Jurnal Teknik Industri, Vol. 27, No. 2, December 2025
ISSN 1411-2485 print / ISSN 2087-7439 online
DOI: https://doi.org/10.9744/jti.27.2.225-236

Adaptive Zone-Based Inventory Framework using Self-Supervised Learning for Cost-Efficient Restocking in the Food and Beverage Industry

Anindya Annisa Agung[1], Juniwati [1*], Intan Mardiono[1,2], Yu-Chieh Wang[2]
[1)] Industrial Engineering Study Program, Institut Teknologi Sumatera
Jl. Terusan Ryacudu, Desa Way Hui, Kecamatan Jatiagung, Lampung Selatan 35365, Indonesia
[2)] Department of Mechanical Engineering, National Central University
No. 300, Zhongda Rd, Zhongli District, Taoyuan City, Taiwan 320217, Taiwan (R.O.C.)
Email: anindya31072002@gmail.com, juniwati@ti.itera.ac.id*,
intan.mardiono.im5@gmail.com, wanguchieh@gmail.com
*Corresponding author

Abstract: The food and beverage service industry operates under high demand volatility, requiring inventory systems that are both adaptive and cost-efficient. A central challenge is maintaining product availability without excessive inventory that inflates costs. The objective of this study is to develop a data-driven restocking framework that improves cost efficiency while accounting for real operational constraints. The proposed method integrates K-Means clustering with a decision tree to generate interpretable, rule-based stock recommendations. K-Means clustering was applied as an unsupervised approach to group items into risk-based zones (Green, Yellow, Red), which were then used as labels in a supervised Decision Tree model. The model achieved 99% accuracy and an F1-score of 0.93. When applied to real industry data, it reduced Total Inventory Cost (TIC) by up to 16.9% compared with the company's MOQ-based policy while preserving stable service performance. These findings demonstrate that combining clustering and rule-based machine learning provides a practical, cost-efficient, and interpretable solution for optimizing restocking decisions in complex operational environments.

Keywords: Food and beverage service industry, inventory management, K-means clustering, decision tree.

Introduction

The food and beverage (F&B) service industry in Indonesia is a highly dynamic sector that contributes significantly to the national economy, accounting for over 38% of the non-oil and gas processing industry's Gross Domestic Product (GDP) in 2023 [1]. This rapid growth, driven by increasing demand for practical, ready-to-eat products and expanding digital distribution networks, necessitates adaptive and efficient inventory management systems [2, 3, 4]. A critical challenge for F&B service providers lies in ensuring consistent product availability while avoiding excessive logistical costs associated with overstocking [5] [6]. Consequently, effective inventory management is paramount for maintaining operational continuity and service reliability within this sector [7, 8, 9].

Despite industry demands, the company is one of the major players in Indonesia's food and beverage service sector, faces significant challenges with inventory management. Internal reports show that 59% of its product portfolio exceeds the company's maximum Days in Inventory (DII) threshold of 21 days, while the average Cycle Service Level (CSL) remains at only 70%, below the 95–99% industry benchmark [10]. This discrepancy creates a strategic dilemma: attempts to ensure product availability result in significant cost inefficiencies without commensurate improvements in service quality. The underlying issues are complex, encompassing suboptimal interdepartmental coordination, reliance on subjective decision-making rather than data-driven insights, and the absence of systematic inventory planning to optimize order quantities, safety stock, and reorder points. Furthermore, conventional inventory models, such as the Economic Order Quantity (EOQ), face practical limitations under volatile demand and constraint-heavy environments [11, 12].

Cycle Service Level (CSL) is widely recognized as a key performance indicator in supply chain planning, essential for balancing service performance and minimizing holding costs [10, 13, 14]. Its traditional application as a fixed numerical target (e.g., 95%) often fails to capture item-level variability [15]. This static approach tends to overlook differences in demand patterns, usage frequency, and uncertainty, resulting in excessive inventory, stockouts, or delayed deliveries. Consequently, identical CSL values may correspond to distinct risk profiles depending on demand volatility and lead-time fluctuations. These limitations highlight the need for a more adaptive classification method in which CSL functions as a contextual reference rather than an absolute benchmark in restocking decisions.

Recent advancements in machine learning (ML) provide new opportunities to capture these complexities. Studies by Mohammad [16] have demonstrated that ML algorithms, particularly Decision Trees, offer higher classification accuracy and interpretability than traditional approaches such as FSN-Fuzzy, thereby improving stock movement and warehouse efficiency. Nevertheless, a critical gap persists in the existing literature: the direct integration of service performance indicators, such as CSL, into ML-driven classification models, as well as a comprehensive exploration of their practical implications for dynamic restocking policies, remains largely underexplored. However, most prior research has not fully integrated service-level indicators into data-driven frameworks, nor evaluated their operational implications for adaptive restocking [17].

To address these identified gaps, this study develops an adaptive, data-driven framework designed to generate order-size recommendations through machine learning. The proposed model integrates K-Means clustering and Decision Tree rules to determine SKU-level order quantities (Q) and evaluate their Total Inventory Cost (TIC) performance against the company's MOQ-based policy. Conversely, rather than treating CSL as a static determinant, the framework repositions it as a contextual indicator for interpreting service adequacy, showing that high performance can still be achieved even when conventional CSL thresholds (90–95%) are not strictly met. The primary objective is to enhance cost efficiency through zone-based restocking rules derived from empirical data patterns, demonstrating measurable efficiency improvements while maintaining acceptable service levels.

Methods

This study developed an adaptive, data-driven inventory framework using K-Means clustering and a Decision Tree algorithm to generate SKU-level order size recommendations. Figure 1 shows that the research methodology was structured into several distinct phases: data acquisition and pre-processing, risk zone segmentation through clustering, classification model development, and rigorous model evaluation.

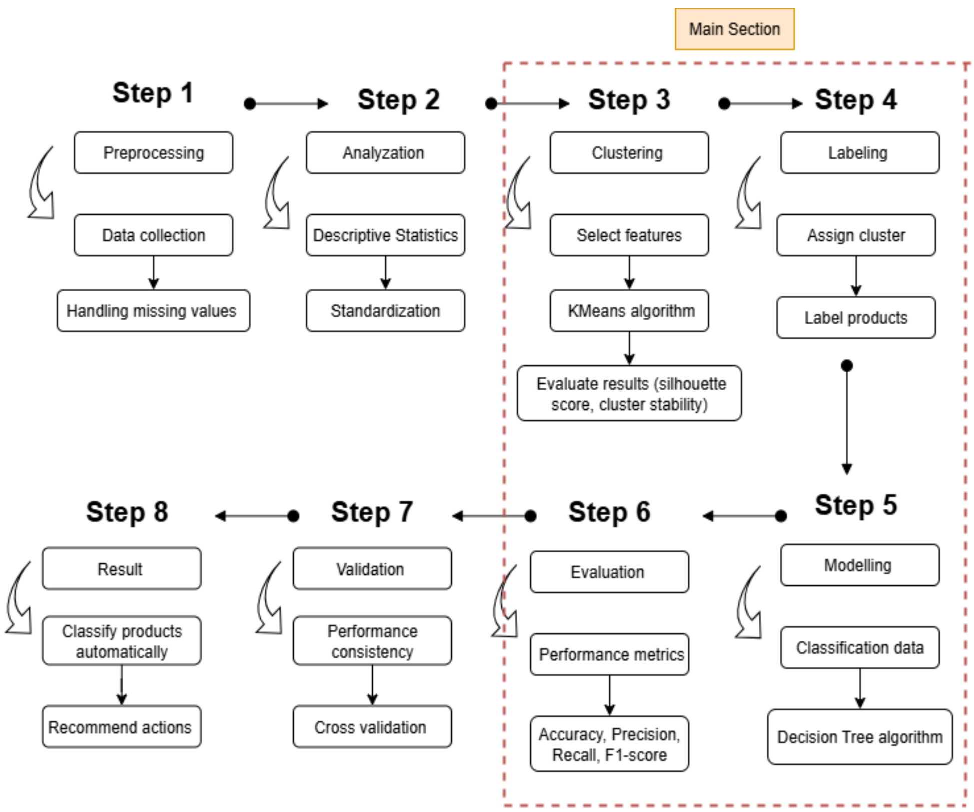

Figure 1. Research flow framework

Data Collection and Preprocessing

This study utilized both primary and secondary data sources. Primary data were obtained through structured interviews with the procurement division of the company to identify undocumented cost components and verify current inventory storage practices. Meanwhile, secondary data were extracted from the company's Supply Chain Management System (MRPS), focusing on operational records for frozen food items from all distribution centers.

The modeling dataset shown in Table 1 comprises inventory records for 494 SKUs, collected during the second week of June 2024, a period selected for its data completeness and alignment with ongoing operational cycles. The dataset included 24 attributes, comprising both numerical and categorical features such as weekly average demand (DEMAND_1 to DEMAND_12), current stock levels, actual lead time (in days), and other key product identifiers. Prior to analysis, missing values were handled using appropriate deletion and imputation techniques. Feature selection was conducted using correlation analysis and domain validation to retain attributes most relevant to the modeling objectives, while avoiding potential data leakage.

Table 1. List of input variables

Variable	Description	Type	Unit
SOH	The current physical stock available in the warehouse	Numerical	Unit
COMMITTED	Stock that has already been allocated or requested by outlets but has not yet been deducted from SOH.	Numerical	Unit
AVAILABLE	Stock that is truly ready for use, calculated as SOH minus Committed.	Numerical	Unit
AVG_DEMANDSALES_DAY	Stock that is truly ready for use, calculated as SOH minus Committed.	Numerical	Unit
*DEMAND*_MONTH	Average monthly demand	Numerical	Unit
*DEMAND*_YEAR	Average annual demand	Numerical	Unit
*DEMAND*_1 - 12	Weekly demand from Week 1 to Week 12	Numerical	Unit
LEAD_TIME	The time required from placing an order until the goods are received	Numerical	Days
DAY_SAFETY	Buffer days for additional safety stock are applied on a situational basis, as needed and per policy.	Numerical	Days
DL	Total estimated demand during the lead time period (calculated as average demand × lead time)	Numerical	Unit
STD L	Standard deviation of weekly demand over 12 weeks, adjusted for lead time (i.e., √lead time × standard deviation)	Numerical	Unit
MOQ	Minimum Order Quantity required when placing a restocking order	Numerical	Unit
CSL 2	The actual Cycle Service Level is based on demand patterns and the company's inventory policy.	Numerical	Percentage
WHS	Warehouse code where the product is stored	Categorical	-
PLU_CUST	Unique SKU (Stock Keeping Unit) code representing each product's identity	Categorical	-
DESCRIPTON	Product name or description	Categorical	-

To ensure comparability among numerical variables and to optimize the performance of distance-based algorithms, all continuous attributes were standardized using Z-score normalization. This transformation adjusted each variable to have a mean of zero and a standard deviation of one, thereby eliminating scale differences that could bias the model. Standardization was particularly crucial for clustering and classification, as unscaled data could lead to distorted distance calculations and suboptimal decision boundaries. By applying this preprocessing step, the model ensured that each feature contributed proportionately to the learning process, enhancing both the interpretability and stability of the analytical outcomes [18].

Clustering Analysis

An unsupervised learning approach using the K-Means clustering algorithm was employed to segment inventory items into distinct risk-based categories. This method was selected because there were no predefined labels, allowing the algorithm to autonomously detect patterns and form clusters based on procurement characteristics, such as recent demand levels and inventory during the lead time.

To determine the optimal number of clusters, the Silhouette Coefficient was used, a metric that measures the consistency of each data point within its assigned cluster relative to other clusters. The coefficient ranges from -1 to 1, with values closer to 1 indicating more defined and cohesive clusters. This validation ensured that the resulting segments were meaningful and actionable for inventory management.

An additional step was conducted to identify the most influential variables in shaping the clusters. For each candidate feature, the variance of its values across the cluster centroids was calculated to assess how strongly the feature differentiated between clusters. Features with higher between-cluster variance were considered more discriminative. The two features with the largest variances, which are demand during lead time (DL) and average demand at the 12th period (DEMAND_12), were selected as dominant drivers of the clusters and later confirmed by the Decision Tree as primary split criteria.

The clustering analysis produced three well-separated groups, which were subsequently interpreted and labeled according to their relative stockout risk: Green Zone (low risk), Yellow Zone (moderate risk), and Red Zone (high risk). These zone labels served as the basis for the next modeling phase, in which the clustering output served as class targets for supervised classification.

Decision Tree Algorithm

A supervised classification model was developed using the Decision Tree Classifier to predict each product's risk zone. The model was trained using cluster labels (Green, Yellow, Red) derived from the previous K-Means clustering step as target variables. To ensure generalizability, the dataset was split into training (80%) and testing (20%) subsets. Hyperparameter tuning was conducted using GridSearchCV, optimizing parameters such as ccp_alpha, max_depth, min_samples_leaf, and min_samples_split to enhance performance and reduce overfitting.

One of the key advantages of the Decision Tree model lies in its interpretability. The resulting tree structure allows for the extraction of clear decision rules that support stock adjustment decisions. These threshold-based splits offer a logical foundation for determining order quantities, ensuring SKU-level recommendations align with the characteristics of the Green Zone. This transparency enhances the model's practical applicability for inventory managers.

Model performance was evaluated using a classification report that included precision, recall, F1-score, and overall accuracy, calculated on the test set. To assess the model's generalizability and mitigate overfitting, a k-fold cross-validation procedure (k=5) was also conducted. The validation results demonstrated consistent performance across folds, supporting the robustness of the Decision Tree model.

Total Inventory Cost Analysis Cost Evaluation Using TIC

This study conducted a financial evaluation of the proposed inventory adjustment strategy using the Total Inventory Cost (TIC) framework. TIC was used to compare the baseline cost under the company's existing Minimum Order Quantity (MOQ) policy with the cost after implementing the model's recommendations. The analysis focused on two simulation tracks: (1) shifting items classified in the Red and Yellow Zones toward the Green Zone through stock adjustments, and (2) testing the impact of reduced service levels for selected Green Zone SKUs to assess the risk of over-service and explore potential cost savings from slight reductions in order quantities.

In this study, all SKUs were treated as non-perishable for cost-calculation purposes because the company does not record spoilage or deterioration costs. Using assumed values would reduce the validity of the comparison. Consequently, TIC was computed using only standard non-perishable holding and ordering cost components. The Total Inventory Cost was calculated using the classical inventory cost formula [19]:

$$TIC = \left(\frac{D}{Q}\right) S + \left(\frac{Q}{2}\right) H \tag{1}$$

Notation:
D = Annual demand
Q = Order quantity

S = Ordering cost per order
H = Holding cost per unit per year

All relevant cost parameters were obtained through structured interviews with the procurement division and validated against internal financial records. The resulting TIC values were used to assess the economic viability of the proposed adaptive inventory strategy and determine whether it offers cost advantages over the static MOQ-based approach. This analysis provided a quantitative basis for identifying whether the proposed adjustments offer measurable cost efficiencies compared to the current policy.

Results and Discussions

Clustering for Risk Zone Classification

The K-Means clustering algorithm successfully segmented 494 inventory items into three risk-based categories: Green Zone (low risk), Yellow Zone (moderate risk), and Red Zone (high risk), as shown in Figure 2. The optimal number of clusters (k = 3) was selected using the Silhouette Coefficient, which yielded a value of 0.80, indicating well-separated, internally cohesive clusters. This selection not only reflects a strong mathematical fit but also aligns with business needs for classifying inventory by stockout risk levels.

The clustering process primarily relied on two variables: demand in the 12th week (DEMAND_12) and estimated demand during lead time (DL). These two attributes were identified as the key drivers for segmentation based on their ability to highlight SKU-level variations in stock pressure. Analysis of cluster centroids revealed that items in the Green Zone exhibited low, stable demand patterns with manageable lead-time coverage, while those in the Yellow Zone exhibited moderate demand variability. In contrast, Red Zone SKUs were characterized by high average demand and large lead-time coverage requirements, increasing their vulnerability to stockouts. The scatter plot of the clustering result, shown in Figure 2, visually confirms that the clusters are clearly separable, with Red Zone items tending toward higher values in both DL and DEMAND_12 dimensions.

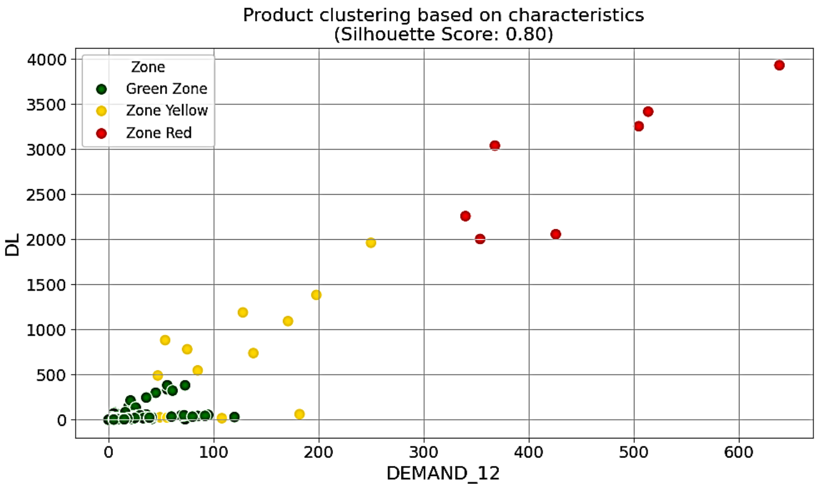

Figure 2. Zone labeling visualization

Further insight was gained from a descriptive statistical summary in Table 2, which quantitatively distinguishes the three zones. Products in the Green Zone averaged 8.49 units in monthly demand and 13.67 units in DL, suggesting low urgency and minimal service risk. In contrast, Red Zone SKUs had an average monthly demand of 449.43 units and DL values reaching 2,851.93 units, indicating high replenishment vulnerability. Despite representing only 15% of all items, Red Zone products account for a disproportionately large portion of potential stockout events, emphasizing the critical need for their prioritization.

Table 2. Distribution and descriptive statistics per zone

Zone	Items	Average *DEMAND*_12	Average DL
Green Zone	473	8.49	13.67
Yellow Zone	13	118.54	708.30
Red Zone	7	449.43	2851.93

However, it should be noted that the distribution of items across the three zones is inherently imbalanced (473 in Green, 13 in Yellow, and 7 in Red). This imbalance reflects the company's actual MOQ-based restocking practices, which tend to generate relatively high stock levels for most SKUs. As a result, many items are sufficiently supplied and classified into the Green Zone rather than the Yellow or Red. Consequently, the two features with the highest variance, DL and DEMAND_12, naturally emerged as dominant in distinguishing clusters and subsequently reappeared as the main split criteria in the Decision Tree. This alignment does not indicate overfitting but instead highlights the dataset's intrinsic structure and the central role of these features in explaining inventory risk.

This clustering analysis provided a strong analytical foundation for the subsequent classification model and zone-based inventory adjustment strategy. By segmenting items based on demand characteristics and lead-time exposure, the analysis provided a data-driven understanding of the distribution of inventory risk. The interpretability and operational relevance of the clustering output make it a valuable first step in developing a more targeted, responsive, and risk-sensitive inventory control system.

Decision Tree Model

Following the clustering process, a Decision Tree Classifier was developed to predict each SKU's risk zone based on its numerical attributes. The model was trained using the labels generated from the K-Means clustering step (Green, Yellow, Red), transforming the unsupervised segmentation into a supervised classification task to enable predictive applications. The classifier's hyperparameters were optimized using GridSearchCV, yielding a maximum tree depth of five and a minimum of one sample per leaf. The tuning process revealed that the model maintained strong classification capabilities without signs of overfitting, and pruning was not necessary at this stage. The final configuration, including selected hyperparameter values, is summarized in Table 3.

Table 3. Parameter tuning results

Parameter	Value
ccp_alpha	0.0
max_depth	5
min_samples_leaf	1
min_samples_split	2

The trained model demonstrated high interpretability, a key strength of Decision Tree algorithms. It produced a hierarchy of rules to explain and justify stock adjustment decisions. Instead of relying solely on fixed formulas such as EOQ, this approach incorporates dynamic classification logic that adapts to real operational data. The resulting tree structure provided intuitive, data-driven decision paths that guide SKU movement from higher-risk zones (Red, Yellow) into the Green Zone, using thresholds automatically derived during training. For instance, the tree identified that SKUs with DEMAND_12 greater than 1,200 units and DL (demand during lead time) exceeding 80 days had an 85% likelihood of falling into the Red Zone. These decision splits were derived from the tree's impurity-minimization process and serve as operational guides for risk-based inventory planning. The complete set of extracted classification rules is provided in Table 4.

After mapping each SKU into its respective risk zone using the trained model, targeted stock adjustments were developed based on the Decision Tree's rule outputs. These rules specify quantitative thresholds that distinguish each zone, allowing stock interventions to be made objectively and consistently across products. In this approach, SKUs classified into the Red or Yellow Zones were analyzed to identify the necessary quantity adjustments required to reach the Green Zone, representing stable availability and optimal cost efficiency. All adjustment thresholds were derived directly from the Decision Tree's learned split points, ensuring that recommendations are data-driven and not subject to managerial subjectivity. The resulting SKU-level adjustment summaries for the Jakarta distribution center are presented in Table 5.

Table 4. Decision tree rules

Rules	Description
Rule 1	IF DL <= 437.1 AND DEMAND_12 <= 47.0 THEN Zona = Zona Hijau
Rule 2	IF DL <= 437.1 AND DEMAND_12 > 47.0 AND DL <= 29.255 AND DL <= 14.535 THEN Zona = Zona Hijau
Rule 3	IF DL <= 437.1 AND DEMAND_12 > 47.0 AND DL <= 29.255 AND DL > 14.535 THEN Zona = Zona Kuning
Rule 4	IF DL <= 437.1 AND DEMAND_12 > 47.0 AND DL > 29.255 THEN Zona = Zona Hijau
Rule 5	IF DL > 437.1 AND DL <= 1983.825 THEN Zona = Zona Kuning
Rule 6	IF DL > 437.1 AND DL > 1983.825 THEN Zona = Zona Merah

Table 5. Classification results for Yellow and Red zones

PLU_CUST	CSL	Zone	Adjustment to Green Zone
PZZ-AA00005C	100%	Yellow Zone	+1935 *stock*
PZZ-AA00005H	12%	Red Zone	+1345 *stock*
PZZ-AA00006G	12%	Red Zone	+16 44 *stock*
PZZ-AA00007H	8%	Red Zone	+1479 *stock*
PZZ-AA00007I	100%	Yellow Zone	+1876 *stock*
PZZ-AA00008G	8%	Red Zone	+1470 *stock*
PZZ-AA00011F	12%	Yellow Zone	+1937 *stock*
PZZ-AA00032B	12%	Yellow Zone	+1856 *stock*
PZZ-AA00049B	8%	Red Zone	+1558 *stock*
PZZ-AA00049C	100%	Yellow Zone	+1802 *stock*
PZZ-AA00059A	10%	Yellow Zone	+1846 *stock*
PZZ-AA00096A	6%	Yellow Zone	+1930 *stock*
PZZ-AA00096E	4%	Yellow Zone	+1734 *stock*
PZZ-AA00101C	1%	Red Zone	+1616 *stock*
PZZ-AA00101D	16%	Yellow Zone	+1909 *stock*
PZZ-AA00102D	9%	Yellow Zone	+1813 *stock*
PZZ-AA00156B	14%	Red Zone	+1630 *stock*
PZZ-AA00181C	100%	Yellow Zone	+1928 *stock*
PZZ-AA00201A	13%	Yellow Zone	+1786 *stock*
PZZ-AA00230B	6%	Yellow Zone	+1899 *stock*

In addition to these rule-based recommendations, Table 5 also presents each SKU's Cycle Service Level (CSL), calculated using the standard service-level formula. Here, however, CSL serves as a contextual performance indicator, complementing the zone-based classification rather than defining it. This contextual interpretation helps reveal that static CSL values can sometimes misrepresent actual inventory risks: for instance, products with low demand and short lead times may show low CSL numerically while still maintaining acceptable availability. Hence, within this framework, CSL provides supportive insight into service performance but does not dictate restocking decisions.

Overall, the integration of data-driven classification and rule-based stock adjustment establishes an operationally transparent complement to traditional decision models such as EOQ. By grounding restocking policies in interpretable Decision Tree logic, this approach enhances accuracy, aligns with real operational constraints, and promotes adaptive decision-making across dynamic product portfolios—particularly relevant for complex sectors such as frozen food logistics.

Evaluation and Validation

Model evaluation was conducted to assess how effectively the Decision Tree classifier could distinguish between the three inventory risk zones. The model achieved a high overall accuracy of 99% on the test set, with substantial precision and F1-scores across most classes. As shown in Table 6, classification performance was robust for the Green and Red Zones. However, the recall score for the Yellow Zone was lower, at 0.67, indicating that some items in this category were misclassified. This discrepancy is likely due to class imbalance, as the Yellow Zone accounts for a smaller proportion of the data, limiting the model's ability to generalize to that class. Addressing this limitation in future work may involve collecting a more balanced dataset to improve classification sensitivity for minority classes.

Table 6. Classification report result

Zone	Precision	Recall	F1-Score	Support
Zona Hijau	0.99	1.00	0.99	95
Zona Kuning	1.00	0.67	0.80	3
Zona Merah	1.00	1.00	1.00	1
Macro Avg	1.00	0.89	0.93	99
Weighted Avg	0.99	0.99	0.99	99
Accuracy			0.99	99

To further validate the model's robustness, a 5-fold cross-validation procedure was conducted. As shown in Table 7, the model achieved accuracy scores ranging from 90.82% to 100%, with an average of 96.75% across all folds. These findings demonstrate the model's consistent performance across varying data partitions, indicating high generalizability and minimal risk of overfitting [20].

Table 7. Cross-validation value

Fold	Cross Validation Score
Fold 1	0.95
Fold 2	0.99
Fold 3	1.00
Fold 4	0.99
Fold 5	0.91

The combination of strong test accuracy and consistent cross-validation performance supports the reliability of the Decision Tree model for practical deployment in inventory risk classification tasks. These results confirm that the model is not only accurate under static test conditions but also generalizable across varying data splits. Such consistency is essential for real-world applications, especially in dynamic inventory environments where model robustness is crucial.

Cost Efficiency Analysis

After generating order-size recommendations through the Decision Tree model, a financial evaluation was conducted to compare the cost efficiency of the proposed framework against two benchmarks: (i) the company's existing MOQ-based restocking policy, and (ii) the classical Economic Order Quantity (EOQ) model used as a theoretical reference. Total Inventory Cost (TIC) was employed as the evaluation metric, using identical ordering (S) and holding (H) cost parameters across all scenarios to ensure methodological consistency. The only parameter that differed among the three approaches was the order quantity (Q): MOQ for the company's policy, Q_DT for the Decision Tree framework, and Q_EOQ for the EOQ baseline.

Cost parameters were sourced directly from the company's actual expense structure. Holding cost (H) components included handling/picking costs of Rp300 per carton and monthly pallet storage fees—Rp3,000 for dry food and Rp6,000 for frozen items—per pallet position. Ordering costs (S) accounted for truck delivery fees of Rp 1,200,000 per 12 m^3, allocated proportionally to outlet usage. A breakdown of the specific S and H values used in the simulation is provided in Table 6.

Table 8. Details of cost components

Cost	Type	Monthly Cost Components (Rp)	Total Cost Per Year (Rp)
Holding	Dry	3,300	39,600
	Frozen	6,300	75,600
Set up	Delivery service	1,200,000	14,400,000

Although some SKUs, particularly frozen products, may be perishable in practice, the absence of measurable deterioration or spoilage records prevents the use of perishable-inventory cost models. Therefore, all TIC calculations consistently adopt the non-perishable formulation to ensure methodological consistency across items and scenarios.

To illustrate the TIC calculation process, consider SKU PZZ-AA00005C, which has an annual demand (D) of 8,289 units. The model-recommended order quantity (Q_DT) is 1,935 units, with an ordering cost (S) of Rp14,400,000 and a holding cost (H) of Rp79,200 per unit per year.

This value was then compared to the TIC obtained using the company's MOQ and the EOQ value for the same SKU to quantify potential savings. The simulation was extended to all SKUs in the Red and Yellow Zones for which quantity adjustments were recommended. Table 9 presents the comparative results, listing Q_MOQ, Q_DT, and Q_EOQ, along with their corresponding TIC values.

Across all evaluated items, the company's MOQ-based policy produced an average TIC of Rp248,369,545. The Decision Tree framework reduced this cost to Rp206,439,553 (a 16.9% reduction), while the EOQ benchmark reduced it to Rp177,625,591 (a 28.5% reduction). This outcome is expected, as EOQ inherently yields the lowest theoretical TIC under its idealized assumptions. However, the Decision Tree framework achieves substantial cost savings while remaining feasible under real operational constraints—such as highly variable demand per cycle (DEMAND_12), fluctuating demand during lead time (DL), and differing storage cost structures for dry and frozen items. Because the Decision Tree model is trained on these real operational variables, these recommendations capture risk patterns and constraints that EOQ does not account for. Therefore, EOQ serves as a theoretical lower bound, while the Decision Tree framework provides an operationally feasible solution aligned with real constraints.

Table 9. Comparison of the TIC values of the Yellow and Red zones

PLU_CUST	MOQ	EOQ	Q_DT	DEMAND_YEAR	TIC	TIC_EOQ	TIC_DT
PZZ-AA00005C	580	1736	1935	8289	228,691,394	137,502,282	138,311,581
PZZ-AA00005H	2752	3781	1345	39312	314,687,150	299,447,991	474,148,840
PZZ-AA00006G	1581	2866	1644	22584	268,317,134	226,965,197	262,918,458
PZZ-AA00007H	2278	3440	1479	32547	295,934,570	272,467,255	375,456,027
PZZ-AA00007I	371	1389	1876	5307	220,425,290	110,022,974	115,025,634
PZZ-AA00008G	2392	3525	1470	34173	300,441,842	279,190,340	392,967,918
PZZ-AA00011F	344	1336	1937	4911	219,327,578	105,838,531	113,214,441
PZZ-AA00032B	833	2081	1856	11904	238,712,174	164,780,302	165,856,221
PZZ-AA00049B	1440	2735	1558	20574	262,745,414	216,629,802	251,854,438
PZZ-AA00049C	1304	2603	1802	18630	257,356,646	206,141,419	220,233,784
PZZ-AA00059A	518	1640	1846	7395	226,213,226	129,875,707	130,787,407
PZZ-AA00096A	412	1463	1930	5889	222,038,594	115,898,979	120,366,653
PZZ-AA00096E	916	2182	1734	13089	241,996,994	172,787,400	177,363,978
PZZ-AA00101C	1419	2715	1616	20268	261,897,182	215,012,784	244,599,541
PZZ-AA00101D	365	1377	1909	5214	220,167,494	109,054,690	114,926,730
PZZ-AA00102D	766	1994	1813	10938	236,034,422	157,952,969	158,671,358
PZZ-AA00156B	1403	2700	1630	20043	261,273,482	213,815,999	241,614,994
PZZ-AA00181C	516	1637	1928	7368	226,138,382	129,638,394	131,379,505
PZZ-AA00201A	969	2244	1786	13845	244,092,626	177,707,319	182,353,819
PZZ-AA00230B	383	1411	1899	5478	220,899,302	111,781,478	116,739,736

Trade-Off for Green Zone

After implementing stock adjustments for items in the Red and Yellow Zones, an additional sensitivity analysis was conducted for SKUs in the Green Zone. The purpose of this trade-off was to test whether products identified as Green might in fact represent an over-service condition, given that the framework does not explicitly define a minimum threshold for this zone. Because the dataset represents a snapshot during a period of sufficient stock, some SKUs may appear as Green even though they are adequately manageable under the Yellow Zone.

To perform this test, adjustments were expressed through the two inventory components recognized by the model: actual stock (stock), which is stock to fulfill actual demand derived from average demand (DEMAND_12), and safety stock (ss), which is stock to reserve demand in lead time, derived from expected demand during lead time (DL). For items in the Red and Yellow Zones that were shifted into the Green Zone, the Decision Tree rules prescribed adjustments only to actual stock. This is because increasing DEMAND_12 to the required threshold inherently secures sufficient coverage for lead-time demand as well. In other words, shortages in these zones are driven mainly by insufficient cycle stock, and correcting this component alone is sufficient to restore both routine fulfillment and lead-time protection.

In contrast, Green Zone items typically maintain substantial buffers in both actual stock and safety stock. To reclassify them into the Yellow Zone, the Decision Tree rules required simultaneous reductions in both DEMAND_12 and DL, since their overstock condition extended across both dimensions. This dual condition arises directly from the model's learned classification rules, meaning that while understocked items can be corrected simply by raising actual stock, overstocked Green items require explicit adjustments to both the routine cycle stock and the lead-time buffer. Consequently, the results are reported as paired reductions (e.g., "-336 ss, -58 stock"), reflecting the two-dimensional requirement imposed by the Decision Tree. For instance, "-336 ss, -58 stock" indicates that 336 units were reduced from the safety stock buffer (DL) and 58 units from the cycle stock (DEMAND_12). The complete outcomes of these Green-to-Yellow adjustments are summarized in Table 10.

Table 10. Comparison of tic values for trade-off

PLU_CUST	Adjustment to Yellow Zone	Q	MOQ	DEMAND_YEAR	TIC DT	TIC
PZZ-AA00011E	-336 ss, -58 *stock*	786	1,180	3,828	101,265,313	93,442,575
PZZ-AA00012I	+36 ss, -60 *stock*	1,090	1,114	3,414	88,266,021	88,245,099
PZZ-AA00019E	-95 ss, -4 *stock*	620	719	1,422	57,577,863	56,951,954
PZZ-AA00030F	-253 ss, -30 *stock*	761	1,044	3,000	86,886,727	82,721,702
PZZ-AA00030H	-199 ss, -21 *stock*	726	946	2,463	77,591,965	74,953,349
PZZ-AA00101E	+41 ss, -58 *stock*	793	810	1,803	64,143,766	64,129,329
PZZ-AA00111B	-167 ss, -6 *stock*	650	883	2,145	73,254,103	69,947,546
PZZ-AA00130A	-290 ss, -41 *stock*	777	1,108	3,375	93,324,688	87,739,615
PZZ-AA00225A	-336 ss, -41 *stock*	803	1,180	3,831	100,485,731	93,479,184
PZZ-AA00260A	-34 ss, +9 *stock*	517	542	807	42,951,683	42,903,784
PZZ-AA00395A	-17 ss, +7 *stock*	473	483	642	38,275,539	38,267,170
PZZ-AA00402A	-9 ss, -2 *stock*	531	542	807	42,912,813	42,903,784
PZZ-AA00403A	-89 ss, -11 *stock*	603	703	1,359	56,332,806	55,676,069
PZZ-AA00404A	-276 ss, -46 *stock*	802	1,084	3,234	89,811,770	85,887,279
PZZ-AA00049C	+45 ss, -1 *stock*	475	519	741	41,273,308	41,111,937
PZZ-AA00005H	+17 ss, -2 *stock*	729	744	1,521	58,913,332	58,901,105
PZZ-AA00007H	+15 ss, -2 *stock*	748	761	1,593	60,288,041	60,279,095
PZZ-AA00008G	+14 ss, -4 *stock*	770	780	1,674	61,797,755	61,792,613
PZZ-AA00049C	+2 ss, -12 *stock*	891	901	2,232	71,356,409	71,351,964
PZZ-AA00156B	+11 ss, -12 *stock*	803	804	1,776	63,647,398	63,647,348
PZZ-AA00007H	+17 ss, -26 *stock*	1,033	1,042	2,988	82,559,197	82,556,093
PZZ-AA00005H	-7 ss, -2 *stock*	433	442	537	35,005,631	34,998,222
PZZ-AA00005C	+7 ss, -8 *stock*	740	741	1,509	58,668,347	58,668,293
PZZ-AA00005H	+18 ss, -20 *stock*	622	624	1,071	49,426,039	49,425,784
PZZ-AA00007I	+17 ss, -24 *stock*	629	636	1,113	50,388,682	50,385,598
PZZ-AA00008G	+9 ss, -19 *stock*	709	719	1,422	56,957,539	56,951,954
PZZ-AA00049B	+14 ss, -22 *stock*	663	671	1,239	53,164,986	53,161,165
PZZ-AA00156B	+13 ss, -17 *stock*	681	685	1,290	54,245,178	54,244,248
PZZ-AA00005H	-11 ss, -21 *stock*	428	460	582	36,529,865	36,435,130
PZZ-AA00005H	-38 ss, -1 *stock*	517	556	849	44,122,665	44,006,080
PZZ-AA00005C	-3 ss, -15 *stock*	1,154	1,172	3,780	92,865,995	92,854,880
PZZ-AA00005H	+16 ss, -105 *stock*	829	918	2,316	73,060,716	72,682,208
PZZ-AA00101C	+3 ss, -77 *stock*	1,018	1,092	3,282	86,735,234	86,522,313

The simulation results revealed that shifting Green Zone items into Yellow increased Total Inventory Cost (TIC) from Rp63,855,287 to Rp64,972,336 (+1.75%). This outcome suggests that the company's current procurement policy for low-risk items is already operating near an efficient cost boundary. Furthermore, since many of these SKUs are frozen products with higher storage costs, excessive reductions in order size—whether cycle stock or safety stock—can paradoxically increase overall costs and reduce efficiency. This finding highlights that service levels should not be lowered uniformly, but rather managed in alignment with product-specific risk profiles and storage cost structures.

Conclusions

This study proposed a zone-based inventory control framework that integrates K-Means clustering, Decision Tree classification, and Total Inventory Cost (TIC) analysis to support more adaptive restocking decisions. The segmentation successfully categorized SKUs into Green, Yellow, and Red Zones, each reflecting distinct risk and urgency profiles. DEMAND_12 and demand during lead time (DL) emerged as key variables influencing the restocking strategy, offering interpretable thresholds for action.

The findings show that restocking recommendations derived from Decision Tree rules significantly reduced total inventory costs, particularly for high-risk items, while maintaining service levels. By tailoring recommendations to SKU-specific conditions, the system enabled a more contextual interpretation of CSL values, demonstrating that service levels should not be treated as absolute indicators.

Overall, integrating predictive modeling and cost evaluation offers a more responsive alternative to traditional, formula-based policies. Nevertheless, this study has several limitations. First, the framework has not yet been integrated into real-time operational platforms, which may limit its immediate applicability. Second, the imbalance in class distribution—particularly the limited number of Yellow Zone items—may reduce the model's sensitivity to minority patterns. Third, the framework does not explicitly define a minimum threshold within the Green Zone, which may lead to over-service by maintaining stock levels higher than necessary, limiting opportunities for cost efficiency.

These limitations open several avenues for future research. One direction is the integration of the framework into a real-time operational dashboard, which would improve both usability and decision speed. Another important step is addressing class imbalance through data enrichment or balancing techniques to enhance the model's ability to generalize. In addition, refining the cost analysis by differentiating holding cost structures between product categories, such as frozen and dry goods, would increase the accuracy of TIC estimation. Finally, the establishment of explicit lower-bound thresholds for Green Zone items, supported by sensitivity or trade-off analysis, could help prevent over-service and ensure a more efficient allocation of resources.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author(s) used ChatGPT by OpenAI to improve the readability and language clarity of the manuscript during its preparation. The author(s) reviewed and edited all content generated by the tool and accepted full responsibility for the accuracy and integrity of the published work.

References

[1] BPS-Statistics, Badan Pusat Statistik, *Statistik Penyediaan Makanan dan Minuman* 2023, vol. 7, 2024. https://www.bps.go.id/id/publication/2024/12/23/f2c7743c4712aaeaa4abf694/statistik-penyediaan-makanan-dan-minuman-2023.html

[2] U. Mehmood, J. Broderick, S. Davies, A. K. Bashir, and K. Rabie, "Machine learning-based predictive inventory for a vending machine warehouse," *IEEE Internet of Things Magazine*, vol. 7, no. 6, pp. 94-100, Nov. 2024. https://doi.org/10.1109/IOTM.001.2300271

[3] D. Safitri, W. Tiswiyanti and F. Olimsar, "Comparative analysis of the financial performance of food and beverage," *Management Studies and Entrepreneurship Journal,* vol. 5, no. (2), pp. 7581-7595, 2023. https://journal.yrpipku.com/index.php/msej/article/download/5373/3055/29041

[4] Y. A. M. 'Ainul, M. J. Rosyid, V. A. Leksono and A. D. Wantira, "A preference-oriented multi-criteria decision model for stunting-prevention food basket ranking using AHP-TOPSIS," *Jurnal Teknik Industri: Jurnal Keilmuan dan Aplikasi Teknik Industri,* vol. 26, no. 2,pp. 145-156, 2024. https://doi.org/10.9744/jti.26.2.145-156

[5] Y. Tan, L. Gu, S. Xu, and M. Li, "Supply chain inventory management from the perspective of 'cloud supply chain'—a data driven approach," *Mathematics*, vol. 12, no. 4, p. 573, 2024. https://doi.org/10.3390/math12040573

[6] S. M. Mose and A. Osoro, "Inventory management practices and performance of food and beverage manufacturing firms in Kisumu City County, Kenya," *Journal of Economics, Management Sciences & Procurement,* vol. 4, 2025. https://www.jemspro.org/index.php/pages/article/view/52/63

[7] S. D. Kırmızı, Z. Ceylan and S. Bulkan, "Enhancing inventory management through safety-stock strategies—a case study," *System,* vol. 12, no. 7, p. 260, 2024. https://doi.org/10.3390/systems12070260

[8] I. G. P. Vergara, M. C. L. Gómez, I. L. Martínez and J. V. Hernández, "Strategies for the preservation of service levels in the inventory management during COVID-19: A case study in a company of biosafety products," *Global Journal of Flexible Systems Management,* vol. 22, pp. 65-80, 2021. https://doi.org/10.1007/s40171-021-00271-z

[9] R. R. Panigrahi, A. K. Shrivastava and P. K. Kapur, "Impact of inventory management practices on the operational performances of SMEs: review and future research directions," *International Journal of System Assurance Engineering and Management,* vol. 15, pp. 1934-1955, 2024. https://doi.org/10.1007/s13198-023-02216-4

[10] Achkar, V. G., Braulio, B. Héctor, P. R. M. Carlos, M. Ignacio and Grossmann, "Extensions to the guaranteed service model for industrial applications," *European Journal of Operational Research,* vol. 313, pp. 192-206, 2024. https://doi.org/10.1016/j.ejor.2023.08.013

[11] C.-H. Chuang and C.-Y. Chiang, "Dynamic and stochastic behavior of coefficient of demand uncertainty incorporated with EOQ variables: An application in finished-goods inventory from General Motors' dealerships," *International Journal of Production Economics,* vol. 172, pp. 95-109, 2022. doi: https://doi.org/10.1016/j.ijpe.2015.10.019

[12] M. Y. Jaber and J. Peltokorpi, "Economic order/production quantity (EOQ/EPQ) models with product recovery: A review of mathematical," *Applied Mathematical Modelling,* vol. 129, pp. 655-672, 2024. https://doi.org/10.1016/j.apm.2024.02.022

[13] T. B. Abid, O. Ayadi and F. Masmoudi, "Enhancing apparel supply chain resilience: a robust-stochastic approach to integrated production-distribution planning," *International Journal of Management Science and Engineering Management,* vol. 20, no. 2, pp. 286-305, 2025. https://doi.org/10.1080/17509653.2025.2455962

[14] S. Sivankutty and B. Elahi, "Proposing a new dynamic safety stock adjustment system: Balancing service level targets and financial constraints," in *9th North American Conference on Industrial Engineering and Operations Management*, Washington D.C., USA, 2024. doi:https://doi.org/10.46254/NA09.20240162

[15] Mandal, I. A. Mohammed and Joydeb, "The impact of lead time variability on supply chain management," *International Journal of Supply Chain Management,* vol. 8, no. 2,, pp. 41 - 55, 2023. doi:https://doi.org/10.47604/ijscm.3075

[16] M. N. A. Z. Abidin, N. E. N. Bazin, D. A. Zebari, A. N. Rahmat and R. R. Asaad, "Inventory categorization using multiple criteria classification," *Journal of Soft Computing and Data Mining,* vol. 5, no. 1, pp. 142-151, 2024. doi: https://doi.org/10.30880/jscdm.2024.05.01.012

[17] Z. Farshadfar, T. Mucha and K. Tanskanen, "Leveraging machine learning for advancing circular supply chains: A systematic literature review," *Logistics,* vol. 8, no. 4, 2024. https://doi.org/10.3390/logistics8040108

[18] C. Wongoutong, "The impact of neglecting feature scaling in k-means clustering," *PLOS ONE,* vol. 19, p. 12, 2024. https://doi.org/10.1371/journal.pone.0310839

[19] P. Liu, A. Hendalianpour, M. H. and M. Feylizadeh, "Cost reduction of inventory-production-system in multi-echelon supply chain using game theory and fuzzy demand forecasting," *International Journal of Fuzzy Systems,* vol. 24, pp. 1793-1813, 2022. https://doi.org/10.1007/s40815-021-01240-5

[20] B. Naderalvojoud, C. M. Curtin, C. Yanover, T. El-Hay, B. Choi, R. W. Park, J. G. Tabuenca, M. P. Reeve, T. F. K. Humphreys and S. M. Asch, "Towards global model generalizability: Independent cross-site feature evaluation for patient-level risk prediction models using the OHDSI network," *JAMIA,* vol. 31, no. 5, pp. 1051-1061, 2024. doi: https://doi.org/10.1093/jamia/ocae028