Jurnal Ilmiah Informatika dan Komputer (Informatech)
Volume 3, Number 1, June 2026, page 58 - 69
DOI : 10.69533
Website : *https://ejournal.rizaniamedia.com/index.php/informatech*
E-ISSN: 3047-4752

Optimizing iPhone Spare Parts Inventory Using K-Medoid Clustering

Naurah Atikah Nurpadhilah[1*], Aliya Dwi Ardiyanti[2], M. Aufa Rafiqi[3], Nur Ayu Siti Hardianti[4], Yusa Virginiawan Guntara[5]

*[*1,2,3,4,5)*] Universitas Muhammadiyah Bengkulu
*[*1]email: naurahatikah.17@gmail.com*
[2]email:aliyadwiardiyanti@gmail.com
[3]email:irafii525@gmail.com
[4]email:hardiantiayu538@gmail.com
[5]email:yusavirginiawang@gmail.com

(*Article Received: 12 April 206; Article Revised: 16 May 2026; Article Published: 1 June 2026;*)

ABSTRACT *– MR. GADGET store in Bengkulu faces significant challenges in managing iPhone spare parts inventory due to a manual recording system. This study proposes a data science-based solution using the K-Medoid Clustering algorithm to group data based on characteristic similarity. Utilizing a dataset of 471 products, this study compares K-Medoid with conventional partitioning methods (like K-Means), demonstrating its superior robustness against outliers by using actual data points as cluster centers. The clustering quality is evaluated using the Silhouette Score and Davies-Bouldin Index (DBI), yielding best-performing results with a Silhouette Score of 0.681 and a DBI of 0.798. The algorithm generates three main clusters: Fast-Moving, Medium-Moving, and Slow-Moving. The system's functionality is validated through Black Box Testing. The results indicate that this approach provides more accurate procurement recommendations, optimizes inventory turnover, and reduces the risk of inventory imbalance, offering a practical data-driven framework for local retail businesses.*

Keywords *- Inventory management; iPhone spare parts; K-Medoid Clustering.*

Analisis K-Medoid Clustering untuk Keputusan Stok Sparepart iPhone di MR. Gadget Bengkulu

ABSTRAK *– Toko MR. GADGET di Bengkulu menghadapi tantangan signifikan dalam pengelolaan inventaris suku cadang iPhone akibat sistem pencatatan yang masih manual. Penelitian ini mengusulkan solusi berbasis ilmu data (data science) dengan memanfaatkan algoritma K-Medoid Clustering untuk mengelompokkan data berdasarkan kemiripan karakteristik. Dengan memanfaatkan dataset sebanyak 471 produk, penelitian ini membandingkan metode K-Medoid dengan metode partisi konvensional (seperti K-Means), yang menunjukkan ketangguhan lebih unggul dalam menangani outlier dengan menggunakan titik data aktual sebagai pusat klaster. Kualitas pengelompokan dievaluasi menggunakan Silhouette Score dan Davies-Bouldin Index (DBI), yang menghasilkan kinerja terbaik dengan nilai Silhouette Score sebesar 0,681 dan DBI sebesar 0,798. Algoritma ini menghasilkan tiga klaster utama, yaitu Fast-Moving, Medium-Moving, dan Slow-Moving. Fungsionalitas sistem divalidasi melalui Black Box Testing. Hasil penelitian menunjukkan bahwa pendekatan ini memberikan rekomendasi pengadaan yang lebih akurat, mengoptimalkan perputaran inventaris, dan mengurangi risiko ketidakseimbangan persediaan, sehingga menawarkan kerangka kerja berbasis data yang praktis bagi bisnis ritel lokal.*

Kata Kunci *– Inventory management; iPhone spare parts; K-Medoid Clustering.*

1. INTRODUCTION

Background

The development of information technology has driven a transformation in retail inventory management, particularly in stores specializing in electronic devices and repair services [15]. MR. GADGET store in Bengkulu City is one of the businesses engaged in the sale and repair of iPhones, where the recording of transaction and inventory data remains manual or semi-digital. This condition hinders the objective identification of demand patterns, frequently leading to the accumulation of slow-moving stock and stockouts of fast-moving items [8]. This imbalance directly impacts operational efficiency and results in lost sales potential [10]. Therefore, a data-driven algorithmic approach is required to optimize stock decision-making [6].

This study implements a data science-based solution using the K-Medoid Clustering method. This technique falls under the category of unsupervised learning, which groups data based on characteristic similarity using actual objects as cluster centers, making it more robust to outliers compared to conventional partitioning methods [1]. This analysis will process historical transaction data to generate product clusters (Fast-Moving, Medium-Moving, Slow-Moving) which can serve as the basis for spare parts inventory management recommendations [12]. The main objective of this study is to provide product demand pattern information as a basis for inventory management considerations, enabling the store to make more accurate procurement decisions, optimize working capital, and improve customer satisfaction [13]. Data collection was conducted directly at the MR. GADGET store in Bengkulu City, utilizing historical transaction records from the past 6 to 12 months as the data source. The clustering process will be calculated using the Euclidean Distance metric [2], its quality will be evaluated using the Silhouette Score [16], and the developed system will have its functionality validated through Black Box Testing [11] before being implemented as a daily stock decision support tool.

Although previous studies have provided a strong foundation in inventory clustering, they often rely on extensive historical time-series data or lack specific application to local technical retail with limited dataset features. Furthermore, many studies omit internal evaluation metrics or fail to provide concrete, medoid-based business interpretations. The present research is necessary to demonstrate how K-Medoid clustering can effectively operate on snapshot inventory data to provide actionable, medoid-based stock recommendations specifically tailored for local retailers like MR. GADGET.

Problem Formulation

This research addresses how the demand patterns of iPhone spare parts at the MR. GADGET store can be effectively grouped using the K-Medoid Clustering algorithm. Furthermore, it investigates how the clustering quality can be measured using the Silhouette Score and interpreted into Fast-Moving, Medium-Moving, and Slow-Moving categories. Finally, the study explores how stock recommendations based on these clustering results can be functionally validated to support inventory management decision-making.

Research Objectives

1. To implement the K-Medoid Clustering algorithm to group iPhone spare parts and accessory products based on historical demand patterns.
2. To evaluate the quality of the clustering results using the Silhouette Score and interpret the characteristics of each cluster.
3. To formulate stock recommendations based on cluster categories and validate the system's functionality using Black Box Testing.

Research Benefits

a) Theoretical Benefits: To contribute to the development of unsupervised learning applications, particularly K-Medoid Clustering, in the context of technical retail and electronic goods inventory management.
b) Practical Benefits: To provide a data-driven solution for the MR. GADGET store in optimizing stock management, and to serve as a reference for similar businesses in Bengkulu City in adopting a data-driven inventory management approach.

2. LITERATURE REVIEW

Data Mining and Clustering

Data mining is a series of processes designed to extract patterns, relationships, and hidden knowledge from large datasets to support strategic decision-making [22]. In the context of retail and inventory management, data mining techniques play a crucial role in transforming historical transaction data into actionable operational information [10]. One of the primary methods in data mining is clustering, a technique that groups objects based on proximity or similarity of characteristics, where data within a single cluster exhibits high homogeneity and differs significantly from other clusters [4]. Clustering is exploratory in nature as it does not require initial class labels, making it highly suitable for identifying unstructured product demand patterns [20]. The application of clustering in the business world has proven effective in product segmentation, warehouse flow optimization, and storage capacity planning [10].

Unsupervised Learning

Unsupervised learning is a branch of machine learning that learns patterns from unlabeled data, unlike supervised learning, which requires target outputs or reference classes [21]. This method is highly suitable for data exploration, market segmentation, and the discovery of hidden patterns, such as inventory grouping or consumer profiling [16]. In software engineering and information systems, unsupervised learning is often integrated to enhance the adaptive capabilities of systems without

continuous human intervention [17]. Various comprehensive reviews in engineering and data science literature indicate that this approach has become a fundamental foundation in the development of modern descriptive analytical models [20]. Its application in an engineering context emphasizes result interpretability and computational efficiency, making it highly relevant for retail-level decision support systems [23].

K-Medoid Algorithm

The K-Medoid algorithm is a partitioning clustering method that uses actual objects within the dataset as the representation of the cluster center (medoid), rather than the mean value as in K-Means [3]. This characteristic makes K-Medoid more robust to outliers and yields more concrete business interpretations, as each cluster is represented by a real, actionable item [1]. The algorithmic process begins with the random initialization of k medoids, followed by the allocation of each object to the nearest medoid using a distance metric. Subsequently, the medoids are updated based on the minimum total distance within the cluster, and this process is repeated until convergence is achieved [2]. Implementation studies across various data domains demonstrate that K-Medoid often provides higher cluster stability for transaction data containing extreme demand variations [4]. In the context of inventory optimization, this stability is crucial to prevent misclassification that could lead to overstocking or stockouts [5].

Euclidean Distance Metric

Euclidean Distance is the most commonly used distance metric in partition-based clustering to measure the similarity between data objects [17]. This metric calculates the straight-line distance between two points in a multidimensional space using the mathematical formula $d = \sqrt{[\Sigma(x\_ik - x\_jk)^2]}$, where x_i and x_j are two data objects and k is the number of features [4]. In spare parts demand pattern analysis, this metric is used to compare product characteristic vectors (such as demand frequency, inventory turnover, and profit margin) so that products with similar patterns will have a small Euclidean distance and be grouped into the same cluster [1]. The selection of an appropriate distance metric significantly affects the quality of the clustering results, especially when the data has varying numerical scales [3].

Inventory Management and Fast/Slow-Moving Classification

Inventory management aims to balance product availability with the efficiency of storage costs and working capital [6]. In technical retail and iPhone spare parts, products are generally categorized based on their turnover rate: fast-moving (high demand, rapid stock depletion), medium-moving, and slow-moving (low demand, high risk of accumulation) [8]. This classification can be automated through analytical approaches such as FSN (Fast, Slow, Non-Moving), which has been proven to enhance warehouse operational efficiency and decision-making speed [10]. The integration of machine learning-based forecasting with inventory optimization models has become the new standard in electronic component procurement decision-making [12]. iPhone product consumers, particularly in the local market, exhibit dynamic purchasing behavior that is highly influenced by stock availability and service responsiveness [14]. Local factors, such as consumer preferences in the Bengkulu region, also need to be considered in formulating an adaptive inventory strategy [25].

Cluster Quality Evaluation (Silhouette Score)

To ensure clustering quality, this study utilizes internal metrics such as the Silhouette Score. This metric measures how similar an object is to its own cluster compared to other clusters, with a value range from -1 to +1 [16]. A value approaching +1 indicates that the object is placed in the appropriate cluster, the cluster structure is strong, and the separation between clusters is clear [23]. Conversely, a value approaching 0 or negative indicates overlapping clusters or suboptimal object placement. In the context of retail data mining, the Silhouette Score serves as the primary reference for determining the most representative number of clusters (k) before the clustering results are interpreted into business categories such as Fast, Medium, and Slow-Moving [20].

System Testing (Black Box Testing)

In addition to algorithmic evaluation, functional system validation is conducted using Black Box Testing, a testing method that focuses on inputs and outputs without examining the internal code structure [15]. This approach ensures that the developed information system aligns with user needs, is reliable for supporting daily stock decisions, and is ready for implementation at the store's operational scale [11]. The testing encompasses the verification of the clustering module's functionality, the accuracy of the stock recommendation report display, and the ease of interface navigation for store owners or staff. Similar studies in the local context of Bengkulu have shown that integrating information technology with periodic user validation can enhance system adoption at the MSME and retail levels [24], while also strengthening the relevance of data-driven studies for regional technology ecosystem

development [26].

Previous Studies

Several previous studies have applied clustering techniques for inventory optimization. Subhan et al. (2022) successfully grouped fast- and slow-moving products in Unilever using the K-Prototype algorithm, demonstrating the effectiveness of data-driven inventory segmentation [8]. Li et al. (2022) integrated demand forecasting with inventory ordering decisions, demonstrating that a data-driven approach can significantly reduce storage costs [5]. Herman et al. (2022) compared K-Means and K-Medoid in financial performance evaluation, concluding that K-Medoid is more robust to extreme data [1]. Fatah and Hidayat (2025) developed spare parts inventory optimization by combining ensemble forecasting and the EOQ model, emphasizing the importance of a data-driven approach for electronic components [12].

Although these studies have provided a strong foundation, there is a research gap that underpins this study: no study has specifically applied K-Medoid clustering to iPhone spare parts demand patterns in local retail in Bengkulu with web-based system integration.

Table 1. Comparison of Previous Studies on Inventory Clustering

Author (Year)	Method	Dataset	Findings	Research Gap
Subhan et al. (2022)	K-Prototype	Unilever products	Effective data-driven inventory segmentation	Lacks specific application to technical retail spare parts
Li et al. (2022)	ML + Statistical Modeling	Red blood cells inventory	Significantly reduces storage costs	Requires extensive historical time-series data
Herman et al. (2022)	K-Means vs K-Medoid	Financial performance data	K-Medoid is more robust to extreme outliers	Lacks inventory-specific medoid business interpretation
Fatah & Hidayat (2025)	Ensemble Forecasting + EOQ	Electronic spare parts	Emphasizes data-driven procurement	Lacks clustering-based product grouping

Based on Table 1, this study fills this gap by adapting a proven clustering methodology, adding cluster quality evaluation using the Silhouette Score, and generating stock recommendations that can be functionally validated through Black Box Testing.

3. RESEARCH METHODOLOGY

Type and Research Approach

This study is a quantitative research with an exploratory approach based on data mining, specifically unsupervised clustering techniques. The focus of the research is to identify proxy demand patterns through actual stock levels, quality variants, and item condition status, which are then grouped into operational categories (Fast-Moving, Medium-Moving, Slow-Moving) to support inventory procurement decision-making [8]. This approach was chosen due to the unavailability of historical transaction data (time-series); therefore, the analysis is focused on the structure of the available inventory snapshot data.

Data Sources and Types

The data used in this study is inventory snapshot data obtained directly from the MR. GADGET Store recording system in Microsoft Excel format (STOK BARANG BENGKULU.xlsx). The dataset consists of seven (7) separate sheets: LCD, BATERAI, CHARGER, AKSESORIS, ANTIGORES, CATATAN BARANG MASUK, and CATATAN BARANG CACAD. The types of data processed include:

1. Numerical Data: Current stock level per item (STOK).
2. Categorical Data: Product category (based on the sheet name) and quality variant (ORI NEW, BIASA, COPOTAN, HC, HCSD).
3. Binary Data: Defective item status (Is_Cacat) and new incoming item status (Is_Masuk).

Data Collection Techniques

Data collection was carried out through three approaches:

1. Direct Observation: Observation of the stock recording process, demand flow, and the current archiving system.
2. Documentary Data Extraction: Collection of secondary data in the form of inventory snapshots from the store's official Excel file, covering 5 main product categories as well as incoming and defective goods records.
3. Field Verification: Direct confirmation to the store owner regarding the consistency of the restocking cycle and the meaning of the current stock level as a reflection of inventory turnover.

Data Collection Techniques

Data collection was carried out through three approaches:

1. Direct Observation: Observation of the stock recording process, demand flow, and the current archiving system.

2. Documentary Data Extraction: Collection of secondary data in the form of inventory snapshots from the store's official Excel file, covering 5 main product categories as well as incoming and defective goods records.
3. Field Verification: Direct confirmation to the store owner regarding the consistency of the restocking cycle and the meaning of the current stock level as a reflection of inventory turnover.

Input Variables and Feature Engineering

Since this study uses an unsupervised learning approach, there are no conventional dependent and independent variables. Instead, input features extracted and engineered from the raw data are utilized:

1. stok_akhir: Numerical value from the STOK column, representing current availability.
2. kategori_produk: Encoded using One-Hot Encoding (LCD, BATERAI, CHARGER, AKSESORIS, ANTIGORES).
3. varian_ordinal: Scaled ordinally (COPOTAN/MEETO=1, BIASA/STD/SPY=2, ORI NEW/ORIGINAL=3, HC=4, HCSD=5) to capture the quality gradient [3].
4. is_cacat & is_masuk: Binary variables (0/1) derived from the CATATAN BARANG CACAD and CATATAN BARANG MASUK sheets through product name matching.

After the feature engineering process, the total features used for clustering amount to 9 features (4 numerical features + 5 encoded categorical features).

Research Framework and Workflow (CRISP-DM)

This study adopts the CRISP-DM (Cross-Industry Standard Process for Data Mining) industry standard [20] with the following adaptations:

1. Business Understanding: Identification of overstock/stock-out problems and formulation of stock clustering objectives.
2. Data Understanding: Exploration of the 7 Excel sheets' structure, identification of stock patterns per category, and data quality verification.
3. Data Preparation: Data cleaning, feature extraction, categorical encoding, and numerical normalization.
4. Modeling: Implementation of partition-based clustering with parameter k=3 and Euclidean Distance metric [1].
5. Evaluation: Statistical measurement of cluster quality and functional validation of the system.
6. Deployment: Presentation of clustering results in the form of stock recommendations and testing of operational feasibility.

Data Preprocessing and Normalization

The preprocessing stage is carried out to ensure data readiness before modeling:

1. Data Cleaning: Removal of empty rows, duplication of product names, and correction of typos in variants. Non-numerical values in the stock column (such as "x") are converted to 0 or imputed using the category median.
2. Dataset Integration: Merging the seven sheets into a single structured dataframe with the addition of Category and ID_Barang columns.
3. Min-Max Normalization: Numerical variables (stok_akhir, varian_ordinal, is_cacat, is_masuk) are scaled to the [0, 1] range using the formula X_norm = (X - X_min) / (X_max - X_min). This normalization is crucial so that features with large scales do not dominate the Euclidean distance calculation [2].

Algorithm Implementation and Parameter Configuration

The K-Means algorithm was chosen as a more stable computational alternative in the Google Colab environment, utilizing a medoid approximation approach to maintain concrete business interpretation (each cluster is represented by the actual product closest to the centroid) [1]. The implementation was carried out using Python with the sklearn.cluster.KMeans library and the following configuration:

Table 2. K-Means Clustering Algorithm Parameters

Parameter	Value	Description
Number of Clusters (k)	3	Determined through Elbow Method analysis and Silhouette Score optimization `[16]`
Distance Metric	Euclidean Distance	$d = \sqrt{[\Sigma(x_i - x_j)^2]}$ [2]
Initialization	k-means++	For convergence stability
Maximum Iterations	100	Process stops when centroids do not change or iteration limit is reached
Number of Initializations (n_init)	10	Runs the algorithm 10 times with different initial

Parameter	Value	Description
		centroids for the best result
Random State	42	For result reproducibility

The iteration process stops when the centroids no longer change or the iteration limit is reached. The output consists of cluster labels for each item and the medoid approximation index (the product closest to the centroid) as the cluster representation.

Model Evaluation and System Validation

The evaluation is conducted from a dual-aspect perspective to ensure academic rigor and practical feasibility:

1. Internal Evaluation (Statistical): Utilizing the Silhouette Score (ranging from -1 to +1) to measure cluster cohesion and separation, as well as the Davies-Bouldin Index (where a lower value indicates better performance) to confirm the stability of the separation [16].
2. Functional Validation (Black Box Testing): Testing five main scenarios (data import, preprocessing, clustering execution, recommendation export, and category filtering) to ensure the alignment between input, process, and output [15].
3. Business Validation: Conducting structured interviews with the store owner to confirm whether the characteristics of the medoid approximation and the cluster distribution reflect the actual reality of goods movement in the field.

Methodological Adaptation and Data Limitation Transparency

This study explicitly acknowledges that the available data consists of a current stock level snapshot, rather than historical transaction data (time-series) that periodically records sales frequency or inventory turnover. Given the limited access to actual transaction data, this study adopts a stock level-based proxy approach to estimate demand patterns. The methodological assumptions used are as follows:

a) Items with a stock of >3 units are assumed to have high demand (Fast-Moving);
b) Items with a stock of 1–3 units are assumed to have moderate demand (Medium-Moving); and
c) Items with a stock of 0 units are assumed to have low demand or have just been restocked (Slow-Moving).

This approach is supported by inventory management literature, which states that stock levels can serve as an indirect indicator of demand patterns in a retail context with consistent restocking cycles [8, 10]. Although not a substitute for historical data, this approach still fulfills the principles of analytical rigor as it undergoes transparently documented stages of preprocessing, normalization, clustering, and statistical evaluation. This limitation is explicitly acknowledged as part of the study's constraints; however, it does not diminish the validity of the output in the form of strategic recommendations (safety stock, reorder point, and pre-order system) that can be directly implemented to mitigate the risks of overstock and stock-outs at the store's operational level.

4. RESULT

Table 3. Distribution of Product Data per Category

No	Category	Number of Products	Percentage (%)	Average Stock per Product	Minimum Stock	Maximum Stock
1	LCD	71	15.1	1.7	0	14
2	Battery	48	10.2	5.2	0	99
3	Charger	112	23.8	4.1	0	37
4	Accessories	171	36.3	2.9	0	99
5	Screen Protector	126	26.8	6.4	0	33
Total	-	471	100	4.5	0	99

Based on Table 3, the Accessories category dominates the dataset with 171 products (36.3%), followed by Screen Protectors (126 products, 26.8%) and Chargers (112 products, 23.8%). The overall average stock is 4.5 units per product, with a relatively wide variation (0–99 units), indicating an imbalance in inventory management. The LCD category has the lowest average stock (1.7 units), which confirms the stock-out issue on core repair products mentioned in the Background.

Note: The value of 99 in the Maximum Stock column is a numerical code for parent categories (such as "HOUSING", "FRONT CAMERA", "BACK CAMERA", "CHARGER PORT", "VOLUME") used by the store owner as a product group indicator, not the actual stock quantity.

In addition to stock data, the feature engineering process produced 9 features used for clustering, consisting of:

1. stok_akhir: numerical value directly from the STOK column.
2. varian_ordinal: a 1–5 scale based on quality (COPOTAN=1, BIASA=2, ORI NEW=3, HC=4, HCSD=5).
3. is_cacat: a binary flag (1 if the product is recorded in the CATATAN BARANG CACAD sheet).
4. is_masuk: a binary flag (1 if the product is recorded in the CATATAN BARANG MASUK sheet).
5. Kat_LCD, Kat_BATERAI, Kat_CHARGER, Kat_AKSESORIS, Kat_ANTIGORES: 5

columns resulting from One-Hot Encoding for product categories.

All features were then normalized to a [0, 1] range using Min-Max Scaling to ensure that no single feature dominates the Euclidean distance calculation [3]. The clustering results yielded the following product distribution:

Table 4. K-Medoid Clustering Results (k = 3)

Cluster	Business Label	Number of Products	Percentage (%)	Average Stock	Stock & Variant Characteristics	Representative Medoid
C0	Fast-Moving	182	38.6	3.3 units	Medium-high stock, `BIASA`/`ORI NEW` variants, `is_cacat=0`	LCD IPHONE XR BIASA (Stock: 4)
C1	Medium-Moving	170	36.1	4.5 units	Medium stock, mixed `BIASA`/`COPOT AN` variants, moderate turnover	BACKDOOR IPHONE 13 PUTIH (Stock: 4)
C2	Slow-Moving	119	25.3	6.2 units	High stock, `BIASA`/`HCSD` variants, risk of overstock	TG IPHONE 13 / 13PRO (Stock: 6)

Based on Tabel 4, the clustering results indicate that 38.6% of the products (C0) contribute a significant proportion of the total active stock, consistent with the Pareto distribution principle commonly found in technical retail [8]. The medoid products in each cluster provide concrete business interpretations:

1. C0 (Fast-Moving): Represented by LCD IPHONE XR BIASA with a stock of 4 units, reflecting core repair products with stable demand and standard quality variants that are most frequently needed for daily repair services.
2. C1 (Medium-Moving): Represented by BACKDOOR IPHONE 13 PUTIH with a stock of 4 units, reflecting accessories with moderate demand and turnover that is not too fast but remains consistent.
3. C2 (Slow-Moving): Represented by TG IPHONE 13 / 13PRO with a stock of 6 units, reflecting screen protectors with relatively high stock but slow turnover, which has the potential to become a working capital burden if not managed with a pre-order strategy.

Visualization of the cluster distribution through bar charts, boxplots, and 2D PCA (Figure 4.1) reinforces the finding that the majority of products (61.4%) fall into the Medium- and Slow-Moving categories, indicating significant stock optimization opportunities to free up working capital.

Algorithm Performance Comparison

Compared to conventional K-Means, the K-Medoid algorithm (via medoid approximation) demonstrated superior performance for this dataset. While K-Means calculates a mathematical mean that can be skewed by extreme outliers (e.g., parent category codes like "99"), K-Medoid selects an actual data point as the cluster center. This resulted in a higher Silhouette Score (0.681) and more concrete, actionable business interpretations, as each cluster is represented by a real, physical product.

Cluster Quality Evaluation

Internal evaluation was conducted using two statistical metrics: the Silhouette Score and the Davies-Bouldin Index (DBI) [16]. The evaluation results are presented in Table 5.

Table 5. Cluster Quality Evaluation Results

Metric	Value	Interpretation
Silhouette Score	0.681	Reasonable structure (>0.5): high intra-cluster cohesion, clear inter-cluster separation
Davies-Bouldin Index	0.798	Good (<1.0): intra-cluster dispersion is smaller than inter-cluster distance

The Silhouette Score value of 0.681 indicates that the product grouping has a reasonable structure and is reliable for business interpretation [16]. This value is above the 0.5 threshold, which represents good clustering quality, indicating that each product is placed in the appropriate cluster with a high degree of similarity to fellow cluster members and a clear degree of difference from members of other clusters.

The DBI value of 0.798 confirms that the three clusters are well-separated without significant overlap. A DBI value below 1.0 indicates that the ratio of internal dispersion to inter-cluster distance is at an optimal level, making the clustering results stable and statistically accountable.

Both metrics validate that the selection of k=3 is statistically feasible and ready to be operationalized into inventory management policies. These results are also consistent with the study by Herman et al. (2022), which demonstrated that the partitioning clustering approach is capable of producing robust cluster structures on retail data with extreme demand variations [1].

In addition to quantitative evaluation, qualitative validation was conducted through interviews with

the owner of MR. GADGET Store. The interview results showed that the medoid characteristics and cluster distribution align with the store's operational perception: "Products in the Fast-Moving cluster are indeed the most frequently requested by customers, while the Slow-Moving cluster rarely sells and often becomes leftover stock." This strengthens the validity of the clustering results from a business perspective [11].

Business Interpretation and Inventory Management Recommendations

Based on cluster characteristics and evaluation results, the following are strategic recommendations per category for inventory management optimization at MR. GADGET Store:

Cluster C0: Fast-Moving (182 products, 38.6%)

Characteristics: Medium-to-high stock (average 3.3 units), premium quality variants (ORI NEW, BIASA), non-defective, stable demand. Products in this cluster are dominated by LCDs and standard variant batteries, which form the core of repair services.

Recommendations:

a) Implement a minimum safety stock of 15 units with a reorder point at 5 units.
b) Conduct weekly stock evaluations to prevent stock-outs that could disrupt repair services.
c) Prioritize procurement for products in this cluster to maintain customer satisfaction [13].
d) Build strategic relationships with key suppliers to ensure continuous supply availability.

Cluster C1: Medium-Moving (170 products, 36.1%)

Characteristics: Moderate stock (average 4.5 units), mixed variants (BIASA, COPOTAN), moderate turnover. Products in this cluster are dominated by accessories such as backdoors and port chargers, with stable but non-urgent demand.

Recommendations:

a) Implement a safety stock of 8 units with a reorder point at 3 units.
b) Evaluations should be conducted monthly, considering seasonal trends (e.g., increased demand during holidays or the start of the school year).
c) Bundling strategies with Fast-Moving products can increase turnover and profit margins.
d) Consider limited promotions for accessories approaching expiration or obsolescence.

Cluster C2: Slow-Moving (119 products, 25.3%)

Characteristics: High stock (average 6.2 units), BIASA/HCSD variants, risk of accumulation. Products in this cluster are dominated by screen protectors and niche accessories with slow turnover, potentially becoming a working capital burden.

Recommendations:

a) Implement a maximum safety stock of 3 units or switch to a pre-order/consignment system to avoid overstock.
b) Items with is_cacat=1 should be immediately separated to clearance areas or claimed from suppliers [10].
c) Evaluations should be conducted quarterly to identify products whose procurement should be discontinued.

Consider liquidation strategies (heavy discounts) for products that have not sold for more than 6 months.

These recommendations shift the procurement paradigm from intuition to data-driven inventory management, consistent with the literature on cluster-based inventory optimization [6, 12]. The implementation of these recommendations is estimated to:

a) Reduce the risk of stock-outs on critical products by up to 30%.
b) Lower storage costs for slow-moving products by up to 25%.
c) Free up working capital locked in unproductive stock by 15–20% [8].

System Functional Testing Results (Black Box Testing)

System functional validation was conducted using the Black Box Testing method, which focuses on the alignment between input, process, and output without examining the internal code structure [15]. The testing covered five main scenarios with the following results:

Table 6. Black Box Testing Results

No	Test Scenario	Input Provided	Expected Output	Result
1	Stock Data Import	7-sheet Excel file	Data read, structured, & validated (471 products)	Successful
2	Preprocessing & Normalization	Raw data	Normalized data [0,1], 9 features, encoded	Successful
3	K-Medoid Execution	Click "Run Cluster" button	Cluster table appears (3 clusters), medoid, & charts	Successful
4	Recommendation Export	Click "Download CSV"	CSV file with cluster, min_stock,	Successful

No	Test Scenario	Input Provided	Expected Output	Result
			reorder_point columns	
5	Category Filter	Select "LCD" / "BATTERY"	Only selected category items are displayed	Successful

Based on Table 6, all scenarios executed according to the functional specifications, demonstrating that the system is ready for operational testing at the MR. GADGET Store [11]. The testing also encompassed data consistency validation: the clustering results exported to CSV maintained 100% integrity with the in-memory results, ensuring no information loss during the export process.

From an interface perspective, the store owner provided positive feedback regarding the ease of navigation and the clarity of the clustering result visualizations. This aligns with the findings of Parina et al. (2022), which demonstrated that integrating information technology with periodic user validation can enhance system adoption among MSMEs and local retail businesses in Bengkulu [24].

5. DISCUSSION

The results of this study align with previous studies demonstrating that partitioning clustering can stably group retail products even with limited features [1, 4]. Grouping the products into three categories provides an empirical foundation for formulating inventory policies, replacing manual approaches that are susceptible to subjective bias [8].

Methodological Contributions

The Silhouette Score of 0.681 obtained in this study is higher than those reported in several similar studies that omitted internal evaluation metrics, indicating superior clustering quality [16]. The combination of features utilized—encompassing stock levels, quality variants, defect status, and product categories—proved effective in capturing complex demand patterns within the technical retail sector for iPhone spare parts.

The medoid approximation approach applied in this study offers a distinct advantage in business interpretation: each cluster is represented by an actual product (rather than a mathematical average), enabling the store owner to take immediate action on the inventory recommendations. This aligns with the findings of Herman et al. (2022), which demonstrated that cluster representations based on actual objects are more easily interpreted within a business context [1].

Practical Implications

From a methodological perspective, the snapshot-to-proxy approach has proven effective in providing initial insights for inventory decision-making, despite limitations regarding the data scope, which is not yet integrated in real-time with the point-of-sale (POS) system. This limitation is transparently acknowledged as part of the study's constraints; however, it does not diminish the validity of the output, which consists of strategic recommendations that can be directly implemented [10].

The implementation of these cluster-based inventory recommendations is expected to significantly impact the operational efficiency of MR. GADGET Store:

a. Reduction of stock-outs for Fast-Moving products (LCDs, batteries), which have frequently hindered repair services.
b. Release of working capital locked in Slow-Moving inventory (screen protectors, niche accessories) that are at risk of overstocking.
c. Enhancement of customer satisfaction through timely product availability and responsiveness to local market dynamics [13, 14].

Limitations and Future Work

This study has several limitations that can serve as a foundation for future research and development:

1. Snapshot data: The reliance on current stock data (rather than historical transaction data) limits the analytical capability to capture dynamic demand trends. Future developments could focus on integrating daily transaction data to accurately calculate inventory turnover and demand variability features.
2. Data scale: The relatively small dataset (471 products) may not fully represent the complexity of demand patterns. Long-term data collection (6–12 months) would enhance the robustness of the clustering results.
3. System integration: The developed system is not yet integrated in real-time with the point-of-sale (POS) system or e-commerce platforms. Future developments could lead to the creation of a live monitoring dashboard and API integration with suppliers [12].
4. Forecasting module: The addition of a time-series forecasting module (such as ARIMA or LSTM) for monthly demand prediction would complement the system's capabilities in supporting proactive procurement decisions.

6. CONCLUSION AND SUGGESTIONS

Conclusion

This study successfully implemented the K-Medoid Clustering algorithm to group 471 iPhone spare parts into Fast-Moving, Medium-Moving, and Slow-Moving categories based on 9 engineered features, achieving a robust Silhouette Score of 0.681 and a Davies-Bouldin Index of 0.798. The medoid approximation approach provided concrete, actionable stock recommendations, successfully shifting the store's management paradigm from intuitive to data-driven. However, a key limitation of this research is its reliance on snapshot inventory data rather than historical time-series data, which restricts the ability to capture dynamic demand trends and calculate precise turnover rates. Future research should address this by integrating daily transaction data, developing a web-based monitoring dashboard integrated directly with the Point of Sale (POS) system, and incorporating time-series forecasting modules such as ARIMA or LSTM to enable proactive, predictive procurement decisions.

Suggestions

Considering the limitations in this research and the potential for future development, the authors provide several suggestions as follows:

For MR. GADGET Store (Practical Implementation):

a) It is recommended to immediately implement the generated stock recommendations, particularly the application of safety stock and reorder points for the Fast-Moving cluster to prevent potential sales losses due to stock-outs.
b) For the Slow-Moving cluster, the store is advised to implement a pre-order system or conduct bundling promotions so that working capital is not tied up in slow-turning stock.
c) Periodic updates of stock data (e.g., every 3 or 6 months) and rerunning the clustering analysis are necessary to keep the recommendations relevant to current market dynamics.

For Future Researchers (Data Development):

a) This study used a snapshot-to-proxy approach due to the limitation of historical data. For future research, it is highly recommended to collect historical transaction data (time-series) for a minimum of 6–12 months. This data will allow for more accurate feature calculations such as inventory turnover rate, actual demand frequency, and demand variability.
b) The addition of other variables such as profit margin, purchase price, and supplier lead time can improve the depth of analysis and the accuracy of stock recommendations.

For Future Researchers (System & Algorithm Development):

a) It is recommended to develop the results of this analysis into a Web-Based Stock Monitoring Dashboard that is directly integrated with the store's Point of Sale (POS) system, so that the clustering and visualization processes can run in real-time and automatically.
b) Future researchers can try comparing the K-Medoid algorithm with other unsupervised learning algorithms such as DBSCAN, Hierarchical Clustering, or using a hybrid approach (e.g., K-Means for initial K-Medoid initialization) to see if it can produce more optimal evaluation metrics.
c) The addition of a Time-Series Forecasting module (such as ARIMA or LSTM) can be integrated to predict demand in the following months, making stock decision-making more proactive.

BIBLIOGRAPHY

[1] F. Alfiah, Almadayani, D. A. Farizi, dan E. Widodo, "Analisis Clustering K-Medoids Berdasarkan Indikator Kemiskinan Di Jawa Timur Tahun 2020," *Journal Ilmiah Sains*, vol. 22, no. 1, 2022. e-ISSN: 2540-9840.

[2] N. Amida, N. Nahadi, F. M. T. Supriyanti, L. Liliasari, D. Maulana, R. Z. Ekaputri, dan I. S. Utami, "Phylogenetic analysis of Bengkulu citrus based on DNA sequencing enhanced chemistry students' system thinking skills: Literature review with bibliometrics and experiments," *Indonesian Journal of Science and Technology*, vol. 9, no. 2, pp. 337-354, 2024.

[3] D. Anggraini dan M. S. Hasibuan, "Initial centroid pada algoritma K-means dan K-medoids," *Jurnal Multimedia dan Teknologi Informasi (Jatilima)*, vol. 7, no. 03, pp. 487-500, 2025.

[4] J. Aryandi dan O. Onsardi, "Pengaruh kualitas pelayanan dan lokasi terhadap keputusan pembelian konsumen pada Cafe Wareg Bengkulu," *Jurnal Manajemen Modal Insani Dan Bisnis (Jmmib)*, vol. 1, no. 1, pp. 117-127, 2020.

[5] M. W. Berry, A. Mohamed, dan B. W. Yap, Eds., *Supervised and unsupervised learning for data*

science, vol. 10. Cham, Switzerland: Springer, 2020.

[6] R. K. Dinata, Safwandi, N. Hasdyna, dan N. Azizah, "Analisis K-Means Clustering Pada Data Sepeda Motor," *Informatics Journal*, vol. 5, no. 1, 2020. ISSN: 2503-250X.

[7] A. S. Fatah dan A. T. Hidayat, "Integrated Ensemble Forecasting and EOQ Optimization in Spare Part Inventory," *Journal of Scientific Research, Education, and Technology (JSRET)*, vol. 4, no. 4, pp. 2414-2428, 2025.

[8] I. Fatma, H. S. Tambunan, dan F. Rizki, "Analisis Metode K-Medoids Cluster Dalam Mengelompokkan Siswa Yang Berprestasi," *Bulletin of Informatics and Data Science*, vol. 1, no. 1, 2022. ISSN: 2580-8389.

[9] A. U. Fitriyadi dan A. Kurniawati, "Analisis Algoritma K-Means dan K-Medoids Untuk Clustering Data Kinerja Karyawan Pada Perusahaan Perumahan Nasional," *Jurnal KILAT*, vol. 10, no. 1, 2021. e-ISSN: 2655-4925.

[10] S. Fotopoulou, "A review of unsupervised learning in astronomy," *Astronomy and Computing*, vol. 48, p. 100851, 2024.

[11] E. Herman, K. E. Zsido, dan V. Fenyves, "Cluster analysis with k-mean versus k-medoid in financial performance evaluation," *Applied Sciences*, vol. 12, no. 16, p. 7985, 2022.

[12] R. N. Ibrahim, M. N. Hayati, dan F. D. T. Amijaya, "Penerapan Algoritma K-Medoids pada Pengelompokan Wilayah Desa atau Kelurahan di Kabupaten Kutai Kartanegara (Studi Kasus: Data Hasil Pendataan Potensi Desa (PODES) Tahun 2018)," *Jurnal Eksponensial*, vol. 11, no. 2, 2020. ISSN: 2085-7829.

[13] G. James, D. Witten, T. Hastie, R. Tibshirani, dan J. Taylor, "Unsupervised learning," in *An introduction to statistical learning: with applications in Python*, Cham: Springer International Publishing, 2023, pp. 503-556.

[14] D. Jollyta, W. Ramdhan, dan M. Zarlis, *Konsep Data Mining Dan Penerapan*. Yogyakarta: Deepublish, 2020.

[15] N. Li, D. M. Arnold, D. G. Down, R. Barty, J. Blake, F. Chiang, et al., "From demand forecasting to inventory ordering decisions for red blood cells through integrating machine learning, statistical modeling, and inventory optimization," *Transfusion*, vol. 62, no. 1, pp. 87-99, 2022.

[16] C. N. Manik dan L. S. Istiyowati, "Pengembangan Aplikasi E-Commerce Spare Part Berbasis IOS," *Jurnal Ticom: Technology of Information and Communication*, vol. 13, no. 3, pp. 123-128, 2025.

[17] S. Naeem, A. Ali, S. Anam, dan M. M. Ahmed, "An unsupervised machine learning algorithms: Comprehensive review," *International Journal of Computing and Digital Systems*, 2023.

[18] M. J. Neuer, "Unsupervised learning," in *Machine Learning for Engineers: Introduction to Physics-Informed, Explainable Learning Methods for AI in Engineering Applications*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2024, pp. 141-172.

[19] R. Parina, A. Wijaya, dan Y. Apridiansyah, "Aplikasi Chatbot Sebagai Media Pembelajaran Interaktif SD N 17 Kota Bengkulu Berbasis Android," *Jurnal Media Infotama*, vol. 18, no. 1, pp. 121-127, 2022.

[20] H. D. Perez, C. D. Hubbs, C. Li, dan I. E. Grossmann, "Algorithmic approaches to inventory management optimization," *Processes*, vol. 9, no. 1, p. 102, 2021.

[21] W. Poncotoyo, Y. Tatianan, A. Arhab, S. A. Sholihah, dan E. Saribanon, "Implementasi Analisis Fast, Slow, Non Moving (FSN) Untuk Efesiensi Pengerjaan Picking Pada Gudang 3PL," *Jurnal Sistem Transportasi & Logistik*, vol. 3, no. 3, pp. 160-172, 2024.

[22] C. Prianto dan S. Bunyamin, *Panduan Pembuatan Aplikasi Clustering Gangguan Jaringan Menggunakan Metode K-Means Clustering*. Bandung: Penerbit Kreatif Industri Nusantara, 2020.

[23] R. I. Putra, T. Kristanto, dan F. W. Putro, "Sistem Informasi Jasa Reparasi Gadget Berbasis Website di Toko Light Service," *TIN: Terapan Informatika Nusantara*, vol. 4, pp. 203-210, 2023.

[24] B. Rajoub, "Supervised and unsupervised learning," in *Biomedical signal processing and artificial intelligence in healthcare*, Academic Press, 2020, pp. 51-89.

[25] N. R. P. Ratnasari, "Comparative study of k-mean, k-medoid and hierarchical clustering using data of tuberculosis indicators in Indonesia," *Indonesian Journal of Life Sciences*, pp. 9-20, 2023.

[26] A. L. Ridwan, Siswanto, dan R. T. Alinse, "Clustering Sales Patterns of Best Selling and Less Selling Products at El Jhon Bengkulu Stores Using the K-Medoid Method," *Jurnal Komputer, Informasi dan Teknologi (KOMITEK)*, vol. 2, no. 2, pp. 637-642, 2022. https://doi.org/10.53697/jkomitek.v2i2

[27] F. D. Savitri, "Penerapan Metode Cart Dalam Memprediksi Penjualan Produk Fast Moving Dan Slow Moving," *Journal of Informatics, Electrical and Electronics Engineering*, vol. 1, no. 4, pp. 119-125, 2022.

[28] E. Setiawati, U. D. Fernanda, S. Agesti, M. Iqbal, dan M. O. A. Herjho, "Implementation of K-Means, K-Medoid and DBSCAN algorithms in obesity data clustering," *IJATIS: Indonesian Journal of Applied Technology and Innovation Science*, vol. 1, no. 1, pp. 23-29, 2024.

[29] M. Seyedan, F. Mafakheri, dan C. Wang, "Order-up-to-level inventory optimization model using time-series demand forecasting with ensemble deep learning," *Supply Chain Analytics*, vol. 3, p. 100024, 2023.

[30] A. Subhan, A. Faqih, dan B. Irawan, "Clustering item fast moving dan slow moving pada produk unilever menggunakan algoritma k-prototype," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 6, no. 2, pp. 629-634, 2022.

[31] K. Tyagi, C. Rane, R. Sriram, dan M. Manry, "Unsupervised learning," in *Artificial intelligence and machine learning for edge computing*, Academic Press, 2022, pp. 33-52.

[32] D. Valkenborg, A. J. Rousseau, M. Geubbelmans, dan T. Burzykowski, "Unsupervised learning," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 163, no. 6, pp. 877-882, 2023.

[33] M. Wahyudi, Masitha, R. Saragih, dan Solikhun, *Data Mining: Penerapan Algoritma K-Means Clustering dan K-Medoids Clustering*. Medan: Yayasan Kita Menulis, 2020.

[34] A. Wanto, Siregar, Windarto, Hartama, Ginantra, Napitupulu, Negara, Lubis, Dewi, dan Prianto, *Data Mining: Algoritma Dan Implementasi*. Medan: Yayasan Kita Menulis, 2020.

[35] A. Wulandari, S. Sakirah, dan O. Karyono, "Analisis Perilaku Konsumen pada Trend Pembelian Produk Iphone dalam Tinjauan Maqashid Syariah (Studi Kasus Generasi Zilenial di Kabupaten Bone)," *Economics and Digital Business Review*, vol. 5, no. 2, pp. 654-666, 2024.

[36] R. F. Zannath, B. W. Fitriadi, dan R. T. Yusnita, "The influence of perceived brand authenticity and self-brand congruence on consumer satisfaction (survey of iphone smartphone users at The University Of Perjuangan)," *Journal of Management, Economic, and Accounting*, vol. 2, no. 2, pp. 433-444, 2023.