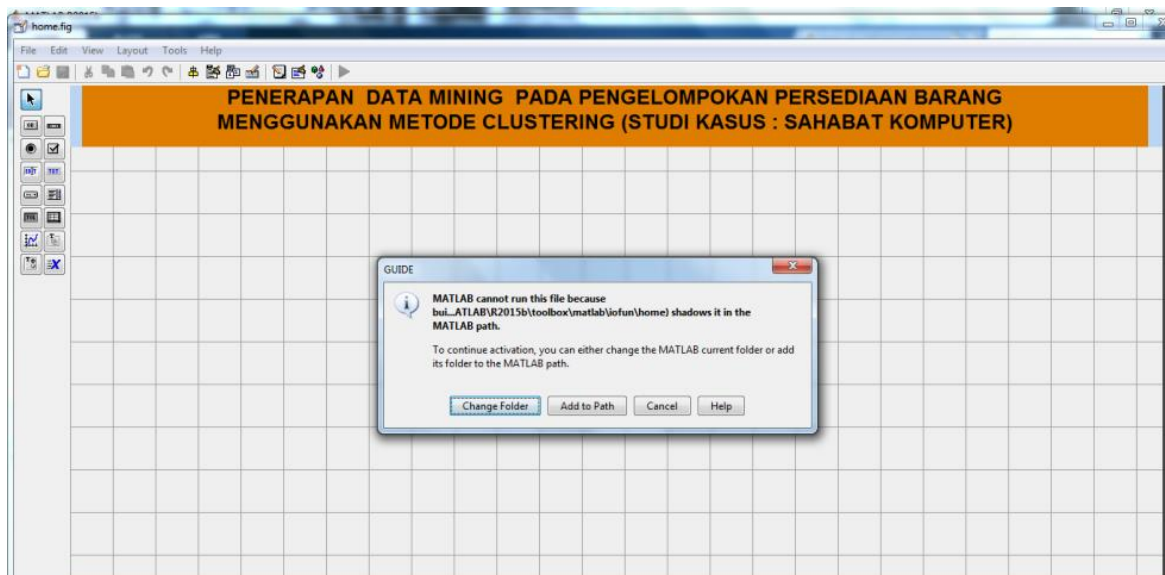


Figure 2. Display of Change Directory Options

c. Main Page Display

This main page display where users can only open the cluster process menu, cluster results, information and exit menu.



Figure 3. Main Page Display

d. Grouping Process Home View

When the user enters the cluster process menu or grouping, first input the data first and determine the number of clusters available on the display. After completing the data input, the next step is to click the cluster process, which will appear the cluster process display where the Centroid Results, centroid information and cluster members as well as diagrams are listed and have shown that the data previously input has been processed on the home display of the process.

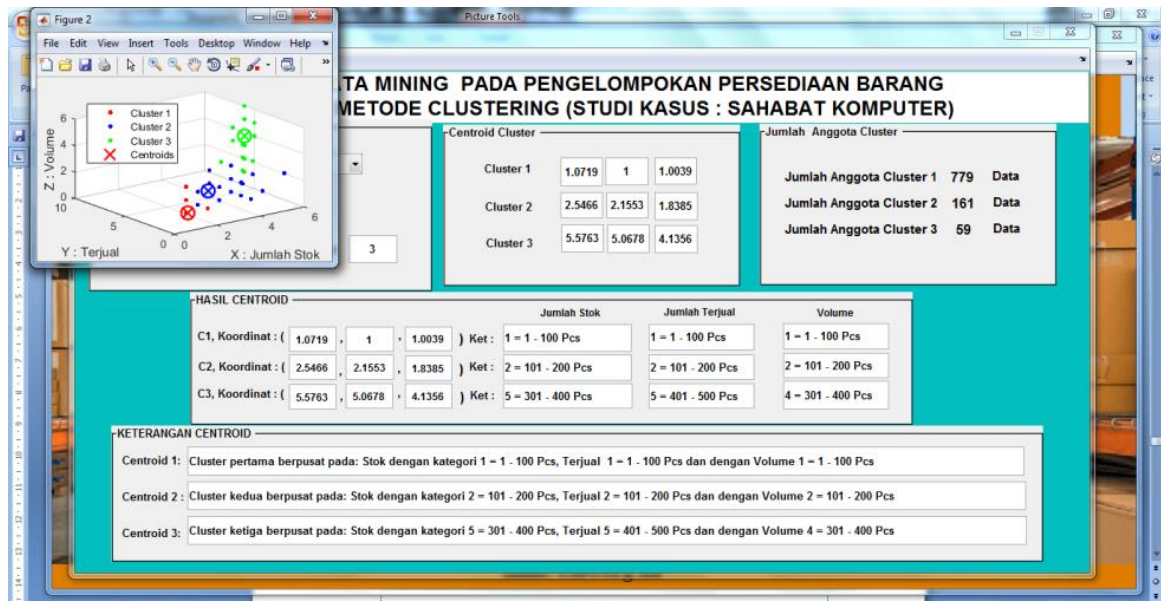


Figure 4. Home View of Grouping Process

e. Result Data Information Display

In this result data information display, when inputting all of the information, it will display several criteria from the inventory grouping of goods, namely the amount of stock, the amount sold and the sales volume.

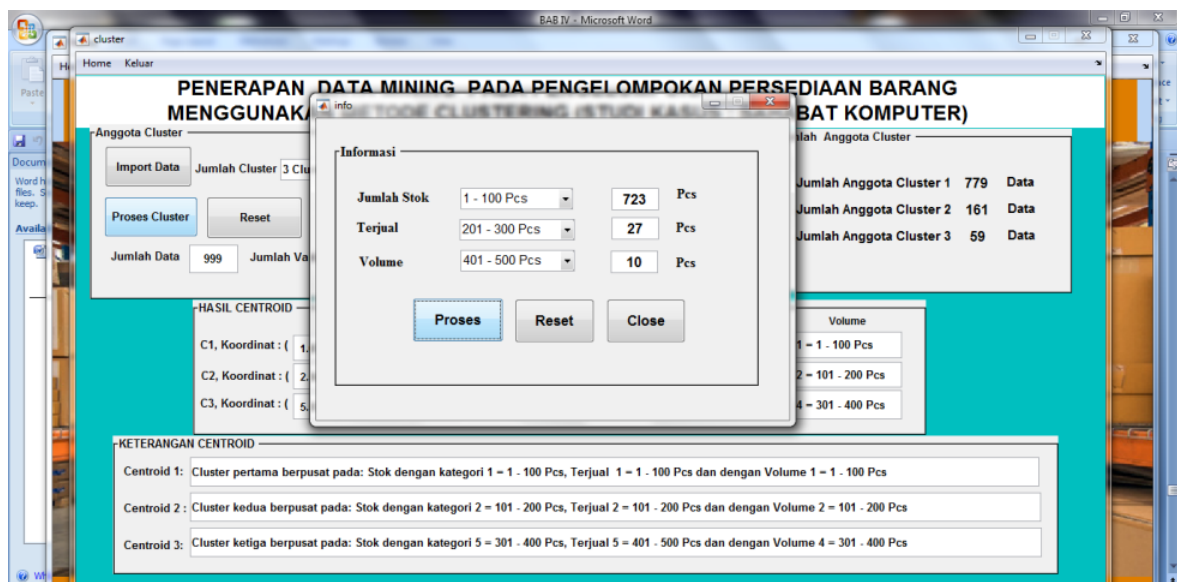


Figure 5. Result Data Information Display

Based on clustering calculations on 999 data, the results of the first cluster centroid centered on stock 1-100 pcs, goods sold 1-100 pcs and sales volume 1-100 pcs. In the second cluster the centroid is centered on stocks with categories 101-200 pcs, items sold 101-200pcs and sales volume 101-200 pcs. In the third cluster, the centroid is centered on stocks with categories of 301-400 pcs, items sold 401-500 pcs, and sales volume of 301-400 pcs.

5. Conclusion

The conclusions that can be drawn from the results of research in the development of a system for applying data mining to grouping inventory of goods are: (a). By using Data Mining Sahabat Komputer can see the grouping of inventory items based on the amount available, the amount sold and the sales volume; (b). With this new application system, Sahabat Komputer can see the grouping of dominant and minimal inventory items that occur in Sahabat Komputer; (c). This system can process large grouping data into useful information, so as to increase excellent knowledge; (d). With the application of this system, Sahabat Komputer can reduce errors in records that were previously done manually, so that the work system becomes more effective and efficient. (e). Furthermore, it can be concluded that in cluster 1 the stock is available on average 1-100 pcs, the number of sales is 1-100 pcs and the sales volume per month is 1-100 units. In cluster 2 there is an average available stock of 101-200 pcs, 101-200 pcs sales quantity of 101-200 units, and monthly sales volume of 101-200 pcs. And in cluster 3 there is an average available stock of 301-400 pcs, the number sold is 401-500 pcs, and the monthly sales volume reaches 301-400 pcs. Future research development suggestions to expand the scope of variables used in data mining analysis, including considering external factors such as market trends and seasonal factors. The integration of artificial intelligence technology can improve the accuracy of inventory prediction and assist in better decision-making. The implication is that more holistic research will provide a deeper understanding of inventory dynamics, helping companies to be more adaptive to market changes. The implementation of artificial intelligence technology can improve operational efficiency and minimize the risk of human error in inventory management, leading to improved performance and long-term profitability for Sahabat Komputer.

References

- [1] J. F. Kros, J. F. Kirchoff, and M. Falasca, "The impact of buyer-supplier relationship quality and information management on industrial vending machine benefits in the healthcare industry," *J. Purch. Supply Manag.*, vol. 25, no. 3, p. 100506, 2019.
- [2] W. C. Copacino, *Supply chain management: The basics and beyond*. Routledge, 2019.
- [3] K. Rubel, "Increasing the Efficiency and Effectiveness of Inventory Management by Optimizing Supply Chain through Enterprise Resource Planning Technology," *EFFLATOUNIA-Multidisciplinary J.*, vol. 5, no. 2, pp. 1739-1756, 2021.
- [4] E. Ramos, T. J. Pettit, M. Flanigan, L. Romero, and K. Huayta, "Inventory management model based on lean supply chain to increase the service level in a distributor of automotive sector," *Int. J. Supply Chain Manag*, vol. 9, no. 2, pp. 113-131, 2020.
- [5] Y. Guo, N. Wang, Z.-Y. Xu, and K. Wu, "The internet of things-based decision support system for information processing in intelligent manufacturing using data mining technology," *Mech. Syst. Signal Process.*, vol. 142, p. 106630, 2020.
- [6] S. Teerasoponpong and A. Sopadang, "Decision support system for adaptive sourcing and inventory management in small-and medium-sized enterprises," *Robot. Comput. Integr. Manuf.*, vol. 73, p. 102226, 2022.
- [7] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf. Sci. (Ny)*, vol. 622, pp. 178-210, 2023.
- [8] Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, and R. Kashef, "Recommendation systems: Algorithms, challenges, metrics, and business opportunities," *Appl. Sci.*, vol. 10, no. 21, p. 7748, 2020.
- [9] M. K. Gupta and P. Chandra, "A comprehensive survey of data mining," *Int. J. Inf. Technol.*, vol. 12, no. 4, pp. 1243-1257, 2020.
- [10] X. Liu, Y. Ding, H. Tang, and F. Xiao, "A data mining-based framework for the identification of daily electricity usage patterns and anomaly detection in building electricity consumption data," *Energy Build.*, vol. 231, p. 110601, 2021.
- [11] A. E. Ezugwu *et al.*, "A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects," *Eng. Appl. Artif. Intell.*, vol. 110, p. 104743, 2022.

- [12] D. Antons, E. Grünwald, P. Cichy, and T. O. Salge, "The application of text mining methods in innovation research: current state, evolution patterns, and development priorities," *R&D Manag.*, vol. 50, no. 3, pp. 329–351, 2020.
- [13] M. Subramaniyan, A. Skoogh, A. S. Muhammad, J. Bokrantz, B. Johansson, and C. Roser, "A generic hierarchical clustering approach for detecting bottlenecks in manufacturing," *J. Manuf. Syst.*, vol. 55, pp. 143–158, 2020.
- [14] M. Z. Rodriguez *et al.*, "Clustering algorithms: A comparative approach," *PLoS One*, vol. 14, no. 1, p. e0210236, 2019.
- [15] P. Govender and V. Sivakumar, "Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019)," *Atmos. Pollut. Res.*, vol. 11, no. 1, pp. 40–56, 2020.
- [16] C. Bouveyron, G. Celeux, T. B. Murphy, and A. E. Raftery, *Model-based clustering and classification for data science: with applications in R*, vol. 50. Cambridge University Press, 2019.
- [17] H. Teichgraber and A. R. Brandt, "Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison," *Appl. Energy*, vol. 239, pp. 1283–1293, 2019.
- [18] B. E. Jin and D. C. Shin, "Changing the game to compete: Innovations in the fashion retail industry from the disruptive business model," *Bus. Horiz.*, vol. 63, no. 3, pp. 301–311, 2020.
- [19] S. Verma and A. Gustafsson, "Investigating the emerging COVID-19 research trends in the field of business and management: A bibliometric analysis approach," *J. Bus. Res.*, vol. 118, pp. 253–261, 2020.
- [20] U. Awan, S. Shamim, Z. Khan, N. U. Zia, S. M. Shariq, and M. N. Khan, "Big data analytics capability and decision-making: The role of data-driven insight on circular economy performance," *Technol. Forecast. Soc. Change*, vol. 168, p. 120766, 2021.
- [21] R. G. Richey Jr, S. Chowdhury, B. Davis-Sramek, M. Giannakis, and Y. K. Dwivedi, "Artificial intelligence in logistics and supply chain management: A primer and roadmap for research," *Journal of Business Logistics*, vol. 44, no. 4. Wiley Online Library, pp. 532–549, 2023.
- [22] M. McMaster, C. Nettleton, C. Tom, B. Xu, C. Cao, and P. Qiao, "Risk management: Rethinking fashion supply chain management for multinational corporations in light of the COVID-19 outbreak," *J. Risk Financ. Manag.*, vol. 13, no. 8, p. 173, 2020.
- [23] J. Huber and H. Stuckenschmidt, "Daily retail demand forecasting using machine learning with emphasis on calendric special days," *Int. J. Forecast.*, vol. 36, no. 4, pp. 1420–1438, 2020.
- [24] G. D. Konstantakopoulos, S. P. Gayialis, and E. P. Kechagias, "Vehicle routing problem and related algorithms for logistics distribution: A literature review and classification," *Oper. Res.*, pp. 1–30, 2020.
- [25] M. Sinurat, M. Heikal, A. Simanjuntak, R. Siahaan, and R. N. Ilham, "Product Quality On Consumer Purchase Interest With Customer Satisfaction As A Variable Intervening In Black Online Store High Click Market: Case Study on Customers of the Tebing Tinggi Black Market Online Store," *Morfai J.*, vol. 1, no. 1, pp. 13–21, 2021.
- [26] N. H. Tien *et al.*, "The formation and development of CRM system at Thien Hoa electronics supermarket in Vietnam," *Int. J. Res. Growth Eval.*, vol. 2, no. 4, pp. 752–760, 2021.
- [27] R. Buaton and M. K. Relita, "Metode Menyelesaikan Data Mining, Sistem Pakar dan Sistem Pendukung Keputusan," *Medan. STMIK KAPUTAMA*, 15AD.