
Enhancing Decision-Making in Local Government through K-Means Clustering of Structural Official's Performance

Adrian Unggul Wirawan¹, Hardika Khusnuliawati^{2*}, Astri Charolina³, Anniez Rachmawati Musslifah⁴,
Rusnandari Retno Cahyani⁵

^{1,2,3,4,5} Sahid Surakarta University

^{1,2,3,4,5} Surakarta, Jawa Tengah 57144

*hardika.khusnulia@gmail.com

Abstract — Employee performance evaluation is a critical process in public sector management. However, in Karanganyar Regency, this process has been traditionally conducted on an individual basis, leading to inefficiencies and a lack of actionable insights. This study addresses the gap by applying the K-Means clustering algorithm to categorize the performance of structural officials based on 2021 Employee Performance Target (SKP) data. Key performance indicators include SKP Value, Service Orientation, Commitment, Cooperation, Leadership, and Work Initiative. Using RapidMiner, the data was clustered into three categories: “very good,” “good,” and “satisfactory.” The clustering quality was validated using the Davies-Bouldin Index (DBI), achieving an optimal value of 0.113, which indicates high intra-cluster similarity. The results provide a data-driven foundation for more efficient performance assessments, aiding decision-making in promotions and personnel management. This study demonstrates the potential of machine learning, specifically K-Means clustering, in improving administrative processes and strategic decision-making within local government.

Keywords – K-Means clustering, performance assessment, machine learning, Davies-Bouldin Index (DBI), employee evaluation

I. INTRODUCTION

Performance evaluation is a systematic assessment of employees' efforts and achievements in carrying out their duties to achieve organizational goals. It plays a crucial role in public services to ensure effective performance aligned with national objectives [1]. The importance of performance evaluation lies in its ability to enhance employee performance, career development, and job satisfaction [2].

The Agency for Personnel and Human Resource Development of Karanganyar Regency is a regional institution responsible for managing personnel affairs, including Civil Servants (PNS) and Government Employees under Work Agreements (PPPK) within the region. One of its main duties is to collect performance evaluation data from all State Civil Apparatus (ASN) in Karanganyar Regency, which is subsequently reported to the National Civil Service Agency. However, the current evaluation process is still conducted individually without utilizing analytical technology that could provide deeper insights.

Artificial intelligence (AI), particularly through machine learning approaches, is becoming increasingly important in the public sector, as it can improve operational efficiency, public service quality, as well as compliance and security aspects [3]. A study

of 250 AI application cases in the European Union revealed that the most significant impact occurred in public service delivery, while its application in internal management and policy decision-making remains limited [4].

One of the most widely used machine learning algorithms across various domains is K-Means clustering, due to its ability to group data based on similar characteristics, thus facilitating pattern analysis and decision-making [5]. K-Means has been successfully implemented in several sectors, such as e-learning for classifying student engagement levels [6], lecturer performance evaluation based on student satisfaction [7], and agriculture for clustering corn production areas based on yield levels [8]. In addition, K-Means has been used to cluster individual performance based on response speed and accuracy [9], group provinces according to COVID-19 pandemic risk levels [10], analyze customer behavior in the retail industry using the RFM model [11], and in maritime safety for ship classification based on inspection data [12]. In the education sector, the algorithm has also been applied to modernize teacher performance evaluation through web-based systems, making the evaluation process more efficient [13].

However, to date, the application of clustering methods, particularly the K-Means algorithm, has not

been utilized in the context of ASN performance evaluation within the Karanganyar Regency Government. The potential of clustering is substantial, especially in identifying hidden patterns, grouping employees based on performance quality, and simplifying the evaluation process, which has traditionally been time-consuming and subjective.

Therefore, this study aims to apply the K-Means algorithm in clustering performance evaluation data of structural officials in Karanganyar Regency. This approach not only offers efficiency in the evaluation process but also provides a more objective and data-driven overview of performance categories—namely excellent, good, and adequate. The clustering results are expected to serve as a strategic consideration for regional leaders in making decisions related to promotions, rotations, and employee development.

II. RESEARCH METHOD

A. Analysis Model

This section presents the overall system process, as illustrated in the block diagram shown in Fig. 1.

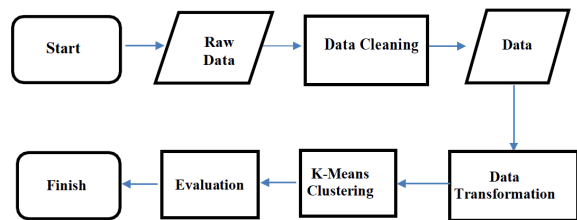


Fig. 1. Block diagram of the clustering analysis process.

The clustering process begins with raw data, namely the Employee Performance Targets (Sasaran Kinerja Pegawai / SKP) of structural officials in Excel format, which contains various performance metrics. The first stage is data cleaning, which removes inconsistencies and errors to produce a refined dataset ready for further processing.

The next stage is data transformation, where alphanumeric identifiers, such as the names of structural officials, are converted into numerical format. The third stage involves applying the K-Means clustering algorithm to categorize the data into three distinct performance groups: Excellent, Good, and Fair.

Subsequently, the clustering results are analyzed and evaluated using the Davies–Bouldin Index (DBI) to measure the accuracy of the clustering process. The DBI functions as a cluster performance evaluation metric, where a positive correlation indicates within-class dispersion, while a negative correlation reflects between-class separation [14]. This study aims to determine the optimal cluster value using the Davies–Bouldin Index (DBI), in which a smaller DBI value indicates better cluster formation [15].

The results of the analysis are then submitted to the Agency for Personnel and Human Resource Development (BKPSDM) of the Karanganyar

Regency Government as valuable input for assessing the performance of structural officials.

B. Data Analysis

This study focuses on the clustering process of structural officials' performance evaluation results using primary data from the Employee Performance Targets (SKP) of structural officials in 2021, comprising a total of 460 entries. The data were obtained directly from the Agency for Personnel and Human Resource Development of Karanganyar Regency through the Personnel Information System (SIMPeg) application, specifically within the e-SKP (Employee Performance Targets) module.

The dataset includes key performance attributes such as SKP Score, Service Orientation, Commitment, Teamwork, Leadership, and Work Initiative. An example of the dataset used in this study is shown in Table 1.

Table 1. The first 10 rows of employee performance evaluation data

ID	SKP Score	Service Orientation	Commitment	Teamwork	Leadership	Work Initiative
001	109	120	120	120	120	120
002	111,79	109	109	109	109	109
003	109	109	109	109	109	109
004	109	109	109	109	109	109
005	104	109	109	109	109	109
006	100	109	109	109	109	109
007	104,31	101,4	97,6	93,8	105,2	105,2
008	100	109	109	109	109	109
009	100	109	105,2	105,2	105,2	105,2
010	109	109	117,8	109	117,8	117,8

III. RESULT

This section presents the findings from the clustering analysis conducted on the performance data of structural officials. The K-Means algorithm was employed to group the dataset into clusters that reflect different levels of performance. The objective of the clustering process was to classify officials into three categories—Excellent, Good, and Fair—based on various performance metrics such as SKP Score, Service Orientation, Commitment, Teamwork, Leadership, and Initiative.

The evaluation of the clustering results was carried out using the Davies–Bouldin Index (DBI) to assess the effectiveness of the clustering approach. A lower DBI value indicates better separation between clusters and higher coherence within clusters, signifying superior clustering quality. In addition to numerical evaluation, the characteristics of each cluster were analyzed to identify patterns across performance attributes, allowing for deeper insights into the factors influencing officials' performance levels.

A. Data Processing Using RapidMiner Studio Educational 10.0

To perform clustering in RapidMiner, particularly with the K-Means algorithm, the first step is to open RapidMiner Studio and import the dataset. RapidMiner is a flexible data mining software that can function both as a standalone application and as an

integrated data mining engine within other products [16]. The software provides a comprehensive suite of features, including data integration, analytical ETL processes, data analysis, and reporting capabilities.

Once the data are successfully loaded, the K-Means operator can be located in the Operators panel. Drag the operator into the Process window, then connect the output of the data reader to the input of the K-Means operator. Next, link the output of the K-Means operator to the “Results” output port. Configure the K-Means operator by specifying the desired number of clusters (for instance, three clusters representing the categories Excellent, Good, and Fair).

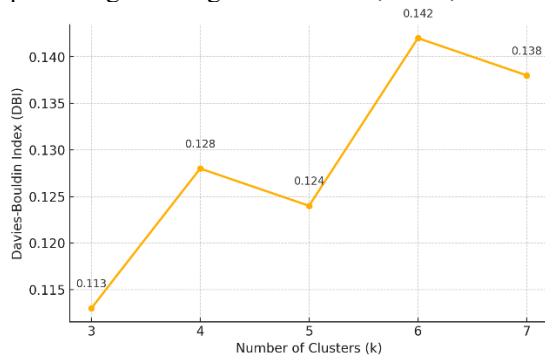


Fig. 3. Relationship between the number of clusters (k) and the Davies–Bouldin Index (DBI)

From each tested k value, it was found that the optimal value among the five options was k = 3, where the dataset was divided into three clusters with a Davies–Bouldin Index (DBI) value of 0.113. As illustrated in Fig. 3, the line diagram of DBI values across different k values (k = 3 to 7) shows a non-linear relationship. The DBI initially increases after k = 3, indicating that clustering quality decreases as the number of clusters grows. Although the average within-cluster distance becomes smaller with higher k, the increase in DBI reflects more overlap and less distinct separation among clusters. Therefore, k = 3 provides the most balanced and well-separated grouping structure.

To identify which individuals belong to each of the three clusters, as well as the number of data points in clusters 0, 1, and 2, the “Description” function was used. The results showed that the first cluster contained 134 records (31.38%), the second cluster contained 275 records (64.40%), and the third cluster contained 18 records (4.22%). To view the names of the officials in each cluster, the “Folder View” option was selected, displaying the corresponding names.

B. Cluster Analysis Results

The clustering results of the 2021 performance evaluation data for structural officials within the Government of Karanganyar Regency, obtained using the K-Means algorithm, are presented in Fig. 4. These results are visualized in the form of a graphical plot.

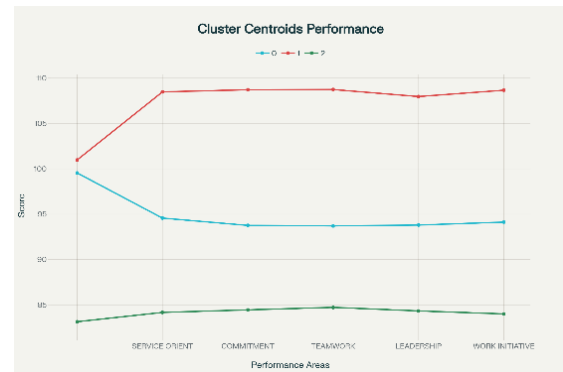


Fig. 4. Visualization of clustering results using the K-Means algorithm on the performance evaluation dataset.

From Fig. 4, it can be seen that the data distribution in the first cluster is represented by the color blue, the second cluster by green, and the third cluster by orange.

The explanation for each cluster is as follows:

- Cluster 1, with its centroid values positioned between those of the second and third clusters, represents the intermediate group or the “Good” category. This cluster contains 134 data points with centroid values of (99.545, 94.561, 93.757, 93.702, 93.796, 94.116).
- Cluster 2 has the highest centroid values among all clusters and represents the “Excellent” category. This group consists of 275 data points, with centroid values of (100.978, 108.478, 108.712, 108.737, 107.961, 108.659). This indicates that the cluster corresponds to the group with the highest performance.
- Cluster 3, which has the lowest centroid values, is categorized as “Fair”, indicating the group with the lowest performance level. This cluster contains 18 data points, with centroid values of (83.139, 84.167, 84.444, 84.722, 84.333, 84.000).

The data were successfully divided into three clusters that reflect varying levels of performance. The largest cluster, representing 64.41% of the data (275 records), falls into the “Excellent” category, as it has the highest centroid values and represents the group with superior performance. The “Good” category includes 134 records (31.38%), with centroid values lying between the other two clusters, indicating moderate performance. Meanwhile, the smallest cluster, consisting of only 18 records (4.21%), falls into the “Fair” category, with the lowest centroid values, reflecting the lowest performance among the three groups.

The analysis for each cluster was performed with a maximum of 10, 50, 100, and 500 iterations, resulting in an optimal Davies–Bouldin Index (DBI) value of 0.113. Based on this DBI value, it can be concluded that the intra-cluster similarity is

approximately 89%, indicating a strong degree of cohesion within clusters.

IV. DISCUSSION

The clustering results demonstrate that the K-Means algorithm effectively categorized officials into distinct performance levels with a high degree of intra-cluster similarity (approximately 89%). This suggests that officials within the same cluster share consistent performance characteristics, while differentiation between clusters remains clearly defined.

The centroid values presented in Table 2 further support this observation. The “Excellent” cluster (Cluster 2) exhibits the highest centroid scores across all performance indicators, particularly in Commitment (108.712), Teamwork (108.737), and Initiative (108.659), signifying strong interpersonal collaboration and proactive work behavior. The “Good” cluster (Cluster 1) maintains stable but moderate centroid values—ranging from 93.702 to 99.545—indicating consistent yet less outstanding performance across key areas. Conversely, the “Fair” cluster (Cluster 3) shows noticeably lower centroid values—between 83.139 and 84.722—highlighting potential weaknesses in all evaluated dimensions, especially in Initiative (84.000) and Service Orientation (84.167).

Table 2. Centroid Values of Each Cluster

Centroid	SKP Score	Service Orientation	Commitment	Teamwork	Leadership	Work Initiative
1	99.545	94.561	93.757	93.702	93.796	94.116
2	100.978	108.478	108.712	108.737	107.961	108.659
3	83.139	84.167	84.444	84.722	84.333	84.000

These centroid differences quantitatively confirm that each cluster represents a unique performance tier, with clear performance gradients separating one group from another. The “Excellent” cluster (64.40%) dominates the dataset, indicating that the majority of officials perform at a high level across all indicators. The “Good” cluster (31.38%) represents officials with moderate and consistent performance, whereas the “Fair” cluster (4.22%) highlights individuals who may require further evaluation or performance improvement initiatives.

These insights can be leveraged by the Agency for Personnel and Human Resource Development (BKPSDM) and the Position and Rank Advisory Board (Baperjakat) to support data-driven decision-making in promotion and performance evaluation processes. By adopting clustering-based analysis, the evaluation process becomes more objective, transparent, and evidence-based, offering a reliable alternative to conventional assessment methods.

V. CONCLUSION

The application of the K-Means algorithm proved effective in clustering the performance data of structural officials in Karanganyar Regency, providing significant insights into employee

performance levels. The clustering results divided the officials into three categories: Excellent (275 officials, 64.41%), Good (134 officials, 31.38%), and Fair (18 officials, 4.21%). These results indicate that most officials fall into the “Excellent” performance category, while only a small proportion belong to the “Fair” group. The clustering process was validated using the Davies–Bouldin Index (DBI), which achieved an optimal value of 0.113, signifying high intra-cluster similarity and well-formed cluster structures.

This method offers a clearer and data-driven alternative to traditional individual performance assessments, reducing evaluation time while improving accuracy. The clustering analysis enables local government leaders to make more precise decisions regarding job promotions and human resource management. Overall, this study highlights the potential of machine learning techniques, such as K-Means clustering, in enhancing efficiency and decision-making processes within public administration.

The clustering results are expected not only to serve as an evaluation tool but also as a foundation for career development planning, position rotation, and targeted training policies. The local government may consider developing an integrated personnel information system that incorporates the K-Means algorithm to enable automated and periodic performance evaluations.

Moreover, to maintain the accuracy of the results, each work unit must ensure the quality and completeness of SKP data submitted for analysis, as the success of this method highly depends on data validity. Training sessions for policymakers and relevant staff should also be conducted to help them understand and appropriately utilize the clustering outcomes in the decision-making process. For future development, further research could explore the use of other clustering algorithms, such as DBSCAN or Hierarchical Clustering, and integrate predictive approaches to achieve a more comprehensive assessment of long-term performance.

REFERENCES

- [1] Y. Dawanaka, W. Waworundeng, and A. Purwanto, “Employee Performance in Improving Population Administration Services in the Population and Civil Registration Office of East Halmahera Regency,” *J. La Bisecoman*, vol. 3, no. 6, pp. 271–277, 2023, doi: 10.37899/journallabisecoman.v3i6.928.
- [2] D. Podungge, I. Mashudi, and K. Napu, “Analysis of Performance Assessment System Model of Civil Servants in Gorontalo Province Training and Education Agency,” *J. La Soc.*, vol. 1, no. 4, pp. 27–32, 2020, doi: 10.37899/journal-la-sociale.v1i4.143.
- [3] P. Henman, “Improving public services using artificial intelligence: possibilities, pitfalls, governance,” *Asia Pacific J. Public Adm.*, vol. 42, no. 4, pp. 209–221, 2020, doi: 10.1080/23276665.2020.1816188.
- [4] C. van Noordt and G. Misuraca, “Artificial intelligence for the public sector: results of landscaping the use of

- AI in government across the European Union,” *Gov. Inf. Q.*, vol. 39, no. 3, p. 101714, Jul. 2022, doi: 10.1016/j.giq.2022.101714.
- [5] I. H. Sarker, “Machine Learning: Algorithms, Real-World Applications and Research Directions,” *SN Comput. Sci.*, vol. 2, no. 3, p. 160, May 2021, doi: 10.1007/s42979-021-00592-x.
- [6] A. Moubayed, M. Injadat, A. Shami, and H. Lutfiyya, “Student Engagement Level in an e-Learning Environment: Clustering Using K-means,” *Am. J. Distance Educ.*, vol. 34, no. 2, pp. 137–156, 2020, doi: 10.1080/08923647.2020.1696140.
- [7] M. Mala Khairunnisa, A. Triayudi, and E. Tri Esti Handayani, “Application of K-Means Clustering on the Performance Evaluation of Lecturers Based on Student Questionnaire,” *J. Mantik*, vol. 4, no. 1, pp. 760–766, 2020, [Online]. Available: <https://iocscience.org/ejournal/index.php/mantik>
- [8] A. A. Aldino, D. Darwis, A. T. Prastowo, and C. Sujana, “Implementation of K-Means Algorithm for Clustering Corn Planting Feasibility Area in South Lampung Regency,” *J. Phys. Conf. Ser.*, vol. 1751, no. 1, 2021, doi: 10.1088/1742-6596/1751/1/012038.
- [9] R. Vankayalapati, K. B. Ghutugade, R. Vannapuram, and B. P. S. Prasanna, “K-means algorithm for clustering of learners performance levels using machine learning techniques,” *Rev. d'Intelligence Artif.*, vol. 35, no. 1, pp. 99–104, 2021, doi: 10.18280/ria.350112.
- [10] D. Abdullah, S. Susilo, A. S. Ahmar, R. Rusli, and R. Hidayat, “The application of K-means clustering for province clustering in Indonesia of the risk of the COVID-19 pandemic based on COVID-19 data,” *Qual. Quant.*, vol. 56, no. 3, pp. 1283–1291, 2022, doi: 10.1007/s11135-021-01176-w.
- [11] 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, doi: 10.1016/j.jksuci.2019.12.011.
- [12] Z. Hou, R. Yan, and S. Wang, “On the K-Means Clustering Model for Performance Enhancement of Port State Control,” *J. Mar. Sci. Eng.*, vol. 10, no. 11, 2022, doi: 10.3390/jmse10111608.
- [13] B. Firmansyah and U. Chotijah, “Implementation of The K-Means Clustering for Teacher Performance Assessment Grouping (PKG) at MI Bani Hasyim Cerme,” *J. Ris. Inform.*, vol. 5, no. 1, pp. 499–506, 2022, doi: 10.34288/jri.v5i1.475.
- [14] D. A. Tarigan, “Optimization of the K-Means Clustering Algorithm Using Davies Bouldin Index in Iris Data Classification,” *Media Online*, vol. 4, no. 1, pp. 545–552, 2023, doi: 10.30865/klik.v4i1.964.
- [15] A. Idrus, N. Tarihoran, U. Supriatna, A. Tohir, S. Suwarni, and R. Rahim, “Distance Analysis Measuring for Clustering using K-Means and Davies Bouldin Index Algorithm,” *TEM J.*, vol. 11, no. 4, pp. 1871–1876, 2022, doi: 10.18421/TEM114-55.
- [16] R. Oktarina and Junita, “Determine the clustering of cities in Indonesia for disaster management using K-Means by excel and RapidMiner,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 794, no. 1, 2021, doi: 10.1088/1755-1315/794/1/012094.