Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th February 2025. Vol. 4. No. 2; e-ISSN: 2808-4519

# Optimizing Grocery Sales Data Grouping Using the Fuzzy C-Means Algorithm: Case Study of Nafhan Mart Store

**Nafhan Khairuddin Fathin[1]*, Rudi Kurniawan[2], Saeful Anwar[3]**

[1,2,3]*STMIK IKMI CIREBON*

naphankhairudin@gmail.com[1]*, rudi226@gmail.com[2], saefulanwar419@gmail.com[3]

**Abstract**

The sale of staple food products at Nafhanmart Store, Cirebon Regency, includes essential household items such as rice, cooking oil, sugar, and flour, which maintain stable demand as basic necessities. This study focuses on improving sales clustering models at Nafhanmart using the Fuzzy C-Means (FCM) algorithm, a prominent method in data mining. Key factors influencing sales include price, sales volume, demand, and remaining stock. Accurate clustering analysis is vital for strategic inventory management and profit maximization. The research applies the Knowledge Discovery in Database (KDD) methodology, encompassing data selection, preprocessing, transformation, FCM implementation, and evaluation using the Davies-Bouldin Index (DBI). Attributes analyzed include price, sales volume, demand, and remaining stock. The FCM algorithm clusters data based on patterns, with DBI evaluating clustering quality and determining optimal clusters. Data analysis and visualization were conducted using RapidMiner. Results show that the FCM algorithm achieves optimal clustering quality with a DBI score of 0.452 for two clusters, outperforming three clusters (DBI 0.474) and four clusters (DBI 0.536). Price and demand are identified as critical factors influencing clustering outcomes. These findings enhance the clustering model, offering actionable insights for inventory management and sales strategy, while showcasing the FCM algorithm's adaptability for other SMEs to support data-driven decision-making.

***Keywords***: Fuzzy C-Means, sales clustering, staple foods, clustering, Davies boldin index

## 1. Introduction

The rapid growth in the field of informatics has had a significant impact on various aspects of life, including business, technology and education [1]. One of the main applications of informatics is the use of data mining-based algorithms to process data into meaningful information [2]. In the business world, effective data management is very important to assist the strategic decision making process, especially in the sales sector. The Nafhan Mart store in Cirebon Regency, as a provider of basic necessities (sembako), faces challenges in grouping product sales patterns in a relevant manner. The use of algorithms such as Fuzzy C-Means (FCM) is relevant to help group complex data more systematically and measurably, thus producing a reliable grouping model. Fuzzy C-Means clustering is a method in data mining that is used to group data into groups based on similarities in the degree of membership or certain characteristics [3].

The main challenge in the sales management process is the inability of manual methods to provide relevant groupings, so that there is often a mismatch between demand and stock [4]. In addition, the literature shows that clustering algorithms such as Fuzzy C-Means have not been widely applied to the local business sector, especially in optimizing grouping of sales of basic food products. Therefore, this problem is relevant to help operational efficiency and improve customer service at basic food stores [5], the Fuzzy C-Means (FCM) algorithm has become a tool that is often used in various research, but its application is in analyzing basic food sales in the market Local companies still face challenges in a shop that is still manual. Reliance on this kind of data can create obstacles in getting consistent analysis results. In addition, the current literature shows that there is a picture of optimizing FCM for specific applications, such as analysis of basic food sales. Previous research tends to focus more on the general application of FCM without considering 2 unique needs of the local market. Recent trends show that combining FCM with other algorithms can improve a model's ability to understand consumer behavior. This combination of algorithms also has the potential to produce more relevant and adaptive clustering of sales patterns. However, implementing this adaptive approach requires in-depth analysis to ensure its effectiveness in various market conditions. Therefore, further research is needed to overcome this problem, as well as develop more relevant and applicable solutions. It is hoped that the results of this kind of research can make an important contribution to the development of analytical data technology that supports the local business sector.

Previous studies have demonstrated the application of the Fuzzy C-Means (FCM) algorithm in various fields, such as customer data clustering [3], stock management strategy determination [6], and employee performance clustering [7]. However, the use of FCM for

predicting retail product sales, particularly in the staple food sector, remains underexplored. These studies also highlight limitations in integrating external factors, such as market trends and local consumption patterns, which present opportunities for further development.

This study aims to: Describe the clustering model for staple food sales using the Fuzzy C-Means (FCM) algorithm, as it incorporates membership degrees [8]. FCM was chosen for its ability to handle uncertainty, flexibility in clustering, and suitability for complex data such as sales patterns. Evaluate the quality of optimal clusters using the Davies Bouldin Index (DBI). Provide data-driven recommendations for stock management and sales planning [9]. The main contribution of this research is offering a data mining-based alternative solution that can be applied in the SME sector [10]. Practically, the findings are expected to assist staple food stores in optimizing decision-making processes by developing a sales clustering model for Nafhanmart Store using the FCM algorithm. This study also clusters sales data to gain deeper insights into consumer purchasing patterns. Its primary contribution lies in enriching the literature on FCM applications in staple food sales and providing practical benefits for business decision-making in the local retail sector. Furthermore, the study aims to serve as a reference for developing more adaptive data analysis methods in the future.

The Fuzzy C-Means (FCM) algorithm will be employed in this study using a quantitative approach and experimental techniques to develop a sales model for Nafhanmart Store in Cirebon Regency. The model will be trained using data from Nafhanmart, incorporating variables such as product prices, sales volume, remaining stock, and others. Utilizing machine learning data processing techniques, the model will be developed and evaluated to identify inventory needs and produce more custering. To ensure optimal performance, the study will assess the model using clustering techniques and other relevant clustering. By improving the accuracy of stock supply estimates, the proposed model aims to provide practical solutions for business inventory management.The research adopts a Knowledge Discovery in Database (KDD) framework, encompassing data selection, preprocessing, data transformation, FCM algorithm application, and model evaluation. Sales data is analyzed using the RapidMiner software, focusing on key attributes such as product price, sales volume, and remaining stock. Clustering quality is evaluated using the Davies-Bouldin Index to ensure the effectiveness of the grouping results. This approach is expected to enhance inventory management and support decision-making processes in retail business operations

This study is expected to significantly advance our understanding of the Fuzzy C-Means algorithm for clustering in the trade sector if it successfully achieves its objectives [11]. Beyond contributing to the field of informatics, the findings aim to establish a solid foundation for further research on machine learning applications in inventory management. This clustering model will assist practitioners, particularly in the trade sector, in managing the stock of staple food products more effectively, thus improving outcomes and reducing losses.Furthermore, this research could serve as a reference for future researchers interested in developing similar Fuzzy C-Means clustering algorithms in different contexts, fostering innovation in information technology. The proposed implementation approach is anticipated to drive the adoption of advanced technologies in inventory management, paving the way for broader data-driven solutions across the agricultural industry. Thus, this study not only contributes to the development of more adaptive and market-responsive technologies but also holds the potential to directly influence practical managerial practices.

The findings of this study will make a significant contribution to the advancement of Fuzzy C-Means (FCM)-based clustering methods [12]. The proposed clustering model can be utilized by small and medium-sized enterprises (SMEs) to enhance operational efficiency and support strategic decision-making [13]. Additionally, this research may serve as a reference for further studies aimed at integrating FCM with other algorithms for more complex applications [3]. The results are expected to improve the understanding of FCM implementation in analyzing staple food sales data [14]. The developed model can assist business operators in designing more effective marketing strategies and optimizing inventory management. Furthermore, it has the potential to provide deeper and more accurate insights into consumer needs. The study's findings may inspire the development of more innovative clustering algorithms, extending beyond the business sector into fields such as education and healthcare. The practical application of this model will support data-driven decision-making in local business environments, providing a competitive edge, particularly for SMEs in the era of digital transformation. Ultimately, the research also contributes significantly to advancing data analysis methods across various sectors [15].

## 2. Research Method

The research method uses the Fuzzy C-means algorithm to group data by partitioning the data into different groups. This method uses the KDD (Knowledge Discovery Database) stage, this stage is practical for carrying out data mining.

### 2.1. Data Analysis Techniques

The Fuzzy C-Means algorithm approach is used in the data mining analysis process using data techniques. The purpose of this procedure is to identify data patterns and relationships that will help reflect the need for basic food products at the Nafhan Mart Store. The data analysis process uses the Knowledge Discovery in Database (KDD) approach with the Fuzzy C-Means algorithm as the main method.

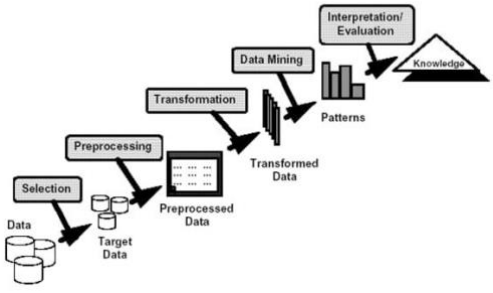


**Fig. 1:** Knowledge Discovery in Databases [8]

a. Data selection; At this stage, research data was systematically collected. A direct interview with the owner of Nafhanmart was conducted on October 8, 2024, to verify the accuracy and reliability of the data. Among the 500 data entries collected, key attributes included product names, sizes, prices, quantities received, quantities sold, remaining stock, units, and other characteristics related to inventory needs analysis. These attributes were specifically selected to assist in inventory management and to ensure that the clustering model could be developed using data relevant to the study's objectives.
b. Preprocessing; Carried out to improve data quality and prepare it for use at the next stage. The main activities at this stage include Data cleaning: removing inconsistent, empty, or invalid data.
c. Transformasi data; Transformations are carried out to organize and structure data to increase its significance. The transformation process in this research consists of: Grouping values (clustering): to find optimal clusters, data was collected using the K-Means clustering approach and the Davies Bouldin Index (DBI).The grouping results show that cluster 2 has the optimal value
d. Data mining; At this stage, the grouping model is built using the Fuzzy C-Means technique. To produce rules that can be used to group basic food product needs at Nafhanmart stores, Fuzzy C-Means looks for patterns in data sets. The weight of the item's price is highest between Size, Name of item, Quantity entered, Number of sizes, Remaining stock, Unit.
e. Evaluation; The clustering model that is built is evaluated using clusters from data such as: Davies Bouldin Index (DBI): used to evaluate the quality of clustering. -0.452 is the optimal DBI value, indicating that the resulting cluster is very representative. The results of this research show that the model developed can be applied well.
f. Knowledge; This stage is the final part of the Knowledge Discovery in Database (KDD) process which aims to produce useful information or knowledge from sales data of basic food products at the Nafhan Mart store. The knowledge obtained comes from the results of analysis using the Fuzzy C-Means algorithm which has gone through the stages of pre-processing, data transformation and data mining processes. The resulting knowledge is in the form of Clustering patterns which can be used to group basic food product needs. By using Fuzzy C-Means, rules or regulations are found that explain the relationship between attributes, such as item name, size, item price, incoming quantity, outgoing quantity, remaining stock, units, to the grouping of stock requirements. insight into which attributes influence the Davies Boldin Index value from the research results. Price of goods and number of entries have high and medium weights for clustering grouping.

After collecting data related to the Nafhanmart shop, Cirebon Regency, the data analysis technique that will be used in this research is the Fuzzy C-Means algorithm which will be implemented for What is the K value based on the best dbi value, What attributes influence the K value with the best dbi, How The application of the Fuzzy C-Means algorithm can help in product inventory management at the Nafhanmart Store. This analysis was carried out using the rapidminer tool which allows processing large amounts of data to produce efficient clusters.

## 3. Result and Discussion

The results of this research present data from the FUZZY C-MEANS ALGORITHM TO IMPROVE THE CLUSTERING GROUPING MODEL OF SALES DATA FOR GROCERY PRODUCTS SALES AT THE NAFHANMART STORE, CIREBON DISTRICT, which is taken from the attributes of item name, size, item price, incoming quantity, outgoing quantity of remaining stock, units. The next step is that the data will be grouped using the Fuzzy C-Means algorithm by applying the Knowledge Discovery in Database (KDD) method.

### 3.1. Data Selection

The first stage is data selection. At this stage, data and attributes are selected that are relevant and appropriate to the research to be conducted. The data to be processed is data on SALES OF FOOD PRODUCTS SALES AT NAFHANMART STORE, CIREBON DISTRICT with the attributes item name, size, item price, quantity, quantity out, remaining stock, units. With a total of 500 records.

**Table 1**: Results of the Read Excel operator

| No | Uraian | Keterangan |
|---|---|---|
| 1 | Record | 500 |
| 2 | Special atributes | 2 |
| 3 | Regular atributes | 5 |
| 4 | Attributes | |
| 5 | Nama barang | Nominal |
| 6 | Ukuran | Integer |
| 7 | Harga barang | Integer |
| 8 | Jumlah masuk | Integer |
| 9 | Jumlah keluar | Integer |
| 10 | Sisa stok | Integer |
| 11 | Satuan | Nominal |

Read Excel results show that the dataset used in this research consists of 500 records with 5 regular attributes and 2 special attributes. Each attribute has varying data types, such as Integer for numeric attributes Size, Item price, Incoming quantity, Outgoing quantity, and Remaining stock and Polynominal for categorical attributes Item name, Unit. The UNIT attribute is used as a label, while the other attributes have relevant values. This data structure provides an illustration that the dataset has good quality to be explained further in the transformation and classification process. With complete and structured data, this research can focus analysis on grouping needs for basic food sales products using Fuzzy C-Means.

### 3.2. Praprocessing

**Table 2**: Results of the Set Role operator

| No | Uraian | Keterangan |
|---|---|---|
| 1 | Record | 500 |
| 2 | Special attributes | 2 |
| 3 | Regular attributes | 5 |
| 4 | Attributes | |
| 5 | Nama barang (id) | Nominal 0 |
| 6 | Ukuran | Integer 0 |
| 7 | Harga barang | Integer 0 |
| 8 | Jumlah masuk | Integer 0 |
| 9 | Jumlah keluar | Integer 0 |
| 10 | Sisa stok | Integer 0 |
| 11 | Satuan ( label ) | Nominal 0 |

The Set Role results show that the dataset consists of 500 records, with 2 special attributes and 5 regular attributes. Attribute The item name is designated as an ID or Special Attribute because it functions as a unique identifier for each row of data in the dataset. And the unit attribute as a label because each unit is different in identifying the unit. This determination is important to ensure that each data entry can be clearly differentiated, avoid duplication, and make tracking easier during the analysis process. Apart from that, the item name attribute has a nominal data type with a missing value of 0, which indicates there is no empty data in this column, making it ideal as a primary key in data processing. Other attributes, such as size, incoming quantity, outgoing quantity, remaining stock, act as descriptive attributes used in the clustering and analysis process. By using the item name as the ID, the dataset structure becomes organized and makes it easier to further data processing in the grouping analysis stage using the Fuzzy C-Means algorithm.

### 3.3. Transformasi data

In the data transformation stage, the data that has been cleaned will be processed by changing the attribute type. At this stage the researcher uses the Nominal to Numerical operator to change the nominal data category into a numeric form of unit attributes which were initially in nominal form to become numeric because the clustering algorithm is more suitable for numeric data types.

**Table 3:** Result of nominal to numerical

| No | Nama barang | Satuan | Ukuran | Harga barang | Jumlah masuk | Jumlah keluar | Sisa stok |
|---|---|---|---|---|---|---|---|
| 1 | cappucino | Sachet | 1 | 1.250 | 20 | 2 | 18 |
| 2 | Kapal api mix | Sachet | 1 | 1.700 | 10 | 3 | 7 |
| 3 | Kopi gula aren | Sachet | 1 | 1.000 | 10 | 2 | 8 |
| 4 | Kopi creamy latte | Sachet | 1 | 1.500 | 15 | 5 | 10 |
| 5 | Kopi luwak | Sachet | 1 | 1.500 | 10 | 2 | 8 |
| …. | ….. | ….. | ….. | ….. | …… | ….. | ….. |
| 496 | Freascare aroma buah | Buah | 1 | 11.500 | 10 | 2 | 8 |
| 497 | Mama lemon jeruk tanggung | Reviel | 1 | 2.000 | 15 | 7 | 8 |
| 498 | Mama lemon carcol tanggung | Reviel | 1 | 2.000 | 15 | 5 | 10 |
| 499 | Sitrun | Sachet | 1 | 3.000 | 20 | 4 | 16 |
| 500 | Mie goreng sakura | Bungkus | 1 | 2.000 | 40 | 10 | 30 |

The results of this process explain the attributes used in the parameter process in the form of Subsets because it is possible to choose which attributes such as "product type, item name, size, item price, incoming quantity, outgoing quantity, remaining stock, units" will be changed. By doing this, it can be ensured that only related attributes are handled, so that other attributes are not affected. Then dummy coding assigns a unique number to each group (, category A=1, Category B=2), a grouping method that uses numerical data.

### 3.4. Data mining

At the data mining stage, researchers used the k-means algorithm to group sales data for Nafhanmart stores in Cirebon Regency. At this stage there are several stages that the researcher will carry out, namely the first stage is carrying out the clustering process using the Fuzzy C-Means algorithm, the second stage is testing and broadcasting the results of the clustering using the Cluster Distance Performance operator with the David Boulding Index (DBI) evaluation method used. The first stage is to create a clustering model using the Fuzzy C-means algorithm. Then it will carry out the process of checking the best cluster value (K value) which is carried out from testing k=2 to k-10.

**Table 4:** Result cluster model

| No | Cluster | Keterangan |
|---|---|---|
| 1 | Cluster 0 | 48 items |
| 2 | Cluster 1 | 452 items |
| | Total | 500 items |

Based on the cluster results, it shows that the best cluster value, namely the one with the lowest DBI value or close to 0, is in cluster 2 with a DBI value - 0.452, whose members are cluster 0: 48 items, cluster 2: 452 items.

Below are the cluster members, there are 2 clusters with the lowest DBI value - 0.452 which is considered good, the lower the value the better for cluster division. The following is a table of cluster 0 results.

**Table 5:** Cluster 0 result

| No | Nama barang | Ukuran | Harga barang | Jumlah masuk | Jumlah keluar | Sisa stok | Satuan |
|---|---|---|---|---|---|---|---|
| 1. | Rokok neslite | 1 | 27.000 | 10 | 2 | 8 | Bungkus |
| 2. | Rokok samsu kretek | 1 | 20.000 | 10 | 4 | 6 | Bungkus |
| 3. | Rokok samsu reviel | 1 | 21.500 | 10 | 2 | 8 | Bungkus |
| 4. | Rokok gugang garam merah | 1 | 15.500 | 10 | 3 | 7 | Bungkus |
| 5. | Rokok gudang garam filter | 1 | 25.500 | 10 | 1 | 9 | Bungkus |
| 6. | Rokok gudang garam jaya | 1 | 15.000 | 10 | 2 | 8 | Bungkus |
| 7. | Rokok galang baru 12 | 1 | 17.000 | 10 | 5 | 5 | Bungkus |
| 8. | Rokok galang baru 16 | 1 | 21.500 | 10 | 2 | 8 | Bungkus |
| 9. | Rokok sampoerna kretek | 1 | 15.500 | 10 | 3 | 7 | Bungkus |
| 10. | Rokok sampoerna filter | 1 | 34.000 | 10 | 3 | 7 | Bungkus |
| 11. | Rokok 303 | 1 | 13.500 | 10 | 3 | 7 | Bungkus |
| 12. | Rokok djarum coklat | 1 | 16.500 | 10 | 2 | 8 | Bungkus |
| 13. | Rokok surya pro merah | 1 | 31.000 | 10 | 1 | 9 | Bungkus |
| 14. | Rokok surya 12 | 1 | 26.000 | 10 | 1 | 9 | Bungkus |
| 15. | Rokok surya 16 | 1 | 34.600 | 10 | 1 | 9 | Bungkus |
| 16. | Rokok esse juicy | 1 | 40.000 | 10 | 3 | 7 | Bungkus |
| 17. | Water presh | 1 | 16.500 | 5 | 2 | 3 | Dus |
| 18. | Sarden abc kaleng kecil | 1 | 11.500 | 5 | 1 | 4 | Kaleng |
| 19. | Susu kaleng bendera putih 370 g | 1 | 12.500 | 5 | 1 | 4 | Kaleng |
| 20. | Susu kaleng bendera coklat 370 g | 1 | 12.500 | 5 | 1 | 4 | Kaleng |
| 21. | Minyak sayur | 1 | 18.000 | 5 | 2 | 3 | Kg |
| 22. | Minyak lentik | 1 | 30.000 | 2 | 1 | 1 | Kg |
| 23. | Gula putih | 1 | 17.000 | 3 | 1 | 2 | Kg |
| 24. | Gula merah | 1 | 18.000 | 3 | 1 | 2 | Kg |
| 25. | Teh rio | 1 | 19.000 | 24 | 5 | 19 | Gelas |
| 26. | Kopi kap | 1 | 18.500 | 24 | 7 | 17 | Gelas |
| 27. | Teh gelas | 1 | 19.000 | 24 | 6 | 18 | Gelas |
| 28. | levontea | 1 | 18.000 | 24 | 10 | 14 | Gelas |
| 29. | Panther | 1 | 18.500 | 24 | 2 | 22 | Gelas |
| 30. | Minyak kayu putih caplang 30 ml | 1 | 11.000 | 6 | 1 | 5 | Botol |
| 31. | Minyak kayu putih caplang 60 ml | 1 | 21.000 | 6 | 1 | 5 | Botol |
| 32. | Minyak kayu putih caplang 120ml | 1 | 39.500 | 6 | 1 | 5 | Botol |
| 33. | Cat rambut hi-top | 1 | 13.000 | 5 | 2 | 3 | Botol |
| 34. | Cat rambut tancho | 1 | 14.500 | 5 | 4 | 1 | Botol |
| 35. | Obat tetes mata insto | 1 | 14.000 | 5 | 2 | 3 | Botol |
| 36. | Obat tetes mata rohto | 1 | 13.000 | 6 | 3 | 3 | Botol |
| 37. | Aci | 1 | 12.000 | 5 | 1 | 4 | Kg |
| 38. | Beras | 1 | 14.000 | 25 | 5 | 20 | Kg |
| 39. | Rokok signature | 1 | 24.000 | 10 | 3 | 7 | Bungkus |
| 40. | Rokok djarum super | 1 | 24.500 | 10 | 3 | 7 | Bungkus |
| 41. | Karet gelang | 1 | 30.000 | 2 | 1 | 1 | Kg |
| 42. | Telur | 1 | 27.000 | 3 | 1 | 2 | Kg |
| 43. | Pepsodent besar | 1 | 12.500 | 6 | 2 | 4 | Biji |
| 44. | Frescare hot | 1 | 11.500 | 5 | 3 | 2 | Buah |
| 45. | Frescare citrus | 1 | 11.500 | 5 | 1 | 4 | Buah |
| 46. | Freascare lavender | 1 | 11.500 | 5 | 2 | 3 | Buah |
| 47. | Freascare grinty | 1 | 11.500 | 10 | 2 | 8 | Buah |
| 48. | Freascare aroma buah | 1 | 11.500 | 10 | 2 | 8 | Buah |

**Table 6:** Cluster 1 result

| no | Nama barang | Ukuran | Harga barang | Jumlah masuk | Jumlah keluar | Sisa stok | Satuan |
|---|---|---|---|---|---|---|---|
| 1. | Cappuccino | 1 | 1.250 | 20 | 2 | 18 | Sachet |
| 2. | Kapal api mix | 1 | 1.700 | 10 | 3 | 7 | Sachet |
| 3. | Kopi gula aren | 1 | 1.000 | 10 | 2 | 8 | Sachet |
| 4. | Kopi creamy latte | 1 | 1.500 | 15 | 5 | 10 | Sachet |
| 5. | Kopi luwak | 1 | 1.500 | 10 | 2 | 8 | Sachet |
| 6. | Kopi abc klepon | 1 | 2.000 | 15 | 1 | 14 | Sachet |
| 7. | Kopi choholatos creamy | 1 | 1.500 | 10 | 4 | 6 | Sachet |
| 8. | Kopi abc susu | 1 | 1.500 | 5 | 4 | 1 | Sachet |
| 9. | Kopi indocafe | 1 | 1.500 | 15 | 10 | 5 | Sachet |
| 10. | Kopi good day | 1 | 1.500 | 15 | 8 | 7 | Sachet |
| …. | …. | ….. | ….. | ….. | ….. | ….. | …… |
| 443. | Susu kental manis bendera gold | 1 | 2.000 | 12 | 7 | 5 | Sachet |
| 444. | Susu kental manis bendera strawberry | 1 | 1.500 | 6 | 4 | 2 | Sachet |
| 445. | Arang | 1 | 7.000 | 5 | 2 | 3 | Kg |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 446. | Solatif kecil hitam | 1 | 1.500 | 12 | 4 | 8 | Buah |
| 447. | Solatif kecil bening | 1 | 1.000 | 12 | 2 | 10 | Buah |
| 448. | Saos bantal | 1 | 4.000 | 5 | 2 | 3 | Buah |
| 449. | Mama lemon jeruk tanggung | 1 | 2.000 | 15 | 7 | 8 | reviel |
| 450. | Mama lemon carcol tanggung | 1 | 2.000 | 15 | 5 | 10 | reviel |
| 451. | Sitrun | 1 | 3.000 | 20 | 4 | 16 | Sachet |
| 452. | Mie goreng sakura | 1 | 2.000 | 40 | 10 | 30 | Bungkus |

### 3.5. Evaluation

At the evaluation stage, clustering results will be tested by applying the DBI value to the Cluster Distance Performance parameter. In DBI testing, the cluster that has the smallest DBI value or close to 0 is used as the best cluster. To find the best cluster, researchers conducted experiments from cluster 2 to cluster 10. The number of k clusters was generated from the Davies Bouldin Index value.

**Table 7:** Cluster value test results (K=2 – K=10)

| Clustering | Jumlah anggota cluster | Hasil nilai DBI |
|---|---|---|
| 2 | Cluster 0 : 48 items<br>Cluster 1 : 452 items<br>Total of number of items 500 | Davies boldin : - 0.452 |
| 3 | Cluster 0 : 58 items<br>Cluster 1 : 425 items<br>Cluster 2 : 17 items<br>Total number of items : 500 | Davies boldin : - 0.474 |
| 4 | Cluster 0 : 139 items<br>Cluster 1 : 13 items<br>Cluster 2 : 44 items<br>Cluster 3 : 304 items<br>Total number of items : 500 | Davies boldin : - 0.536 |
| 5 | Cluster 0 : 11 items<br>Cluster 1 : 37 items<br>Cluster 2 : 153 items<br>Cluster 3 : 279 items<br>Cluster 4 : 20 items<br>Total number of items : 500 | Davies boldin : - 0.516 |
| 6 | Cluster 0 : 225 items<br>Cluster 1 : 10 items<br>Cluster 2 : 20 items<br>Cluster 3 : 163 items<br>Cluster 4 : 32 items<br>Cluster 5 : 50 items<br>Total number of items : 500 | Davies boldin : - 0.521 |
| 7 | Cluster 0 : 153 items<br>Cluster 1 : 112 items<br>Cluster 2 : 21 items<br>Cluster 3 : 151 items<br>Cluster 4 : 35 items<br>Cluster 5 : 19 items<br>Cluster 6 : 9 items<br>Total number of items : 500 | Davies boldin : - 0.524 |
| 8 | Cluster 0 : 110 items<br>Cluster 1 : 9 items<br>Cluster 2 : 18 items<br>Cluster 3 : 31 items<br>Cluster 4 : 154 items<br>Cluster 5 : 150 items<br>Cluster 6 : 24 items<br>Cluster 7 : 4 items<br>Total number of items : 500 | Davies boldin : - 0.494 |
| 9 | Cluster 0 : 129 items<br>Cluster 1 : 17 items<br>Cluster 2 : 9 items<br>Cluster 3 : 4 items<br>Cluster 4 : 150 items<br>Cluster 5 : 109 items<br>Cluster 6 : 37 items<br>Cluster 7 : 27 items<br>Cluster 8 : 18 items<br>Total number of items : 500 | Davies boldin : - 0.487 |
| 10 | Cluster 0 : 9 items<br>Cluster 1 : 129 items<br>Cluster 2 : 19 items<br>Cluster 3 : 15 items<br>Cluster 4 : 13 items<br>Cluster 5 : 4 items<br>Cluster 6 : 24 items | Davies boldin : - 0.469 |

| | | |
|---|---|---|
| | Cluster 7 : 150 items<br>Cluster 8 : 95 items<br>Cluster 9 : 42 items<br>Total number of items : 500 | |

### 3.5. Knowledge

At this stage we will discuss the interpretation and evaluation of the results obtained in the previous stages. Here the researcher will discuss in more detail the results of the DBI values from the cluster test values k=2 to k=10.

**Table 8:** Davies Boldin Index (DBI) value results

| Cluster | Davies Boldin Index (DBI) |
|---|---|
| 2 | - 0.452 |
| 3 | - 0.474 |
| 4 | - 0.536 |
| 5 | - 0.516 |
| 6 | - 0.521 |
| 7 | - 0.524 |
| 8 | - 0.494 |
| 9 | - 0.487 |
| 10 | - 0.469 |

From the David Bouldin Index (DBI) value table above, it can be concluded that the optimal cluster value is cluster 2 (k=2) with a DBI result of -0.452. The lower the DBI value, the better the cluster formed. in cluster 2 (k=2) it is divided into 2 clusters. With cluster_0 members: 48 items cluster 1 452 items.

## 4. Conclusions and Suggestions

### 4.1. Conclusions

Based on the results of this research, the discussion that has been carried out can be concluded as follows:

1. The K value is determined from the best Davies Boldin Index (Davies-Bouldin Index) value, namely the smallest Davies Boldin Index value. The best Davies Boldin Index is 0.452 at K = 2. This means that the best number of clusters for clustering is 2 clusters. Thus, the owner must continue to pay attention to the price of goods because it has a very high influence on sales data.
2. The attributes that influence the results of the best Davies Boldin Index and K values are all the attributes involved in the analysis. Based on the table, these attributes include item name, size, item price, incoming quantity, outgoing quantity, remaining stock, units. However, the main product attributes that influence sales patterns, namely; "price of goods", "quantity in/out", and "remaining stock" usually have a greater influence because they are directly related to sales and inventory trends. used in the Fuzzy C-Means algorithm to calculate the distance between data, determine the cluster center, and produce the Davies Boldin Index value.
3. Factors that influence the grouping results are the number of attributes used, data normalization, selection of K values, Fuzzy C-Means algorithm parameters, and data distribution.

### 4.2. Suggestions

Based on the explanation of the research results above, there are several suggestions, namely:

1. Collecting a larger and more complete dataset for the Fuzzy C-Means clustering method to make sales data more accurate, sales data for 2024.
2. To develop this research, the application used for Fuzzy C-Means clustering analysis can use a more modern and sophisticated application.
3. This research can add data to use the clustering method to group various Nafhanmart store sales products.

## References


[1] Ade Andani, Erine Novianta, Adhe Friam Budhi, Adek Arya, Zalika Sabina, and Wahyu Indah Sari, "Pengaruh Perkembangan Ekonomi Digital Terhadap Perusahaan Start-Up di Indonesia," *J. Kaji. dan Penelit. Umum*, vol. 2, no. 1, pp. 01–11, 2024, doi: 10.47861/jkpu-nalanda.v2i1.797.

[2] R. Kurniah, "Analisa Dan Penerapan Metode Klasifikasi Dalam Data Mining Untuk Penerimaan Siswa Jalur Non-Tulis," *J. Ilm. Inform.*, vol. 8, no. 01, p. 9, 2020.

[3] E. S. Rahmadina, B. Irawan, and A. Bahtiar, "Penerapan Algoritma Fuzzy C-Means Pada Data Penjualan Distro," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 2348–2354, 2024, doi: 10.36040/jati.v8i2.8467.

[4] R. Putri *et al.*, "Estimasi Penjualan Pada Toko Damhil Cake and Bakery Menentukan Strategi Penjualan Kue Kering," vol. 8, no. 1, pp. 429–436, 2024.

[5] A. Kineta, "Strategi Pengembangan Toko Kelontong di Kecamatan Gayungan," vol. 2, no. 4, pp. 163–177, 2024.

[6] A. C. Putra and K. D. Hartomo, "Optimalisasi Penyaluran Bantuan Pemerintah Untuk UMKM Menggunakan Metode Fuzzy C-Means," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 1, no. 10, pp. 474–482, 2021, [Online]. Available: https://jurnal.iaii.or.id/index.php/RESTI/article/view/2980/423

[7] A. E. Pramitasari and Y. Nataliani, "Perbandingan Clustering Karyawan Berdasarkan Nilai Kinerja Dengan Algoritma K-Means Dan Fuzzy C-Means," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 3, pp. 1119–1132, 2021, doi: 10.35957/jatisi.v8i3.957.

[8] B. Christian and L. Hakim, "Penerapan Algoritma Fuzzy C-Means Pada Penentuan Lokasi Gudang Pendukung PT. XYZ," *Aiti*, vol. 16, no. 1, pp. 31–48, 2019, doi: 10.24246/aiti.v16i1.31-48.

[9] M. Mardiansyah, A. K. Rivai, and D. Kurnia, "Analisis Perencanaan Produksi Mebel Menggunakan Perangkat Lunak Berbasis Fuzzy Linear

Programming serta Pengujian Akurasi Hasil Menggunakan Data Training dan Data Testing," *J. Inform. Univ. Pamulang*, vol. 6, no. 2, pp. 232–239, 2021, [Online]. Available: http://openjournal.unpam.ac.id/index.php/informatika

[10] R. Putriyandari, W. Yuliyana, and Y. S. Rahayu, "Laporan keuangan UMKM dalam menunjang analisis pemberian kredit perusahaan UMKM," *J. Abdimas BSI J. Pengabdi. Kpd. Masy.*, vol. 2, no. 1, pp. 190–209, 2019, [Online]. Available: http://ejournal.bsi.ac.id/ejurnal/index.php/abdimas/article/view/4994/2982

[11] M. Farhan Nugraha, M. Martano, and U. Hayati, "Clustering Data Indonesian Food Delivery Menggunakan Metode K-Means Pada Gofood Product List," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 3, pp. 3484–3492, 2024, doi: 10.36040/jati.v8i3.9727.

[12] K. A. Ngii, D. Sabrina, and R. A. Saputra, "Implementasi Algoritma Fuzzy C-Means (Fcm) Dalam Memprediksi Hasil Tangkapan Ikan Di Kota Kendari," *METHOMIKA J. Manaj. Inform. dan Komputerisasi Akunt.*, vol. 7, no. 2, pp. 319–324, 2023, doi: 10.46880/jmika.vol7no2.pp319-324.

[13] D. Azlil Huriah and N. Dienwati Nuris, "Klasifikasi Penerima Bantuan Sosial Umkm Menggunakan Algoritma Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 360–365, 2023, doi: 10.36040/jati.v7i1.6300.

[14] E. Sholekha, B. Irawan, and A. Bahtiar, "Analisis Penjualan Produk Snack Dan Minuman Menggunakan Metode K-Means Pada Dataset Transaksi Penjualan," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 3, pp. 2533–2539, 2024, doi: 10.36040/jati.v8i3.9310.

[15] U. S. Munawir Sulistyawati, "Membangun Keunggulan Kompetitif melalui Platform E- Commerce : Studi Kasus Tokopedia Abstrak," vol. 1, no. 1, pp. 43–55, 2024, [Online]. Available: https://journal.stmiki.ac.id/index.php/jmt/article/view/776/558