P-ISSN: 2987-7261
E-ISSN: 2987-7253

Jurnal Riset Ilmu Teknik
Journal homepage: https://jurnaljepip.com/index.php/jurit
Vol 2, No. 2, pp. 178-188, 2026
DOI: doi.org/10.59976/jurit. v4i2.390

JURIT

Web-Based Sales Performance Information System Using K-Means

Inayah Ainina Mawardi[1]*, Rita Wahyuni Arifin[2]
Universitas Bina Insani, Indonesia
Correponding Author: inayahmawardi3@gmail.com

ABSTRACT

PT FAN Integrasi Teknologi is a company in the field of information technology that has sales activities as the basis for recording and evaluating sales performance. The sales performance appraisal process is still carried out manually using spreadsheets and separate documents, causing the recapitulation process to take time and make it difficult for management to evaluate performance. This research aims to develop and build a web-based sales performance information system that can integrate data management and help the sales performance evaluation process. The system development method used to classify sales performance is based on the variables of total score, number of projects, and win rate, applied to a dataset of seven historical sales records at the company. The results of the research are in the form of a sales performance information system that provides data management, dashboards, performance analysis, and reports. The clustering process grouped sales into three exploratory categories, namely High Performer, Medium Performer, and Low Performer, based on similarity in the selected variables rather than an independently validated performance standard. A beta test involving eight respondents obtained an initial satisfaction score of 84.5%. These preliminary results indicate that, as a company-specific prototype, the system has potential as a tool for monitoring and supporting the evaluation of sales performance at PT FAN Integrasi Teknologi, pending further validation with a larger dataset.

Keywords: *Data Analysis, K-Means Clustering, Sales Performance, Web-Based Information System, User Centered Design (UCD).*

Manuscript received 08 July. 2026; revised 23 August. 2026; accepted 30 August. 2026.

INTRODUCTION

An information system is a collection of hardware, software, and data components integrated to support organizational decision-making. In sales-driven organizations, sales performance data functions as a key indicator for management to evaluate employee contribution and guide decisions such as coaching, incentives, and process improvement. At PT FAN Integrasi Teknologi, an information technology company whose revenue depends on active sales operations, this evaluation process is still carried out manually using spreadsheets and separate documents [3]–[8]. This condition slows down data recapitulation, increases the risk of recording errors, and makes it difficult for managers to compare performance across periods. This research focuses on developing a web-based sales performance information system that automates data recapitulation and applies a similarity-based grouping method to support the categorization of sales performance.

Two streams of prior research are directly relevant. The first concerns User-Centered Design (UCD) for business information systems [1], [2], [3] applied UCD to design library, sales, and academic information systems and reported that placing users at the center of the requirements and design stages improved interface usability. More recently, [4], [5] and [6], [7] applied UCD specifically to e-commerce sales systems, both finding that early user involvement increased perceived ease of use. The second stream concerns K-Means Clustering for performance and sales data [8], [9], and [10], [11] used K-Means to classify employee performance or discipline into a small number of categories, while [12], [13] , and [14], [15], [16] applied the same algorithm to group sales and product transaction data. [17] combined K-Means with IT service-desk data to assess staff performance. Methodologically, [18], [19], in a comprehensive review of K-Means variants, confirm that the algorithm remains the most widely used clustering method for its simplicity, while cautioning that its outcome is highly sensitive to the number of clusters and initial centroid chosen by the user.

Despite this substantial body of work, UCD studies [20], [21], [22] and K-Means studies [23], [24], [25] largely proceed as separate lines of research. UCD studies focus on interface usability without an embedded analytical grouping method, while K-Means studies focus on clustering accuracy on a dataset without integrating the method into a complete, user-evaluated web application covering the full workflow from data entry to dashboard visualization and reporting [26], [27], [28], [29], [30]. In addition, few of these studies explicitly discuss how the number of clusters is determined, or acknowledge that when k is fixed to match predefined business categories rather than statistically validated, the resulting groups represent an exploratory rather than an objective performance measure [31]. This leaves a gap between technically oriented clustering studies and practically deployable, user-centered systems that communicate the exploratory nature of their analytical output to management [31].

This research addresses that gap by integrating UCD and K-Means Clustering into a single web-based sales performance information system, iteratively designed with input from company stakeholders and evaluated through beta testing, while explicitly treating the resulting High, Medium, and Low Performer categories as an exploratory, similarity-based grouping rather than a validated measure of objective performance. The novelty lies in combining (1) a UCD-driven development process for a sales-specific dashboard and reporting feature, (2) a transparent K-Means grouping based on total score, number of projects, and win rate, and (3) an explicit acknowledgment of the method's exploratory scope, within one deployed system evaluated at PT FAN Integrasi Teknologi, an aspect not addressed jointly in the UCD or K-Means studies reviewed above.

This research aims to design and build a web-based sales performance information system using the User-Centered Design approach and the K-Means Clustering method, so that PT FAN Integrasi Teknologi can integrate sales data management, visualize performance through an interactive dashboard, and obtain an exploratory grouping of sales performance that supports, rather than replaces, managerial evaluation.

MATERIAL AND METHOD

System Development Method (UCD)

This study uses User-Centered Design (UCD) as a system development approach that puts users as the main focus in each design stage. This method consists of four main stages: Understand Context of Use, Specify User Requirements, Design Solution, and Evaluate Against Requirements [32], [33], [34].

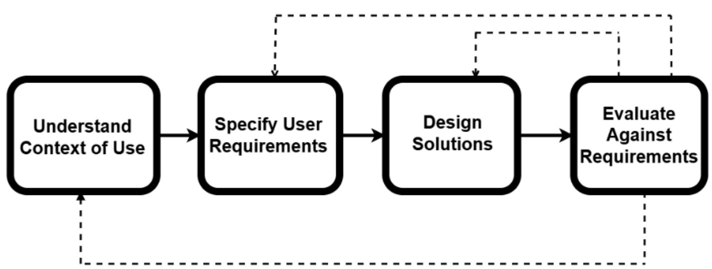

Figure 1. UCD method

K-Means Clustering Method

The data analysis technique used is K-Means Clustering, which is applied to group sales performance data based on the level of similarity. The analysis process includes data normalization, determination of the number of clusters, distance calculation using Euclidean Distance, and an iterative process until the optimal cluster is obtained.

The three variables used in the clustering process are defined as follows: total score is the cumulative performance score obtained by a salesperson from completed projects during the observed period; number of projects is the total number of projects handled by that salesperson during the same period; and win rate is the percentage of proposed projects that were successfully won. The exact scoring formula and the normalization or scaling procedure applied to these variables before the distance calculation were not fully documented in the original computation and should be clarified and made reproducible in future work [35], [36].

Distance calculation in the K-Means Clustering method is carried out using Euclidean Distance to measure the proximity between data. Euclidean Distance is a commonly used method in the clustering process because it is able to accurately calculate distances based on the value of data attributes [37], [38].
The Euclidean Distance formula is as follows:

$$d_e = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2} \tag{1}$$

Description:
d_e = Euclidean distance
a_i = the value of the i-th data attribute
b_i = the value of the i-th centroid attribute
n = the number of variables (features) used in the distance calculation, namely three in this study: total score, number of projects, and win rate

This distance calculation is used to determine the proximity of each data to the centroid so that it can be grouped into appropriate clusters [39].

Stages of the K-Means Clustering process, which starts from the input of sales performance data, determination of the number of clusters, centroid initialization, calculation of distances between data, inward grouping of data, nearest clusters, to centroid updates and produces a stable final cluster.

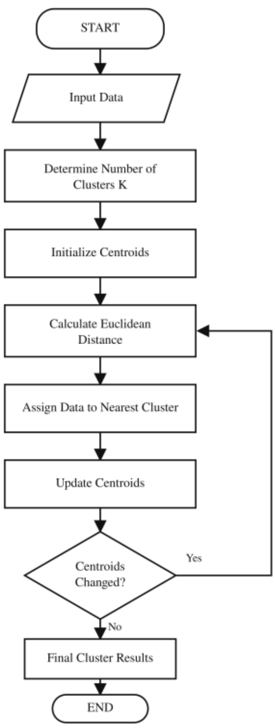

Figure 2. Stages of the k-means clustering process

Data Collection Technique
Data collection is carried out through: (1) direct observation of the sales performance management process at PT FAN Integrasi Teknologi; (2) Interviews with HRD Managers, General Managers, and Sales and Marketing Managers; (3) Literature Study by reviewing literature, journals, and articles related to sales performance information systems and K-Means Clustering.

RESULTS

The results of the study include cluster determination, K-Means calculation, system design, interface implementation, and testing. The system developed integrates sales data management and performance analysis so that clustering results can be used as information to support evaluation.

Determination of the Number of Clusters
This study uses 3 clusters that represent the sales performance category, namely High Performer, Medium Performer, and Low Performer. The number of clusters (k = 3) was set in advance to match these three predefined managerial

categories, rather than determined through a statistical validation method such as the elbow or silhouette approach; validating this choice of k statistically is left for future work.

Table 1. Cluster categories

Cluster	Categories
C1	High Performer
C2	Medium Performer
C3	Low Performer

The use of three clusters aims to simplify the process of evaluating and interpreting the results of sales performance analysis.

Early Centroid Determination

Once the number of clusters is determined, the next stage is to determine the initial centroid as the temporary center in the K-Means Clustering process.

In this study, all sales data is used in the clustering process. However, to start the K-Means iteration process, an initial centroid is required that serves as the initial reference point for data clustering.

In the manual calculation stage of the K-Means Clustering method, all sales performance data is used as data that will be grouped based on the total score, number of projects, and win rate. The data is used as initial data in the clustering process. To protect employee privacy, the seven salespeople are identified in this article using anonymous codes S1 to S7 instead of their real names.

Table 2. Early centroid values

Yes	Sales	Total Score	Number of Projects	Win Rate
1	S1	72	27	22,22
2	S2	65	24	16,67
3	S3	63	25	16,00
4	S4	61	22	18,18
5	S5	57	25	12,00
6	S6	56	22	13,64
7	S7	52	24	12,50

Based on this data, the initial centroid was selected from sales data representing the high, medium, and low-performance categories. The election aims to ensure that the clustering process can produce optimal and representative data grouping.

The initial centroids used are as follows.

Table 3. Results of Initial Centroid Determination

Cluster	Sales	Total Score	Number of Projects	Win Rate	Representation
C1	S1	72	27	22,22	High Performance
C2	S3	63	25	16,00	Medium Performance
C3	S7	52	24	12,50	Low Performance

Data S1 was chosen as the initial centroid of C1 because it has the highest performance value compared to other sales. The data of S3 was chosen as the initial centroid of C2 because it has a medium performance value that can represent the medium category. Meanwhile, S7 was chosen as the initial centroid of C3 because it has a lower performance value compared to other data so that it can represent the low-performance category.

The initial centroid value is then used as the center of the cluster while in the distance calculation process using the Euclidean Distance method.

Euclidean Distance Calculation

The calculation of the distance between the data and the centroid point is carried out using the Euclidean Distance method.

Based on the initial centroid that has been determined beforehand, a calculation is carried out to calculate the distance adapted for each sales data.

Example of calculating the distance of S1's data to the C1 centroid:

$$d = \sqrt{(72-72)^2 + (27-27)^2 + (22{,}22-22{,}22)^2} = 0$$

Example of S3's data distance calculation to C2 centroids:

$$d = \sqrt{(63-63)^2 + (25-25)^2 + (16{,}00-16{,}00)^2} = 0$$

Example of calculating the distance of S7's data to the C3 centroid:

$$d = \sqrt{(52-52)^2 + (24-24)^2 + (12.50-12.50)^2} = 0$$

The results of the Euclidean Distance calculation can be seen in the following table.

Table 4. Euclidean Distance Calculation Results

Sales	Total Score	Project	Win Rate	C1	C2	C3	Cluster
S1	100	30	33.33	0.00	19.43	68.82	C1
S7	90	30	16.67	19.43	0.00	52.94	C2
S2	90	30	16.67	19.43	0.00	52.94	C2
S5	75	30	16.67	30.04	15.00	39.09	C2
S4	75	30	16.67	30.04	15.00	39.09	C2
S3	50	25	0.00	60.30	43.62	10.00	C3
S6	40	25	0.00	68.82	52.94	0.00	C3

Clustering Results

Based on the results of the Euclidean Distance calculation, each sales data is grouped into clusters with the smallest distance to the centroid. Data that has the closest distance to C1 goes into the High Performer category, C2 to Medium Performer, and C3 to Low Performer. This grouping is carried out based on the total score, number of projects, and win rate.

The assignment of each sales record to its nearest cluster (C1, C2, or C3) is shown in the Cluster column of Table 4 above (Euclidean Distance Calculation Results). A separate table that reproduced the same seven records with different, unreconciled values has been removed from this revision, since its figures could not be traced back to the raw data in Table 2.

Final Centroid Values

Using the cluster memberships obtained in the first iteration (Table 4), the centroid of each cluster was recalculated as the average of the total score, number of projects, and win rate of its member records, producing the following updated centroid values

Table 5. Final centroid values

Cluster	Total Score	Total Project	Win Rate	Cluster
C1	100	30	33.33%	High Performer
C2	82.5	30	16.67%	Medium Performer
C3	45	25	0.00%	Low Performer

Based on The Results of The Grouping*:*

This manuscript reports the centroid update after a single iteration only. A full K-Means run would repeat the distance calculation and centroid update until the cluster memberships no longer change; because that second iteration was not performed here, the three-cluster solution above should be read as an illustrative first pass rather than a fully converged result.

1) The C1 cluster consists of the highest-performing salespeople.
2) The C2 cluster consists of sales with medium performance.
3) C3 consists of sales with low performance.

The results of the cluster are then used to calculate the new centroids based on the average values in each cluster.

Use Case Planning

The design of the system is illustrated through a Use Case Diagram which shows the interaction between actors and the features of the system. There are two main actors in this system, namely Sales and Marketing Managers and sales.

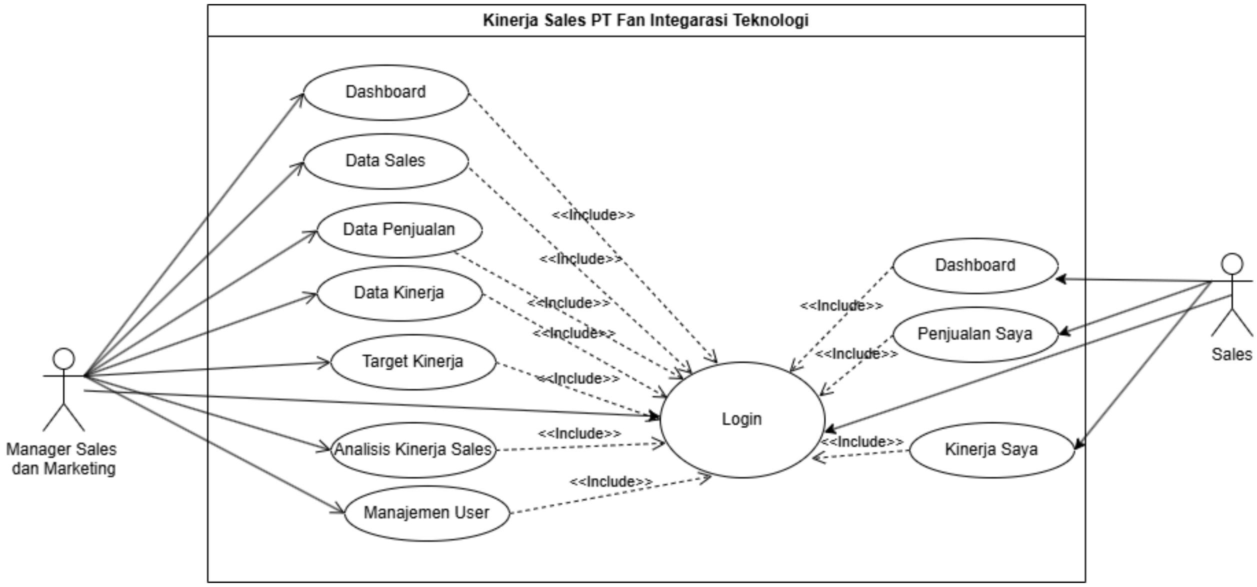

Figure 1. Use case diagram

Figure 3 shows the Use Case Diagram of PT FAN Integrasi Teknologi sales performance information system. Sales and Marketing managers have access to all system features, while Sales have limited access to the Dashboard, My Sales, and My Performance. All features can be accessed after the user logs in through the *include relationship*.

System Implementation

Results The implemented system developed is presented through the following interface view the system provides a performance analysis flow that includes determining the analysis period, creating a performance recap, grouping using K-means, and storing results.

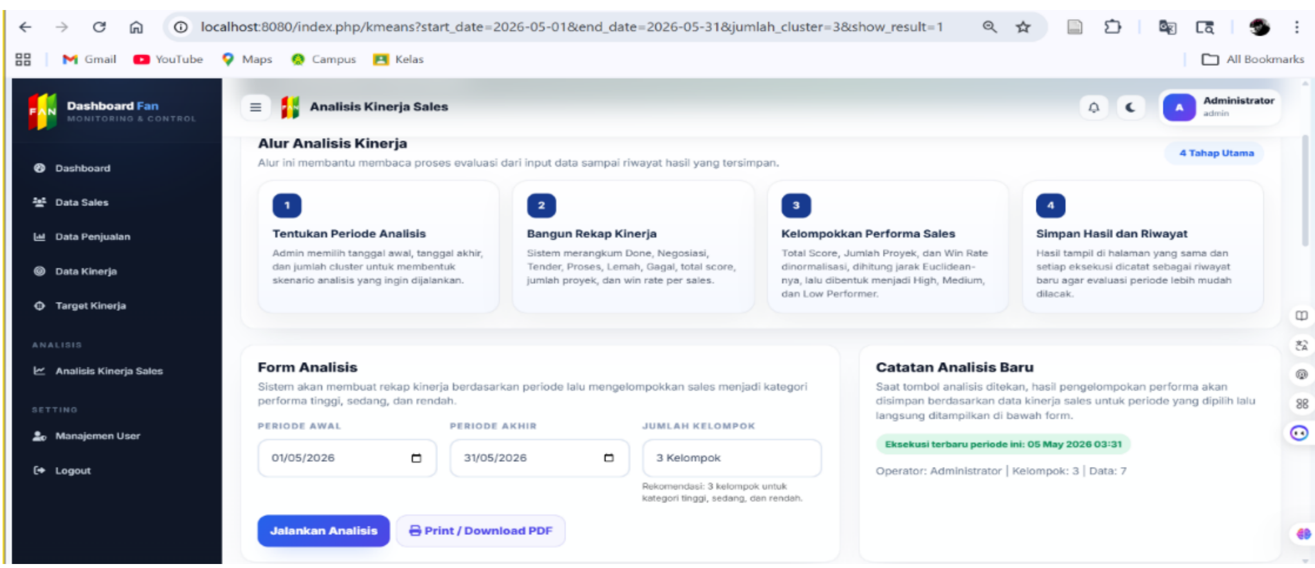

Figure 2. Sales performance analysis flow on the system

Figure 4 shows the flow of the sales performance analysis process in the system, which includes determining the analysis period, creating performance recaps, clustering using K-Means, and storing results.

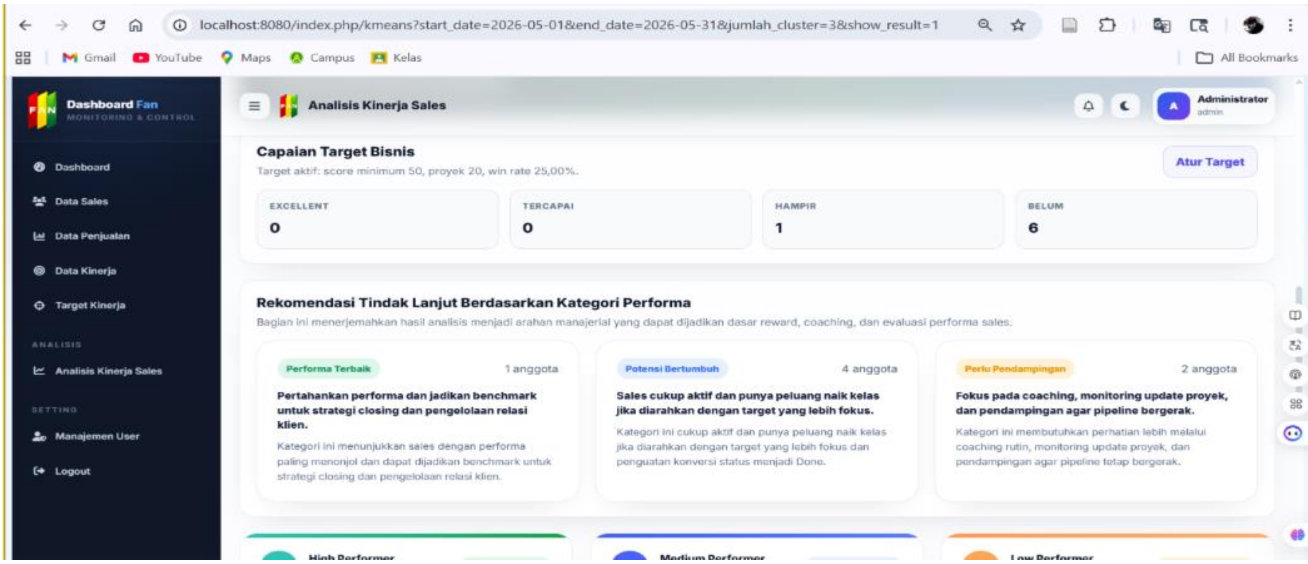

Figure 3. Results of sales performance grouping

The results of the analysis grouped sales performance into three categories, namely High Performer, Medium Performer, and Low Performer, based on the final score, total score, number of projects, and target achievement status. These results are consistent with previous studies indicating that K-Means Clustering can produce a similarity-based, exploratory grouping of performance data, rather than an independently validated or fully objective measure of performance. This research also integrates K-Means Clustering with the User-Centered Design (UCD) approach in a web-based system equipped with dashboards and follow-up recommendations to assist management in monitoring and determining performance improvement strategies.

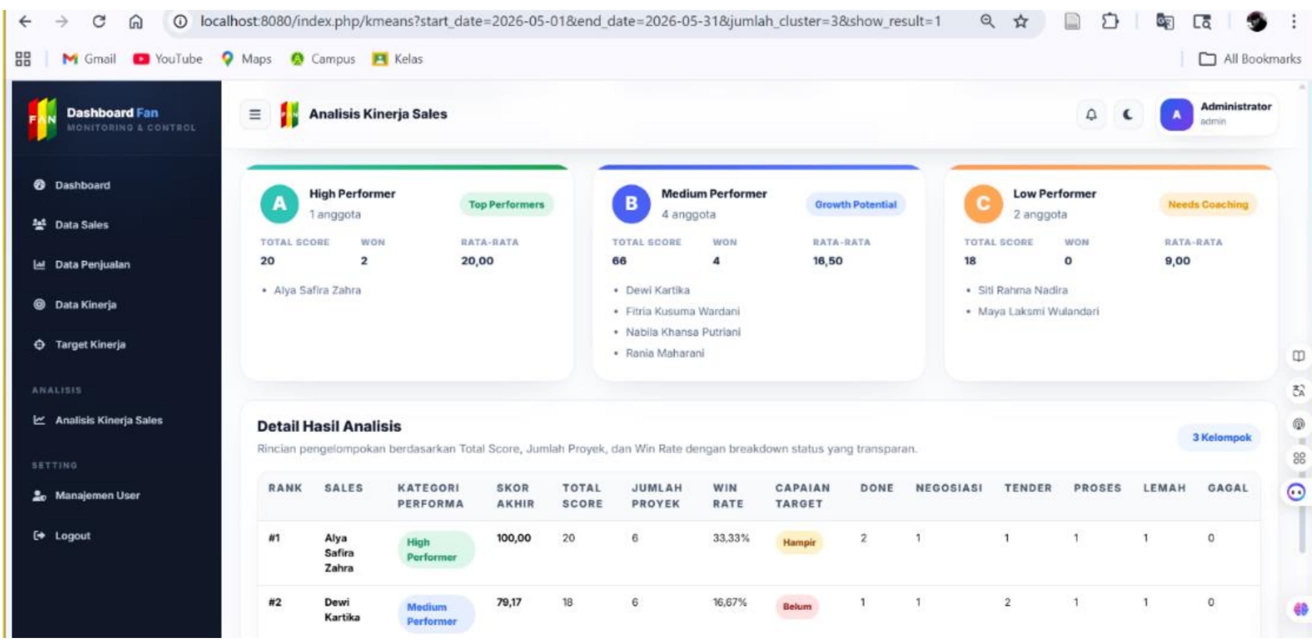

Figure 4. Details of sales performance analysis results

System testing

System testing is carried out using the Black Box Testing and White Box Testing methods to ensure that the system runs according to the needs of users [13]. Black Box Testing is carried out by testing the main functions of the system from the user side, including login, sales data management, performance data, performance targets, K-Means analysis, dashboard clustering, and report creation. Within the scenarios tested, these functions ran according to the expected results; a complete test-case log recording preconditions, inputs, and the pass or fail status of each individual scenario was not documented in this study.

White Box Testing is done by examining the structure and internal logic of the program on the main functions of the system. Testing is done to ensure each logical flow can run correctly. Based on the test results, the logical paths reviewed executed without errors; formal code-coverage metrics such as statement or branch coverage were not measured in this study[40].

Thus, within the scope of the scenarios examined, the Black Box and White Box testing indicate that the system's core functions behaved as expected. Broader verification, such as security testing, load and concurrent-user testing, and regression testing, was not carried out and remains a limitation of this study.

DISCUSSION

The decision to apply K-Means Clustering to sales performance data is consistent with a substantial body of recent work that uses the same algorithm to group employee and workforce records. Ibnu et al. [25] and Dewi et al. [26] each grouped staff performance quality into a small number of categories using K-Means, Murni et al. [27] applied the algorithm to employee discipline records at a village government office, and Rakhmawaty et al. [28] compared K-Means with Fuzzy C-Means for evaluating employee performance and discipline at a manufacturing company. More recently, Sakalessy and Purnomo [30] used K-Means to categorize information-technology staff performance at a mining company based on service-desk activity data. Across these studies, K-Means is consistently treated as an exploratory tool for surfacing similarity-based groups rather than a validated instrument for certifying individual performance, which is in line with the exploratory framing adopted for the High, Medium, and Low Performer categories in this study.

A parallel line of research applies K-Means specifically to sales and transaction data rather than to workforce records. Ridzki et al. [29] clustered best-selling product data at a grocery store chain, Al Rasyid et al. [37] grouped product sales at an online shop, and Yahya and Kurniawan [32] clustered sales transactions based on sales patterns at a retail store, reporting a Euclidean distance formulation consistent with the one used in this study. Chairunnita et al. [33], Miranda and Sriani [34], and Manarung et al. [35] each used K-Means to segment retail sales data for inventory or marketing decisions, while Deli et al. [36] combined K-Means with an R Shiny dashboard to segment marketing budgets and product sales in the retail sector. These studies confirm that K-Means is a common and practical choice for grouping sales-related data in small and medium enterprises, supporting the general approach taken in this study, although most of them worked with considerably larger transaction datasets than the seven sales records available here.

At the same time, the broader clustering literature raises concerns that are directly relevant to the limitations already noted in this study. In their comprehensive review of K-Means variants, Ikotun et al. [31] highlight that the algorithm's performance is highly sensitive to the number of clusters and the initial centroid selection specified by the user, and that determining an appropriate number of clusters for a given dataset remains an open challenge, particularly when that number is fixed in advance rather than derived from the data through methods such as the elbow or silhouette approach. This concern reinforces the position already taken in the Method and Conclusion sections of this study: because $k = 3$ was set to match the three predefined managerial categories rather than validated statistically, and because the clustering was carried out on only seven records, the resulting groups should be read as an illustrative, exploratory grouping rather than a definitive segmentation of sales performance.

On the interface-design side, the use of User-Centered Design in this study mirrors its use in other recent web-based sales systems. Amanda and Putri [38] applied UCD to design an e-commerce sales system, and Alja et al. [39] used the same approach for e-commerce UI/UX design, both reporting that involving users early in the requirements and design stages improved the perceived ease of use of the resulting system. The present study's use of UCD for the dashboard and reporting features is consistent with these findings, although, as noted in the Methodology section, this study interviewed only managers during the context and requirements stages, whereas a fuller application of UCD would also involve the sales staff who use the Dashboard, My Sales, and My Performance features directly.

Taken together, this comparison suggests that the system developed in this study follows established practice in combining UCD with K-Means Clustering for a business information system, and that its clustering results are broadly plausible when set against similar small-scale studies. Its main point of difference from the studies reviewed above is scale: most prior sales-clustering studies worked with dozens to thousands of transaction or employee records, whereas this study clustered only seven sales records from a single company. This difference reinforces, rather than contradicts, the limitations already identified in the Conclusion, and points to a clear direction for future work, namely validating the same UCD and K-Means approach on a larger, multi-period dataset before drawing firm conclusions about its effectiveness for sales performance evaluation.

CONCLUSION

A web-based sales performance information system was successfully developed using the User-Centered Design (UCD) approach at PT FAN Integrasi Teknologi. The system integrates sales data management, displays an interactive dashboard, and produces performance reports in PDF format. The application of K-Means Clustering grouped sales performance into three categories, High Performer, Medium Performer, and Low Performer, based on the variables of total score, number of projects, and win rate. Because the number of clusters (k = 3) was set in advance to match these three managerial categories, rather than determined through a statistical validation method such as the elbow or silhouette approach, and because the analysis was applied to only seven sales records from a single company, the resulting groups should be read as an exploratory, similarity-based grouping rather than an independently validated or fully objective measure of performance. Based on beta testing involving eight respondents, the system received an initial positive response, with an average satisfaction score of 84.5%; a fully documented instrument, a larger and more diverse respondent sample, and a clearer breakdown of respondent roles are needed before this figure can be treated as conclusive evidence of usability.

This study has several limitations that should be considered before wider use. The clustering analysis was based on only seven historical sales records from a single company, which is a very small sample for a three-cluster solution. The procedure used to normalize or scale the raw performance data before the distance calculation was not fully documented and should be clarified and made reproducible in future work. Only one iteration of the K-Means algorithm was reported, so full convergence of the clustering was not verified. The beta test involved a small number of respondents using an instrument that was not described in detail. Because the resulting performance categories could influence how management perceives an individual salesperson, they should be used only as a supporting, human-reviewed input to performance discussions, not as the sole or automatic basis for reward, training, promotion, or disciplinary decisions, and employees should be able to view and question the data and classification used in their own evaluation.

The system is best positioned as a company-specific prototype with a preliminary user evaluation. Broader claims of operational readiness should wait until data quality, cluster validity, fairness, explainability, security, and role-specific usability have been further demonstrated.

Further development can add an automatic notification feature when sales performance is below target and improve security and user access rights. Subsequent research can also compare other clustering methods, such as K-Medoids, and add assessment variables such as customer satisfaction levels to make the evaluation results more comprehensive, validate the choice of k using statistical methods such as the elbow or silhouette approach, run the K-Means process to full convergence across multiple iterations, and evaluate the system on a larger, multi-period dataset covering more than one company.

REFERENCES

[1] M. S. E. Kasem, M. Hamada, and I. Taj-Eddin, "Customer Profiling, Segmentation, and Sales Prediction Using Artificial Intelligence," *Neural Comput. Appl.*, 2024, doi: 10.1007/s00521-023-09339-6.

[2] H. Syahputra, "Clustering Tingkat Penjualan Menu Food and Beverage Menggunakan Algoritma K-Means," *J. KomtekInfo*, vol. 9, no. 1, pp. 29–33, 2022, doi: 10.35134/komtekinfo.v9i1.274.

[3] X. Yan, "Bank Customer Segmentation and Marketing Strategies Based on Machine Learning," *Appl. Sci.*, vol. 15, no. 6, p. 3138, 2025, doi: 10.3390/app15063138.

[4] F. J. Seagull, M. S. Lanham, M. Pomorski, M. Callahan, E. K. Jones, and G. D. Barnes, "Implementing Evidence-Based Anticoagulant Prescribing: User-Centered Design Findings and Recommendations," *Res. Pract. Thromb. Haemost.*, vol. 6, no. 6, p. e12803, 2022, doi: 10.1002/rth2.12803.

[5] H. J. Wilbert, "Recency, Frequency, Monetary Value, Clustering, and Customer Behavior Analysis," *Algorithms*, vol. 16, no. 9, p. 396, 2023, doi: 10.3390/a16090396.

[6] Y. Sonatha, A. Erianda, and R. Fitri, "Sales Segmentation Analysis of Tobacco Products Using the K-Means Clustering Method," *INVOTEK J. Inov. Vokasional dan Teknol.*, vol. 24, no. 2, 2024, doi: 10.24036/invotek.v24i2.1221.

[7] D. A. Titania, L. Kurniawati, and T. Haryanti, "Designing the UI/UX of a Document Archiving Information System Using the User-Centered Design Method," *METIK J.*, vol. 8, no. 1, 2024, doi: 10.47002/metik.v8i1.686.

[8] K. Tabianan, S. Velu, and V. Ravi, "K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data," *Sustainability*, vol. 14, no. 12, p. 7243, 2022, doi: 10.3390/su14127243.

[9] A. Ghosh and A. Roy, "Comparative Study of K-Means, Bisecting K-Means, and DBSCAN for Customer Segmentation," *J. Intell. Fuzzy Syst.*, vol. 45, no. 4, pp. 5321–5332, 2023, doi: 10.3233/JIFS-212245.

[10] L. L. Jauculan and U. B. Patayon, "Enhancing UX/UI: A Mixed-Approach Evaluation of a Web-Based Student

Clearance System at a State University in the Philippines," *Procedia Comput. Sci.*, vol. 234, pp. 1061–1068, 2024, doi: 10.1016/j.procs.2024.03.100.

[11] K. Chowdhary *et al.*, "User-Centered Design to Enhance mHealth Systems for Individuals With Dexterity Impairments: Accessibility and Usability Study," *JMIR Hum. Factors*, vol. 9, no. 1, p. e23794, 2022, doi: 10.2196/23794.

[12] L. Xiao, "GMS-AFKMC2: A New Customer Segmentation Framework Based on Feature Selection and K-Means Clustering," *Electronics*, vol. 13, no. 17, p. 3523, 2024, doi: 10.3390/electronics13173523.

[13] M. Rochmawati, "Implementation of the K-Means Algorithm in Sales Clustering at a Company Using the KDD Methodology," *SISTEMASI*, vol. 13, no. 1, 2024, doi: 10.32520/stmsi.v13i1.3074.

[14] S. N. Lathifah and Z. F. Azzahra, "AI-Driven Customers Segmentation Using K-Means Clustering," *G-Tech J. Teknol. Terap.*, vol. 9, no. 1, 2025, doi: 10.70609/gtech.v9i1.6202.

[15] C. P. Ufeli, "Customer Segmentation Based Approach Using Factor Analysis, K-Means, and Hierarchical Clustering," *Information*, vol. 16, no. 6, p. 441, 2025, doi: 10.3390/info16060441.

[16] M. Puspita Hannah and R. Nur Kholiza, "Perancangan UI/UX Penerimaan Peserta Didik Baru Berbasis Website Menggunakan Metode User Centered Design," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 4, pp. 2342–2356, 2024, doi: 10.30865/klik.v4i4.1722.

[17] B. Mostefai, "Enhancing User Experience in E-Learning Systems: A New Evaluation and User-Centered Design Approach," *Comput. Hum. Behav. Reports*, p. 100643, 2025, doi: 10.1016/j.chbr.2025.100643.

[18] M. Miranda, "Implementation of K-Means Clustering in Grouping Sales Data at Zura Mart," *J. Appl. Informatics Comput.*, 2025, [Online]. Available: https://jurnal.polibatam.ac.id/index.php/JAIC/article/view/9160

[19] Y. Zhao, H. Lin, and W. Zhang, "RFM-Based Clustering and Product Recommendation Using Association Rules," *Expert Syst. Appl.*, vol. 198, p. 116899, 2022, doi: 10.1016/j.eswa.2022.116899.

[20] J. M. John, O. Shobayo, and B. Ogunleye, "An Exploration of Clustering Algorithms for Customer Segmentation in the UK Retail Market," *Analytics*, vol. 2, no. 4, pp. 809–823, 2023, doi: 10.3390/analytics2040042.

[21] A. Twomlow, S. Grainger, K. Cieslik, J. D. Paul, and W. Buytaert, "A User-Centred Design Framework for Disaster Risk Visualisation," *Int. J. Disaster Risk Reduct.*, vol. 77, p. 103067, 2022, doi: 10.1016/j.ijdrr.2022.103067.

[22] S. S. Groos, A. J. Linn, J. I. Kuiper, N. M. van Schoor, N. van der Velde, and J. C. M. van Weert, "Combining User-Centered Design and Behavioral Theory to Enhance Health Technologies: A Personas-Based Approach for a Primary-Care-Based Multifactorial Falls Risk Assessment Tool," *Int. J. Med. Inform.*, vol. 186, p. 105420, 2024, doi: 10.1016/j.ijmedinf.2024.105420.

[23] M. Deldadehasl, "A Hybrid Data Mining and Decision-Making Approach for Customer Segmentation," *Businesses*, vol. 6, no. 2, p. 80, 2025, doi: 10.3390/businesses6020080.

[24] H. Alamsyah, "Customer Segmentation Using RFM and K-Means Clustering," *J. Theor. Appl. Inf. Technol.*, vol. 100, no. 5, pp. 1301–1310, 2022.

[25] A. Ullah, "Customer Analysis Using Machine-Learning-Based Methods for Effective Segmentation," *Sensors*, vol. 23, no. 6, p. 3180, 2023, doi: 10.3390/s23063180.

[26] C. Wongoutong, "The Impact of Neglecting Feature Scaling in K-Means Clustering," *PLoS One*, vol. 19, no. 12, 2024, [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11623793/

[27] G. D. G. E. Sidabutar, J. A. Seah, and Y. A. Singgalen, "Analysis and Design of Web-Based Information System for Coffeeshop Management Using Design Thinking Methodology," *J. Inf. Syst. Informatics*, vol. 5, no. 1, pp. 217–231, 2023, doi: 10.51519/journalisi.v5i1.455.

[28] H. Saputra, M. Algipani, A. N. Mulya, and H. F. Zafira, "UI/UX Design Using User-Centered Design Method on Staycation Website," in *2024 International Conference on Information Technology Research and Innovation*, 2024. doi: 10.1109/INCITEST64888.2024.11121477.

[29] S. N. A. Zahra and S. Suryatiningsih, "Evaluation and Improvement of User Interface Design of Bandung City APBD Website Using Human Centered Design Method," *Procedia Comput. Sci.*, vol. 234, pp. 1791–1798, 2024, doi: 10.1016/j.procs.2024.03.187.

[30] P. Anitha and M. M. Patil, "RFM Model for Customer Purchase Behavior Using K-Means Algorithm," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 5, pp. 1785–1792, 2022, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1319157819309802

[31] C. G. Wong, T. Nguyen, and K. Lee, "Application of Hierarchical and K-Means Clustering in E-Commerce Customer Segmentation," *Int. J. Inf. Manag. Data Insights*, vol. 2, no. 1, p. 100055, 2022, doi: 10.1016/j.jjimei.2021.100055.

[32] N. Kumar, "Intelligent Customer Segmentation: Unveiling Consumer Patterns with Machine Learning," *J. Umm Al-Qura Univ. Eng. Archit.*, 2025, doi: 10.1007/s43995-025-00180-7.

[33] T. Ozcan, "Customer Segmentation Using an Extended RFM Model and Machine-Learning Methods," *J. Theor. Appl. Electron. Commer. Res.*, vol. 21, no. 5, p. 142, 2026, doi: 10.3390/jtaer21050142.

[34] J. Salminen, M. Mustak, M. Sufyan, and B. J. Jansen, "How Can Algorithms Help in Segmenting Users and Customers? A Systematic Review and Research Agenda for Algorithmic Customer Segmentation," *J. Mark. Anal.*, vol. 11, no. 4, pp. 677–692, 2023, doi: 10.1057/s41270-023-00235-5.

[35] R. I. Manarung, E. Widodo, and A. M. Rifai, "Sales Data Clustering Using the K-Means Algorithm to Determine Retail Product Needs," *Int. J. Softw. Eng. Comput. Sci.*, vol. 5, no. 1, pp. 226–234, 2025, doi: 10.35870/ijsecs.v5i1.4090.

[36] A. A. Yunanto, "Design and Implementation of the Prayer Reminder Application Using KISS Principle Based on User Centered Design," *Procedia Comput. Sci.*, vol. 234, pp. 1484–1491, 2024, doi: 10.1016/j.procs.2024.03.149.

[37] A. Mulyawan, A. Bahtiar, and I. Ali, "Transformasi Strategi Penjualan Batik Cirebon dengan Pendekatan Analisis Pengelompokan K-Means," *KOPERTIP J. Ilm. Manaj. Inform. dan Komput.*, vol. 7, no. 1, pp. 1–7, 2023, doi: 10.32485/kopertip.v7i1.309.

[38] Y. Yovian and N. R. Pratama, "Examining The Impact of E-Service Quality and E-Recovery Service Quality in Digital Public Services in Indonesia," *J. World Sci.*, vol. 4, no. 6, pp. 754–768, 2025, doi: 10.58344/jws.v4i6.1426.

[39] I. Sarifah *et al.*, "Development of Android Based Educational Games to Enhance Elementary School Student Interests in Learning Mathematics," *Int. J. Interact. Mob. Technol.*, vol. 16, no. 18, pp. 149–161, 2022, doi: 10.3991/ijim.v16i18.32949.

[40] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf. Sci. (Ny).*, vol. 622, pp. 178–210, 2023, doi: 10.1016/j.ins.2022.11.139.