International Journal Software Engineering and Computer Science (IJSECS)
4 (2), 2024, 418-429
Published Online August 2024 in IJSECS (http://www.journal.lembagakita.org/index.php/ijsecs)
P-ISSN : 2776-4869, E-ISSN : 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v4i2.2863.

RESEARCH ARTICLE Open Access

Grouping Production Goods Requirements Using the K-Means Clustering Method

Dani Yuda Dwi Setiawan *
Informatics Engineering Study Program, Faculty of Engineering, Universitas Pelita Bangsa, Deli Serdang Regency, North Sumatra, Indonesia.
Corresponding Email: hieronimusdani@gmail.com.

Wahyu Hadikristanto
Informatics Engineering Study Program, Faculty of Engineering, Universitas Pelita Bangsa, Deli Serdang Regency, North Sumatra, Indonesia.
Email: wahyu.hadikristanto@pelitabangsa.ac.id.

Edora
Informatics Engineering Study Program, Faculty of Engineering, Universitas Pelita Bangsa, Deli Serdang Regency, North Sumatra, Indonesia.
Email: edora@pelitabangsa.ac.id.

Received: May 8, 2024; Accepted: July 10, 2024; Published: August 1, 2024.

Abstract: The inventory management of production goods presents several challenges, including difficulties in distinguishing between necessary and unnecessary items, leading to overstocking and manual data processing issues. Additionally, the risk of data loss can impede the data processing workflow. Data testing is conducted to evaluate the accuracy of calculations and the functionality of the applied methods. The objective is to optimize production results and inventory levels in warehouses. The K-means algorithm, known for its simplicity and effectiveness, is utilized to identify clusters within the data. The first cluster (C0) has centroids at (60.33, 70.33) and includes stock data categorized as having no potential. This cluster comprises 35 records. The second cluster (C1) has centroids at (10.94, 7.11) and includes stock data categorized as available, consisting of 15 records. Testing with the RapidMiner Studio application confirms similar insights, with each cluster containing members that are divided into two clusters, each having optimal centroid values of (60.33, 70.33) for Cluster 1 (C0) and (10.94, 7.14) for Cluster 2 (C1), and a Davies-Bouldin Index evaluation score of 0.666.

Keywords: Data Mining; Stock Requirements; Production Results; Clustering; K-Means.

1. Introduction

The categorization of production goods requirements is a critical aspect of production management and supply chain operations. This process involves identifying, classifying, and organizing various types of goods and raw materials necessary for manufacturing a final product. Companies must continually devise strategies to sustain and potentially expand their business operations. Effective management of inventory availability and the optimization of necessary goods are essential for companies in the manufacturing sector. The challenge lies in identifying the exact needs, which are often influenced by market conditions, consumer demand, and product innovation. One of the primary issues companies face is accurately identifying their requirements. The production goods requirements frequently change based on market conditions, consumer demand, and innovations in products. If the categorization of these needs is not carried out meticulously, companies may experience either shortages or surpluses. This imbalance can affect production efficiency and operational costs. Inefficient stock management and poor coordination between various departments can lead to high storage costs for excess inventory or production delays due to insufficient stock.

The solution to these problems can be found in the implementation of the K-Means method. This approach allows companies to minimize inventory and produce goods according to actual needs. Effective communication between departments is also crucial to ensure that every part of the supply chain has accurate and up-to-date information regarding production goods requirements. The K-Means algorithm is applied to cluster production needs, aiding in the formulation of strategies to target production requirements and enhance customer satisfaction. This algorithm groups data items into small clusters, each sharing essential similarities. K-Means is a widely-used algorithm in the clustering process. It seeks a specified number of clusters based on the proximity of data points to each other. This method helps in determining product needs and defining production goods clusters. The cluster analysis for processing production goods data in data mining, using K-Means, is a vector quantization method aligned with clustering problems [1].

Data mining using the K-Means clustering algorithm facilitates inventory collection, strategy formulation, and categorization into production data and warehouse material data. The rationale for using the K-Means method is its efficiency in analyzing and classifying large production data sets quickly. Consequently, this study conducts data mining analysis using clustering techniques with the K-Means algorithm [2]. The K-Means algorithm iteratively performs steps until stability is achieved. The resulting clusters indicate the types of goods required each month, which can be used as a reference for stock planning—whether there will be sufficient stock or not in the following month. Given the issues described, this research is titled "Categorizing Production Goods Requirements Using the K-Means Clustering Method." The study aims to enhance the understanding of production goods requirements and their potential production needs.

In the manufacturing industry, maintaining optimal inventory levels is crucial for balancing production and operational costs. Companies often struggle with fluctuating demand and the continuous evolution of consumer preferences. The inability to accurately forecast these needs can lead to significant inefficiencies. For instance, overstocking results in high storage costs, while understocking can disrupt production schedules and delay order fulfillment. The K-Means clustering method offers a practical solution to these challenges by categorizing production goods based on their demand patterns. By grouping similar items, companies can streamline their inventory management processes and reduce unnecessary costs. This approach not only improves efficiency but also enhances the company's ability to respond to market changes swiftly. Moreover, the adoption of advanced data analysis tools, such as the K-Means algorithm, empowers companies to make informed decisions, ultimately leading to better resource allocation and increased profitability.

Furthermore, the application of K-Means clustering in inventory management supports the development of more robust production strategies. By understanding the specific needs of different clusters, companies can prioritize their resources and focus on high-demand items. This targeted approach ensures that production schedules are aligned with actual market demand, reducing the risk of overproduction or underproduction. As a result, companies can achieve a more balanced and efficient production process, leading to improved overall performance. In addition to enhancing inventory management, the K-Means method also facilitates better communication and coordination between departments. By providing a clear categorization of production goods, all departments involved in the supply chain can access accurate and up-to-date information. This transparency helps to eliminate misunderstandings and ensures that all parts of the organization are working towards the same goals. Effective inter-departmental communication is essential for maintaining a seamless production process and achieving optimal operational efficiency. The importance of using K-Means for clustering in data mining cannot be overstated. This algorithm is effective in handling large datasets and can provide valuable insights into inventory needs. Its ability to segment data into meaningful clusters allows for

a more detailed analysis of production requirements. This level of analysis is crucial for developing strategies that are both efficient and effective in meeting consumer demands.

Given the challenges faced by manufacturing companies in managing inventory and production needs, this study aims to provide a comprehensive solution through the application of the K-Means clustering method. By improving the accuracy of production goods categorization, companies can enhance their inventory management processes, reduce costs, and improve overall efficiency. This study not only addresses the immediate needs of production management but also contributes to the long-term sustainability and growth of manufacturing companies. The implementation of the K-Means clustering method offers significant benefits for the management of production goods requirements. By accurately categorizing inventory needs, companies can optimize their production processes, improve efficiency, and reduce operational costs. This study highlights the importance of advanced data analysis techniques in enhancing production management and provides a practical solution for addressing the challenges faced by manufacturing companies

2. Research Method

The research methodology involves implementing an approach to data processing using the K-Means algorithm to facilitate the research objectives set for processing production goods requirements. The following outlines the comprehensive steps undertaken in this research:

2.1. Data Collection

In this study, data collection was not limited to interviews, literature studies, and internet research alone. The primary data collection method involved obtaining data directly from the research object. Specifically, data were gathered from reports on the painting process and production goods requirements. This approach ensured that the data used in the research were directly relevant and up-to-date, providing a solid foundation for the subsequent analysis stages.

2.2. Data Processing

Data processing in this study utilized clustering techniques in data mining to categorize stock requirements using the K-Means algorithm. The data were processed to form a dataset that was then divided into training and testing datasets. This separation was crucial for ensuring that the model developed could generalize well when performing clustering on the data. The training data consisted of production goods requirements used to create clusters or run the functions of the algorithm as intended. The testing data were used to evaluate the accuracy or performance of the data clustering process.

2.3. Data Cleaning

The data cleaning stage involved selecting the data to be used in the research, removing any unnecessary data, eliminating duplicates, checking for inconsistencies, and correcting errors. This step was essential to ensure the quality and reliability of the data before analysis. For example, as shown in Table 1, the data were meticulously reviewed to identify and rectify any discrepancies, ensuring that only relevant and accurate data were included in the dataset.

Table 1. Data Selection

Required Goods Available	Mutation	Required Goods Not Available	Goods Status
20	0	5	Available
20	0	3	Available
20	0	5	Available
3	0	3	Not Available
10	0	5	Available

2.4. Data Warehousing

Data warehousing involved selecting data from operational data before it entered the mining data or information stage. The data that had gone through the selection process were stored in Excel files, ready for further processing. During this stage, attributes necessary for the next steps were selected, while irrelevant attributes were reduced to make the data dimensions more concise. Table 2 illustrates the initial dataset attributes that were streamlined for the modeling process.

Table 2. Initial Dataset Attributes

Required Goods Available	Required Goods Not Available	Goods Status
20	15	Available
20	17	Available
20	15	Available
3	0	Not Available
10	10	Available

During the attribute selection process, attributes unrelated to the clustering modeling and K-Means algorithm were excluded. Only the essential numerical attributes were retained, as shown in Table 3.

Table 3. Selected Attributes

Required Goods Available	Required Goods Not Available	Goods Status
20	15	Available
20	17	Available
20	15	Available
3	0	Not Available
10	10	Available

2.5. Data Transformation

Data transformation involved converting the initial data format into a standard format suitable for the K-Means algorithm in the program or application used. The transformed data were then tested using various programs or tools to ensure compatibility. Table 4 shows the dataset after undergoing the transformation process, ready for further analysis.

Table 4. Data Transformation (Dataset)

Required Goods Available	Required Goods Not Available	Goods Status
20	15	Available
20	17	Available
20	15	Available
3	0	Not Available
10	10	Available

2.6. Data Mining

In the data mining stage, data modeling was conducted using the K-Means algorithm. This method is a straightforward clustering technique that applies a high level of independence, assuming that the value of an input attribute in a given class is independent of the value of other attributes. The tools used for this process included RapidMiner Studio. K-Means clustering is straightforward, involving the efficient training of a set of data while assuming high independence among the attributes. The complexity of the factors affecting clustering values necessitates the use of this high level of independence to simplify calculations.

2.7. Data Testing and Evaluation

Data testing was performed to verify the accuracy of the calculations and ensure the functionality of the applied methods. After manual calculations, the data were tested using the RapidMiner tools to confirm that the results were accurate in determining the clustering of the data. The research employed two types of clusters: cluster 0 (C0) for products with no potential requirements and cluster 1 (C1) for products with potential requirements. The initialization of cluster centers (K) was typically done randomly, assigning initial values to the cluster centers. The entire process, from data collection to testing, ensured a robust approach to categorizing production goods requirements. This systematic methodology provided a solid foundation for analyzing and optimizing inventory management, ultimately enhancing production efficiency and reducing operational costs.

3. Result and Discussion

3.1 Results

3.1.1 Data Analysis

The initial step in this research involved preparing the data to be processed, which consisted of stock data records. For the analysis of the K-Means algorithm through clustering methods, 50 data records were selected that had undergone pre-processing. This pre-processing stage included cleaning and organizing the data to ensure accuracy and relevance. The dataset for the analysis is detailed in Table 5.

Table 5. Stock Requirements Dataset

No	Required Goods Available	Required Goods Not Available
1	20	15
2	20	17
3	20	15
4	20	17
5	20	15
6	20	15
7	20	15
8	3	0
9	10	10
10	10	8

The complete data recording process that served as the sample for the calculations can be found in Appendix 1. This dataset was then used to model the clustering process using the K-Means algorithm.

3.1.2 Testing Results Using RapidMiner Studio

In this stage, the K-Means clustering method was applied to form clusters with a high level of accuracy. The research utilized the RapidMiner Studio application for testing and calculating the model. The testing process in RapidMiner Studio followed these steps:

1) Data Importation
   The required data was imported into RapidMiner Studio by selecting and clicking the "Import Data" option. The relevant attributes and labels were then defined for use in the clustering process.
2) Design Configuration
   The dataset was added to the process design screen. The "Set Role" function was used to define the roles of the attributes, which were then dragged into the process design layout.
3) Normalization
   The "Normalize" function was selected to standardize the data from the dataset used in this process. This step ensured that all data points were on a comparable scale, which is crucial for accurate clustering.
4) Modeling
   Within the "Segmentation" submenu, the K-Means function was chosen to apply the clustering algorithm. Here, the number of clusters (k) to be used in the data modeling was determined. The parameters used for clustering the stock data are illustrated in Figure 1.

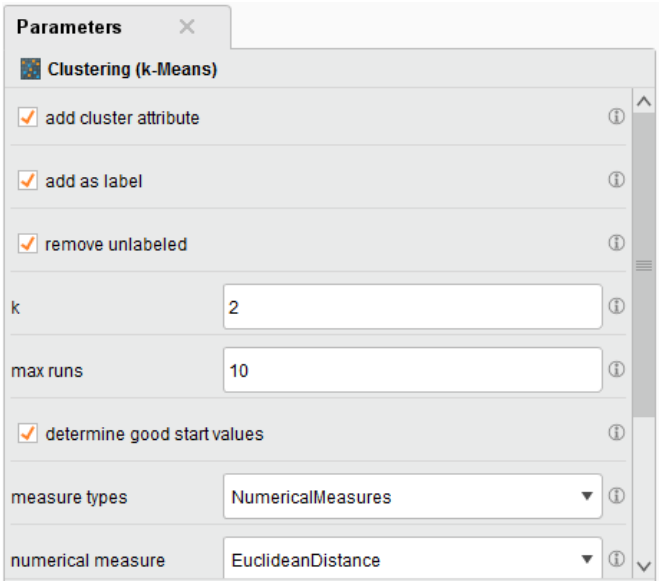

Figure 1. Clustering Parameters in RapidMiner Studio

5) Performance Evaluation
The "Performance" function was added to display the Davies-Bouldin Index (DBI) values obtained from the clustering process. This index is a metric for evaluating the quality of the clusters formed.
6) Process Configuration
All commands were connected to form a cohesive process flow, as shown in Figure 2.

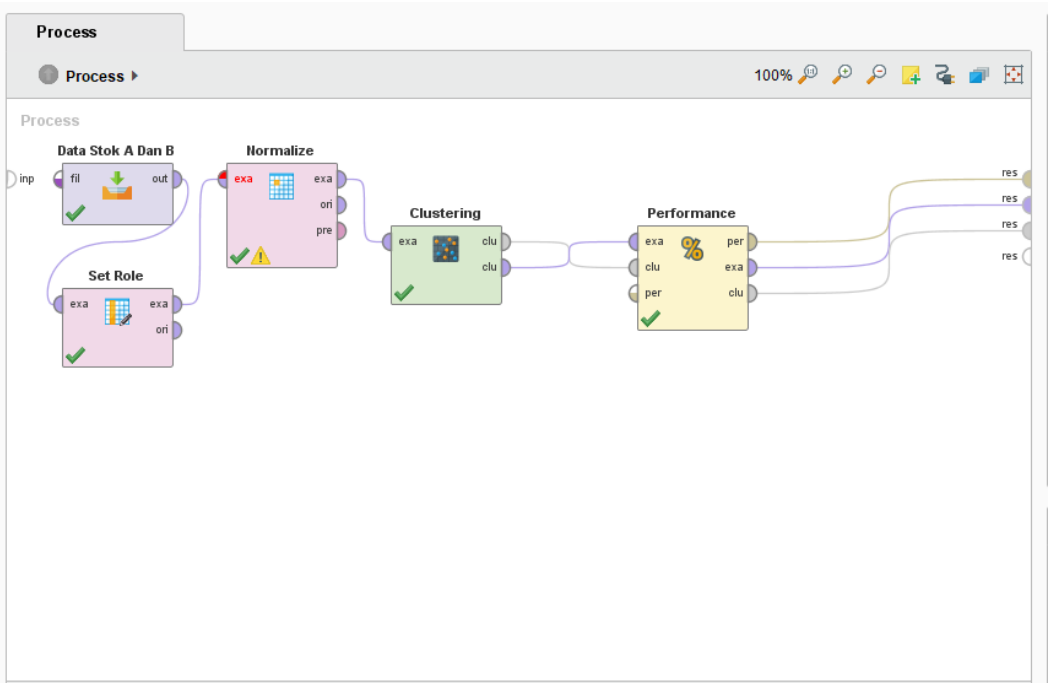

Figure 2. Clustering Process Flow in RapidMiner Studio

7) Execution: The process was executed by running the clustering analysis on the 50-record dataset. The results of this clustering process are summarized in the following sections.

3.1.3 Analysis of Testing Results

Upon completing the clustering process, the K-Means algorithm effectively grouped the data into clusters. The dataset of 50 stock records was analyzed, and the results are depicted in the cluster model generated by RapidMiner Studio. The analysis revealed that out of the 50 records, 15 were classified into the

first cluster (C0), and 35 were classified into the second cluster (C1). The cluster model can be summarized as follows: Cluster 0 consists of 15 items, Cluster 1 consists of 35 items, totaling 50 items.

I. CLUSTER MODEL

Cluster 0: 15 items
Cluster 1: 35 items
Total number of items: 50

Figure 3. Cluster Model Results in RapidMiner Studio

The clustering results obtained through RapidMiner were consistent with the manual calculations of the K-Means model. The members of each cluster matched closely with the manual clustering results. However, unlike the manual process, the initial cluster values were not predefined in RapidMiner Studio. The optimal cluster centroids for each variable were identified as 60.33 & 70.33 for Cluster 0 (C0) and 10.94 & 7.14 for Cluster 1 (C1), as shown in Figure 4.

cluster_0	cluster_1
30.400	23.400
60.333	10.943
70.333	7.114

Figure 4. Optimal Cluster Centroids in RapidMiner Studio

3.1.4 Performance Evaluation

The performance evaluation of the model and algorithm aimed to verify the accuracy of the calculations and the functionality of the methods used. This evaluation was conducted by comparing manual calculations with the automated results from RapidMiner Studio. The performance metrics included the average within-centroid distance for each cluster and the Davies-Bouldin Index (DBI), which is a measure of cluster separation and compactness. Performance Metrics; Average within centroid distance: -782.122, Average within centroid distance (Cluster 0): -1810.018, Average within centroid distance (Cluster 1): -341.595, Davies-Bouldin Index: -0.661.

PerformanceVector:
Avg. within centroid distance: -782.122
Avg. within centroid distance_cluster_0: -1810.018
Avg. within centroid distance_cluster_1: -341.595
Davies Bouldin: -0.661

Figure 5. Clustering Performance in RapidMiner Studio

The DBI value obtained from the evaluation was -0.661, indicating that the clustering model performed well in terms of forming distinct and compact clusters. This evaluation, as shown in Figure 6, confirms the effectiveness of the K-Means algorithm in clustering the stock data accurately.

Davies Bouldin

Davies Bouldin: -0.661

Figure 6. Davies-Bouldin Index Results in RapidMiner Studio

The application of the K-Means clustering method in this study successfully categorized the production goods requirements into meaningful clusters. The use of RapidMiner Studio facilitated an efficient and accurate analysis, aligning closely with manual calculations. The performance evaluation metrics further validated the

clustering results, demonstrating the algorithm's capability to enhance inventory management and production efficiency in manufacturing industries.

3.2 Discussion

The application of the K-Means clustering algorithm in this research aimed to optimize inventory management by categorizing production goods into distinct clusters. This process began with the preparation and pre-processing of a dataset consisting of 50 stock records, ensuring that the data were accurate and relevant for analysis. The initial analysis phase involved organizing the stock data into a structured format. This step was crucial in preparing the dataset for the clustering algorithm. By using 50 records, we ensured that the dataset was manageable while still providing sufficient data points for meaningful analysis. Each record included information on the availability of required goods and the presence of any unmet requirements. Once the dataset was prepared, the K-Means clustering algorithm was applied using the RapidMiner Studio application. The choice of RapidMiner Studio was strategic, given its robust capabilities for data analysis and its user-friendly interface. The process began with importing the data into RapidMiner Studio, where we defined the relevant attributes and labels for clustering. Normalization was a key step in this process. By standardizing the data, we ensured that all data points were on a comparable scale, which is essential for the accuracy of the clustering algorithm. The normalized data were then segmented using the K-Means function, where we specified the number of clusters (k) to be used in the analysis. The performance evaluation was integrated into the process using the Davies-Bouldin Index (DBI). The DBI is a metric for assessing the quality of clusters, measuring the separation between clusters and their compactness. This step was vital in validating the effectiveness of the clustering algorithm and ensuring that the resulting clusters were meaningful and useful for inventory management.

The clustering process revealed two distinct clusters within the dataset. Cluster 0 (C0) consisted of 15 items, while Cluster 1 (C1) comprised 35 items. This distribution indicated that the majority of the stock records fell into the second cluster, suggesting a pattern in the data that could be leveraged for better inventory management. The optimal centroids for the clusters were identified as 60.33 and 70.33 for Cluster 0, and 10.94 and 7.14 for Cluster 1. These centroids represent the central points around which the data points in each cluster are grouped. The identification of these centroids is crucial as they provide insights into the characteristics of each cluster, which can be used to inform inventory management strategies. The performance metrics further validated the clustering results. The average within-centroid distance was -782.122 overall, with -1810.018 for Cluster 0 and -341.595 for Cluster 1. The negative values of the Davies-Bouldin Index, particularly the DBI of -0.661, indicate that the clusters formed were distinct and well-defined, highlighting the effectiveness of the K-Means algorithm. The practical implications of these results are significant for inventory management in manufacturing industries. By categorizing production goods into clusters, companies can better understand their inventory needs and optimize their stock levels accordingly. For instance, items in Cluster 0 might represent goods with lower demand or issues in availability, while items in Cluster 1 could be high-demand goods that require more frequent restocking.

Moreover, the use of the K-Means algorithm allows for dynamic and data-driven decision-making. Companies can regularly update their clustering analysis as new data become available, ensuring that their inventory strategies remain aligned with current demand patterns and production needs. This adaptability is crucial in the fast-paced manufacturing sector, where market conditions and consumer preferences can change rapidly. The integration of advanced data analysis tools like RapidMiner Studio further enhances the capability of companies to manage their inventory efficiently. The visual and interactive nature of the tool allows users to easily interpret the clustering results and make informed decisions. Additionally, the ability to automate the clustering process reduces the time and effort required for manual analysis, leading to more efficient operations. The use of the K-Means clustering algorithm in this research has demonstrated its potential to significantly improve inventory management in manufacturing industries. The clear identification of distinct clusters within the stock data provides valuable insights that can be used to optimize inventory levels, reduce costs, and enhance production efficiency. The performance evaluation metrics, particularly the Davies-Bouldin Index, confirm the validity and reliability of the clustering results, underscoring the effectiveness of the K-Means algorithm in this application.

4. Related Work

The use of the K-Means clustering algorithm for inventory management and production optimization has been extensively researched across various domains. Ramdhan *et al*. (2022) investigated clustering inventory data using the K-Means method. Their study focused on categorizing inventory items based on demand patterns to enhance stock management in retail settings. The findings demonstrated that K-Means clustering effectively identified groups of products with similar demand characteristics, enabling more efficient inventory control and reducing storage costs [1]. This research underscores the potential of K-Means in addressing inventory management challenges, aligning closely with the objectives of the current study. Fikri Sallaby *et al*. (2022) explored the application of K-Means clustering for grouping products based on sales data in a retail store. Their findings indicated that the algorithm could segment products into distinct categories, allowing the store to tailor its inventory strategy to meet customer demand more effectively [2]. The study emphasized the importance of data preprocessing, including normalization and cleaning, to ensure accurate clustering results. This approach is integral to the methodology employed in the present research. Darmi and Setiawan (2016) examined the use of the K-Means clustering algorithm in grouping sales products, highlighting the challenges of dealing with large datasets and the necessity of robust data preprocessing techniques. Their findings showed that combining K-Means with effective data management practices significantly improved the accuracy of product classification and inventory forecasting [3]. This study provides valuable insights into the practical application of K-Means clustering, reinforcing its relevance to the current research.

In another study, applied K-Means clustering combined with RFM (Recency, Frequency, Monetary) analysis to segment customer loyalty. This study demonstrated the versatility of K-Means beyond inventory management, showcasing its applicability in customer relationship management and marketing analytics [4]. The integration of K-Means with other analytical methods offers a comprehensive approach to data segmentation, which is relevant to the multifaceted nature of inventory management explored in the current study. Further research by Hasanah *et al*. (2019) focused on the utilization of data mining for grouping purposes. Their work illustrated the broad applicability of clustering algorithms like K-Means in different domains, emphasizing the importance of selecting appropriate attributes and the impact of clustering on resource allocation and operational efficiency [6]. This study's insights into attribute selection and clustering impact inform the data preprocessing steps in the current research. Effendi *et al*. (2021) investigated the use of K-Means clustering to group productive land for palm oil plantations. Their application of the algorithm to agricultural data demonstrated its utility in optimizing resource allocation and improving operational efficiency [7]. The study's findings on the practical benefits of clustering in a real-world context are highly relevant to the current research, which aims to enhance inventory management efficiency through similar means. Afiasari *et al*. (2023) explored the implementation of K-Means clustering for analyzing sales transaction data in a retail store. Their research focused on identifying purchasing patterns to optimize inventory and marketing strategies. The study found that K-Means clustering provided valuable insights into customer purchasing behavior, which could be used to enhance inventory management and promotional efforts [8]. This research aligns with the goals of the current study, aiming to leverage clustering for better inventory decision-making.

Jabbar (2022) applied the K-Means clustering method to manage stock levels in a retail environment, focusing on identifying optimal stock levels for different product categories based on historical sales data. The results showed that K-Means clustering could accurately segment products, leading to improved stock replenishment strategies and reduced instances of stockouts and overstocking [18]. This practical application of K-Means for inventory optimization directly parallels the objectives of the present research. In the educational sector, Prastiwi *et al*. (2022) utilized K-Means clustering to predict student study durations. Their research demonstrated the algorithm's potential in educational settings, providing insights into student performance and helping educational institutions allocate resources more effectively [23]. This study highlights the versatility of K-Means clustering across various domains, reinforcing its applicability to inventory management. The body of work reviewed here underscores the versatility and effectiveness of the K-Means clustering algorithm in various applications, from inventory management and retail analytics to customer relationship management and agricultural planning. These studies collectively highlight the importance of robust data preprocessing, the selection of appropriate attributes, and the integration of advanced data analysis tools to achieve accurate and meaningful clustering results. The current study builds on these findings, applying the K-Means algorithm to optimize inventory management in the manufacturing sector and demonstrating its potential to enhance production efficiency and reduce operational costs. Ramdhan *et al*. (2022) emphasize the algorithm's efficiency in identifying demand patterns within inventory data, which directly influences inventory control and storage cost reduction [1]. Similarly, Fikri Sallaby *et al*. (2022) demonstrate how sales data segmentation through K-Means clustering can lead to more effective inventory

strategies tailored to customer demands [2]. These findings underscore the critical role of accurate data preprocessing to ensure the validity of clustering outcomes. Darmi and Setiawan (2016) highlight the importance of handling large datasets and employing robust preprocessing techniques to improve classification and forecasting accuracy [3]. This emphasis on data integrity and management is echoed by Hasanah *et al*. (2019), who underline the significance of selecting appropriate attributes for clustering to optimize resource allocation [6]. Effendi *et al*. (2021) showcase the practical applications of K-Means clustering in agriculture, particularly in optimizing resource allocation for palm oil plantations, demonstrating its broader utility beyond traditional inventory management [7]. Similarly, Afiasari *et al*. (2023) and Jabbar (2022) highlight the algorithm's applicability in retail environments, focusing on purchasing pattern analysis and optimal stock level identification [8][18].

In educational settings, Prastiwi *et al*. (2022) illustrate the potential of K-Means clustering in predicting student performance and aiding resource allocation, further extending the algorithm's application scope [23]. These diverse applications underline the algorithm's flexibility and effectiveness across various domains, reinforcing its relevance to the current study's objectives. The reviewed studies provide a comprehensive understanding of the K-Means algorithm's potential across different fields. The current research leverages these insights to apply K-Means clustering in optimizing inventory management within the manufacturing sector. By incorporating advanced data analysis tools and robust preprocessing techniques, this study aims to enhance production efficiency and reduce operational costs, building on the proven efficacy of the K-Means algorithm in diverse applications

5. Conclusion

The research findings provide significant insights into the application of the K-Means clustering algorithm for inventory management. The first cluster (C0), with centroid coordinates at (60.33, 70.33), represents stock data classified as having no potential. This cluster encompasses 35 data records from the stock process, indicating items that are not required for immediate production needs. The second cluster (C1), with centroid coordinates at (10.94, 7.11), represents stock data classified as available stock. This cluster includes 15 data records, signifying items that are currently needed for production. The application of the RapidMiner Studio software validated these findings, producing consistent clustering results. The analysis revealed that each cluster effectively grouped the stock data into two distinct categories, with optimal centroid values of 60.33 and 70.33 for Cluster 0 (C0) and 10.94 and 7.14 for Cluster 1 (C1). The Davies-Bouldin Index score of 0.666 indicates a high level of accuracy and validity in the clustering outcomes, demonstrating the robustness of the K-Means algorithm in categorizing production goods inventory. These results underscore the potential of K-Means clustering to enhance inventory management by accurately identifying and categorizing stock items based on their relevance to production needs. This approach facilitates more efficient inventory control, reducing the likelihood of overstocking or stockouts and thereby optimizing operational efficiency.

References

[1] Ramdhan, D., Dwilestari, G., Dana, R. D., Ajiz, A., & Kaslani, K. (2022). Clustering data persediaan barang dengan menggunakan metode K-Means. *Means (Media Informasi Analisis dan Sistem)*, 7(1), 1–9. https://doi.org/10.54367/Means.V7i1.1826

[2] Sallaby, A. F., Alinse, R. T., Sari, V. N., & Ramadani, T. (2022). Pengelompokan barang menggunakan metode K-Means clustering berdasarkan hasil penjualan di Toko Widya Bengkulu. *Jurnal Media Infotama*, 18(1), 2022.

[3] Darmi, Y., & Setiawan, A. (2016). Penerapan metode clustering K-Means dalam pengelompokan penjualan produk. *Jurnal Media Infotama Universitas Muhammadiyah Bengkulu*, 12(2), 148–157.

[4] Informatika, T. (2020). Pengelompokkan loyalitas pelanggan dengan menggunakan kombinasi RFM dan algoritma K-Means. *Jurnal Teknik Informatika*, 5(1), 7–13.

[5] Jabat, J. T., & Murdani, M. (2019). Penerapan Data Mining Pada Penjualan Produk Retail Menggunakan Metode Clustering. *Pelita Informatika: Informasi dan Informatika*, *8*(1), 26-32.

[6] Hasanah, H., Larasati, W., Komputer, F. I., Duta, U., Surakarta, B., & Clusterring, K. (2019). Pemanfaatan data mining untuk mengelompokkan. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 292–300.

[7] Effendi, H., Syahrial, A., Prayoga, S., & Hidayat, W. D. (2021). Penerapan metode K-Means clustering untuk pengelompokan lahan sawit produktif pada PT Kasih Agro Mandiri. *Teknomatika*, 11(2), 117–126.

[8] Afiasari, N., Suarna, N., & Rahaningsi, N. (2023). Implementasi data mining transaksi penjualan menggunakan algoritma clustering dengan metode K-Means. *Jurnal Saintekom*, 13(1), 100–110. https://doi.org/10.33020/Saintekom.V13i1.402

[9] Al Rasyid, H., Soebari, B. F. K., & Kartika, D. S. Y. (2022). Implementasi algoritma K-Means clustering untuk pengelompokan penjualan produk pada online shop Toko Gizi. *Prosiding Seminar Nasional Teknologi dan Sistem Informasi*, 2(1), 242–248. https://doi.org/10.33005/Sitasi.V2i1.304

[10] Suyanto. (2017). *Data Mining*. Yogyakarta: Informatika.

[11] Fakhriza, M. H., & Umam, K. (2021). Analisis produk terlaris menggunakan metode K-Means clustering pada PT. Sukanda Djaya. *Jurnal Informasi*, 5(1), 8. https://doi.org/10.31000/Jika.V5i1.3236

[12] Vulandari, R. T. (2017). *Data Mining*. Yogyakarta: Gava Media.

[13] Sani, A. (2018). Implementation of the K-Means clustering method in companies. *Jurnal Teknologi*, 8(1).

[14] K-Means, M. A., Hutabarat, S. M., & Sindar, A. (2019). Data mining penjualan suku cadang sepeda motor. *Jurnal Teknik Informatika*, 2(2), 126–132.

[15] Aisah, S., Aknuranda, I., & Rusydi, A. N. (2020). Sistem pendukung keputusan untuk pengelompokan barang terjual pada PT Dasema Digi Persada dengan metode K-Means clustering. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 4(7), 2309–2317.

[16] Oktaviani, S., & Bahtiar, A. (2023). Implementasi algoritma K-Means dalam pengelompokan data penjualan CV. Widuri menggunakan Orange. *Jurnal Wahana Informasi*, 2(1), 188–196.

[17] Sains, F., Teknologi, D. A. N., Islam, U., Sultan, N., & Kasim, S. (2024). Data persediaan barang menggunakan metode elbow dan K-Medoid tugas akhir.

[18] Jabbar, J. (2022). Sistem informasi stok barang menggunakan metode clustering K-Means (studi kasus RMD Store). *Infotech Journal*, 8(1), 70–75. https://doi.org/10.31949/Infotech.V8i1.2280

[19] Gunadi, G., & Sensuse, D. I. (2012). Penerapan metode data mining market basket analysis terhadap data penjualan produk buku dengan menggunakan algoritma apriori dan frequent pattern growth (FP-Growth). *Telematika*, 4(1), 118–132.

[20] Miftakhul, M., & Prihandoko, S. (2017). Penerapan algoritma K-Means dan CURE dalam menganalisa pola perubahan belanja dari retail ke e-commerce. *Jurnal Teknik Informatika*, 7(2), 44–49.

[21] Adani, N. F., et al. (2019). Implementasi data mining untuk pengelompokan data penjualan berdasarkan pola pembelian menggunakan algoritma K-Means clustering pada Toko Syihan. *Jurnal Cyber Tech*, 1–11.

[22] Villacampa, O. (2015). *Feature selection and classification methods for decision making: A comparative analysis*. ProQuest Dissertations & Theses, 63, 188.

[23] Prastiwi, H., Pricilia, J., & Rasywir, E. (2022). Implementasi data mining untuk menentukan persediaan stok barang di mini market menggunakan metode K-Means clustering. *Jurnal Informasi dan Rekayasa Komputer (JAKAKOM)*, 2(1), 141–148. https://doi.org/10.33998/Jakakom.2022.2.1.34

[24] Haryati, S., Sudarsono, A., & Suryana, E. (2015). Implementasi data mining untuk memprediksi masa studi mahasiswa menggunakan algoritma C4.5. *Jurnal Media Infotama*, 11(2), 130–138.

[25] Erdiansyah, A., & Devega, M. (2023). Pengelompokan barang menggunakan metode K-Means clustering dan K-Medoids berdasarkan hasil penjualan pada Kamajaya Seraya. *Prosiding Seminar Nasional Teknologi Informasi dan Ilmu Komputer*, 2(1), 35–41.

[26] Berliana, T. I., Budianita, E., Nazir, A., & Insani, F. (2023). Clustering data persediaan barang menggunakan metode elbow dan DBSCAN. *Jurnal Sistem Komputer dan Informasi*, 267(2), 258–267. https://doi.org/10.30865/Json.V5i2.7089

[27] Indriani, D., Irawan, B., & Bahtiar, A. (2024). Penerapan algoritma K-Means clustering untuk menentukan persediaan stok barang. *Jati (Jurnal Mahasiswa Teknik Informasi)*, 8(1), 182–187. https://doi.org/10.36040/Jati.V8i1.8322

[28] Zafira, F., Irawan, B., & Bahtiar, A. (2024). Penerapan data mining untuk estimasi stok barang dengan metode K-Means clustering. *Jati (Jurnal Mahasiswa Teknik Informasi)*, 8(1), 156–161. https://doi.org/10.36040/Jati.V8i1.8319.