

Application Of K-Means Algorithm For Performance Segmentation Of P2TL Officers As Decision Support At PT PLN (Persero) UP3 Kebon Jeruk

Widi Handoko¹, Arief Wibowo²

¹Master of Management Study Program, Budi Luhur University

²Lecturer, Master of Management Study Program, Budi Luhur University

Email: 2531600498@student.budiluhur.ac.id, arief.wibowo@budiluhur.ac.id

Abstract

PT PLN (Persero) as the national electricity provider has a Electricity Usage Control (P2TL) activity that plays an important role in controlling violations of electricity use and minimizing potential company losses. In its implementation, the performance of P2TL officers shows different levels of productivity based on the number of inspections, P2TL findings, and performance assessments. However, the conventional performance evaluation process has prevented officer performance segmentation from being carried out optimally as a basis for managerial decision-making. This study aims to apply the K-Means algorithm to segment the performance of P2TL officers at PT PLN (Persero) UP3 Kebon Jeruk. The study uses a quantitative approach with the data mining method. Variables used include the number of inspections, P2TL findings, and officer performance assessments. Data processing was carried out using RapidMiner Studio through preprocessing, normalization, clustering, and cluster quality evaluation using the Davies Bouldin Index (DBI).

The results show that the K-Means algorithm successfully grouped officers into three main clusters: high-performance, medium-performance, and low-performance clusters. Evaluation using the Davies Bouldin Index (DBI) obtained the best value of - 0.609 at K=3, making three clusters the most optimal number for this study. The resulting segmentation can provide objective grouping of officers based on their performance characteristics. These findings can be used as a basis for developing coaching strategies, competency development, supervision, and more effective, measurable, and data-driven improvement of P2TL officer performance.

502

IJEBIR, Volume 05 Issue 04, 2026



Copyright at authors some right reserved this work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

Keywords: *K-Means, Clustering, Data Mining, Performance Segmentation, P2TL.*

INTRODUCTION

The development of information technology and data digitalization encourages organizations to adopt data-driven approaches to support decision-making processes. Optimal use of data can help organizations improve operational effectiveness, work efficiency, and the quality of supervision of human resources. One approach widely used in data processing is data mining, which is the process of extracting information and hidden patterns from large amounts of data to produce useful information for organizations (Han, Kamber, & Pei, 2012).

In its implementation, data mining has various analytical techniques, one of which is clustering. The clustering technique is used to group data based on the degree of similarity of certain characteristics so that objects within one group have relatively homogeneous characteristics compared to other groups. One of the most widely used clustering methods is the K-Means algorithm because it has a simple computational process, is easy to implement, and is capable of producing effective data grouping across various research fields (Tan, Steinbach, & Kumar, 2016).

From a human resource management perspective, performance management is one of the important aspects in supporting organizational goal achievement. Systematically conducted performance evaluations can help organizations identify levels of productivity, work effectiveness, and individual contributions toward organizational goals. According to Aguinis (2019), performance management is a continuous process aimed at identifying, measuring, and developing individual performance to align with organizational goals. In line with this, Armstrong (2021) states that an effective performance management system can help organizations improve productivity, develop competencies, and support more objective decision-making processes.

As a national electricity provider, PT PLN (Persero) plays an important role in maintaining the quality of service and distribution of electricity to the public. One of the operational activities carried out is the Electricity Usage Control (P2TL), which is the inspection of customer electricity use to reduce electricity usage violations and minimize potential company losses due to non-compliant electricity use.

In the implementation of P2TL activities, each officer has different levels of productivity and work effectiveness. These differences can be seen from the number of inspections carried out, the number of violation findings, and officer performance assessment results. However, the officer performance evaluation process is still

503



conducted manually and has not optimally utilized data analysis. This condition causes management to have difficulty in objectively, measurably, and systematically identifying the performance segmentation of officers as a basis for managerial decision-making.

One data mining method that can be used to perform data segmentation is the K-Means Clustering algorithm. The K-Means algorithm is a clustering method widely used to group data based on the similarity of characteristics between objects. This method has been applied in various fields, including employee performance evaluation, because it is capable of producing objective and easily interpretable data segmentation. Research by Regina, Sutinah, and Agustina (2021) shows that K-Means is able to group employee performance quality based on their characteristics, thereby supporting the organizational performance evaluation process.

Various previous studies have applied the K-Means algorithm for customer segmentation, employee performance quality, and employee performance assessment. Regina et al. (2021) applied K-Means to group employee performance quality at a chemical company, while Pramitasari and Nataliani (2021) compared K-Means and Fuzzy C-Means algorithms in employee clustering based on performance values. The results of these studies show that clustering methods can help organizations identify data patterns that are not easily recognized through conventional analysis.

Other studies also demonstrate the effectiveness of the K-Means algorithm in the segmentation and performance evaluation process. Agustina and Prihandoko (2018) showed that the K-Means algorithm can effectively group employee discipline levels based on their characteristics. Setiawan (2019) also applied the K-Means algorithm to identify employee capabilities in the field of information technology and obtained groupings that can be used as the basis for organizational decision-making. In addition, Intan et al. (2023) showed that the application of K-Means in employee performance assessment can produce more objective segmentation and help organizations in the performance evaluation process.

Although various previous studies have shown that the K-Means algorithm is effective for data segmentation in various fields, most studies still focus on customer segmentation, consumer behavior, and general employee performance evaluation. Research specifically examining the performance segmentation of Electricity Usage Control (P2TL) officers as support for managerial decision-making at PT PLN (Persero) is still relatively limited. Meanwhile, the work characteristics of P2TL officers have a strategic role in supporting the effectiveness of electricity usage supervision and minimizing potential company losses. Therefore, research is needed that can identify



patterns of officer performance more objectively through a data mining approach in order to produce relevant information as a basis for performance evaluation and data-driven human resource management.

Based on these issues, this study was conducted to apply the K-Means algorithm for the performance segmentation of P2TL officers at PT PLN (Persero) UP3 Kebon Jeruk. The results of the study are expected to support the managerial decision-making process in performance evaluation, strategy development, coaching, and improving the data-driven operational effectiveness of P2TL officers.

PROBLEM FORMULATION

1. How is the K-Means algorithm applied to segment the performance of Electricity Usage Control (P2TL) officers at PT PLN (Persero) UP3 Kebon Jeruk?
2. What are the characteristics of each P2TL officer performance group based on the variables of number of inspections, P2TL findings, and performance assessments produced through the clustering process using the K-Means algorithm?
3. How can the results of P2TL officer performance segmentation be utilized as support for managerial decision-making in the performance evaluation and human resource development processes at PT PLN (Persero) UP3 Kebon Jeruk?

RESEARCH OBJECTIVES

This study aims to apply the K-Means algorithm to segment the performance of Electricity Usage Control (P2TL) officers at PT PLN (Persero) UP3 Kebon Jeruk. In addition, this study aims to identify the characteristics of each officer performance group based on the variables of number of inspections, P2TL findings, and performance assessments. The segmentation results obtained are then expected to be utilized as support for managerial decision-making in the performance evaluation, coaching, and human resource development processes at PT PLN (Persero) UP3 Kebon Jeruk.

RESEARCH METHOD

This study uses a quantitative approach with the data mining method using the K-Means Clustering algorithm. The research object is the Electricity Usage Control (P2TL) officers at PT PLN (Persero) UP3 Kebon Jeruk. The data used is operational data of officers consisting of the variables of number of inspections, number of P2TL findings, and officer performance assessments.



The research stages include data collection, data preprocessing, data normalization, clustering process, cluster evaluation, and interpretation of results. The preprocessing stage is carried out to ensure data quality through checking data completeness, data consistency, and cleaning of irrelevant data. Data normalization is then performed to reduce scale differences between variables so that the grouping process can be carried out more optimally.

The clustering process was carried out using the K-Means algorithm implemented through RapidMiner Studio. The K-Means algorithm is used to group data based on the similarity of characteristics between objects so that data groups with relatively homogeneous characteristics within one cluster and heterogeneous characteristics between clusters are obtained. The clustering process is carried out through the stages of determining the number of clusters, forming initial centroids, grouping data into the nearest cluster, and iteratively updating centroids until a stable cluster condition is obtained.

To determine the most optimal number of clusters, this study uses the Davies Bouldin Index (DBI) method. The DBI method is used to evaluate cluster quality based on the level of data homogeneity within a cluster and heterogeneity between clusters. The smaller the DBI value obtained, the better the quality of the clusters produced. The DBI evaluation results are then used as the basis for determining the most representative number of clusters to describe the performance characteristics of P2TL officers.

The clustering results obtained are then analyzed to identify the characteristics of each officer performance group based on the variables of number of inspections, P2TL findings, and performance assessments. This information is used as the basis for supporting the performance evaluation process, competency development, and managerial decision-making at PT PLN (Persero) UP3 Kebon Jeruk.

DISCUSSION

Research Flowchart

This research was conducted through several stages, starting from the collection of P2TL officer operational data, data preprocessing, data normalization, clustering process using the K-Means algorithm, cluster evaluation using the Davies Bouldin Index (DBI), and interpretation of clustering results.



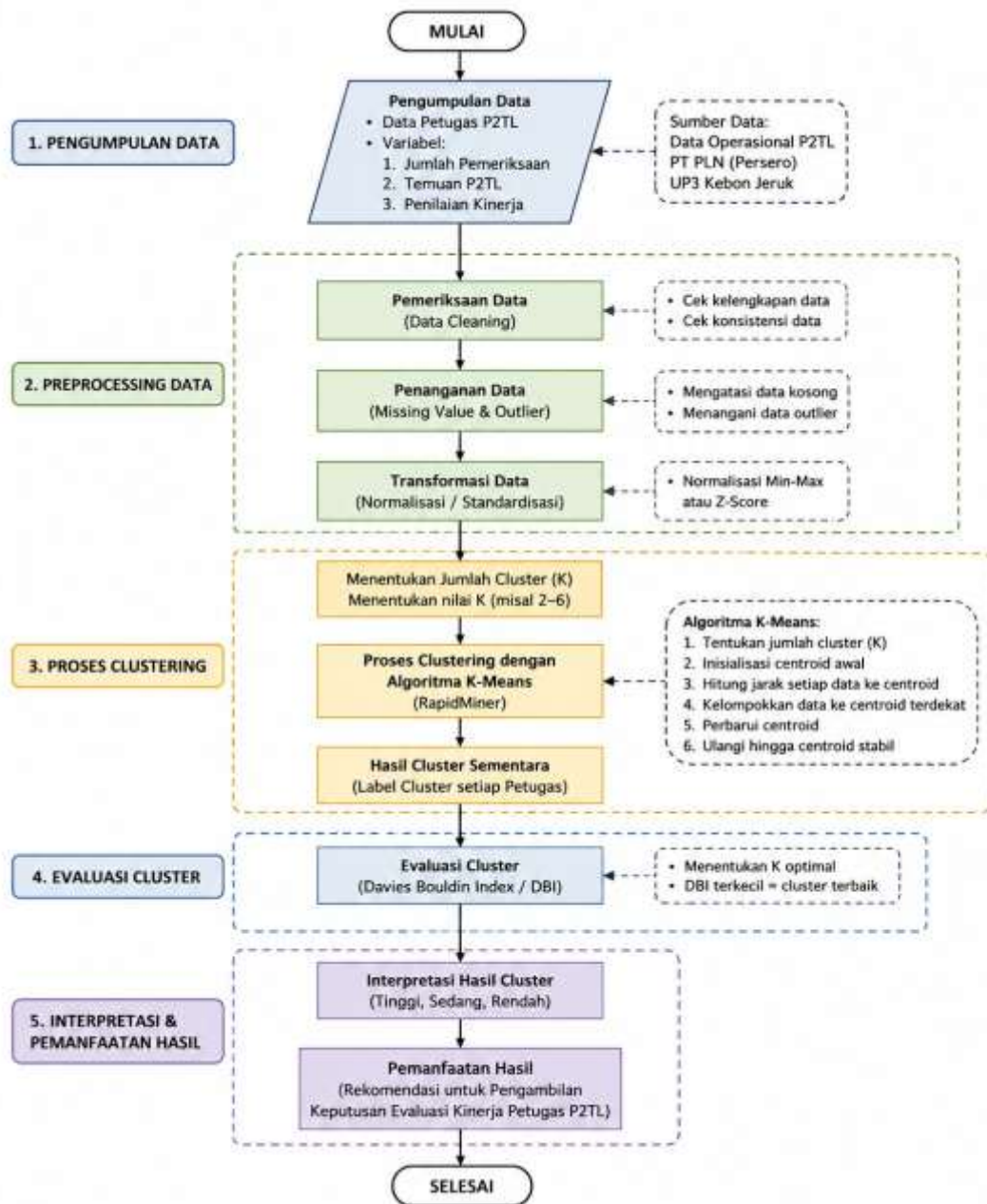


Figure 1.1 Research Flowchart

The research flowchart is used to systematically describe the research stages so that the data processing can be carried out in a structured manner and produce optimal data segmentation as support for managerial decision-making.

Research Data Description

This study uses operational data of Electricity Usage Control (P2TL) officers at PT PLN (Persero) UP3 Kebon Jeruk. The research data consists of the variables of number of inspections, number of P2TL findings, and officer performance assessments. These three variables are used as the basis for the officer performance segmentation process using the K-Means Clustering algorithm.

Before the clustering process, the data first goes through a preprocessing stage to ensure the data is in a clean condition and ready for use in the processing. The preprocessing stage includes checking for empty data, data type conformity, and data normalization to equalize the scale between variables. Normalization is performed so that differences in data scale do not affect the distance calculation process in the K-Means algorithm.



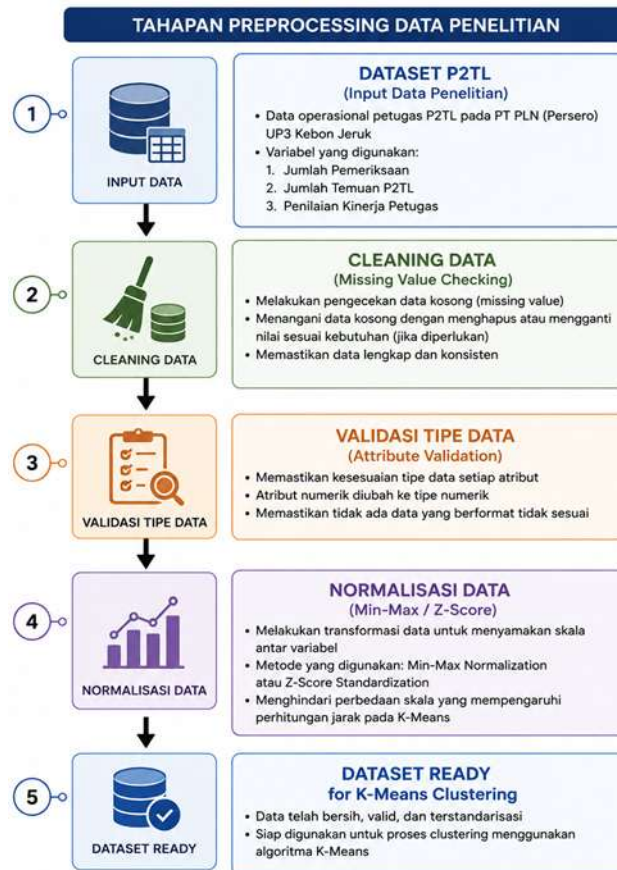


Figure 1.2 Data Preprocessing Stages

Descriptive Statistical Analysis

Descriptive statistical analysis was performed to provide a general overview of the research data based on the variables of number of inspections, number of P2TL findings, and officer performance assessments. Descriptive statistics include minimum, maximum, mean, and standard deviation values.

Table 1.1 Research Dataset

No.	Officer Name	Number of Inspections	P2TL Findings	Performance Assessment
1	RISAD NAMAWI	413	49	88.53
2	TAMIN DJUL FADHLI RAMBE	331	30	61.68
3	NIKSON MARBUN	414	52	91.63
...	...	...	...	...
28	JOAN LATABA	284	27	54.18
29	RASYID HATTAMI	264	22	47.27
30	HANS YEHECHIEL	251	18	42.02

Source: Data processed by author (2026)

Table 1.2 Descriptive Statistics of Research Variables

Variable	N	Minimum	Maximum	Mean	Std. Deviation
Number of Inspections	30	138	425	334	84.83
P2TL Findings	30	7	54	32	14.25
Performance Assessment	30	23.21	94.68	64	21.33

Source: Data processed by author using RapidMiner Studio (2026)

Based on the descriptive statistical results in Table 5.2, it is known that the number of inspections variable has a minimum value of 138 and a maximum value of 425 with a



mean of 334 and a standard deviation of 84.83. These results indicate differences in the level of operational activity between officers in the implementation of P2TL activities.

The P2TL findings variable has a minimum value of 7 and a maximum value of 54 with a mean of 32 and a standard deviation of 14.25. This condition shows that the effectiveness of inspections carried out by officers has different levels of variation.

Meanwhile, the performance assessment variable has a minimum value of 23.21 and a maximum value of 94.68 with a mean of 64 and a standard deviation of 21.33. These results indicate that officer performance achievements have quite diverse characteristics based on company operational indicators.

Overall, the descriptive statistical results indicate that the research data has a sufficient level of heterogeneity to conduct the segmentation process using the K-Means Clustering algorithm. The differences in characteristics between officers serve as the basis for cluster formation based on the level of similarity of operational performance characteristics of P2TL officers.

Determining the Number of Clusters

Determining the number of clusters was carried out through several cluster number experiments using the K-Means Clustering algorithm. Cluster quality evaluation was performed using the Davies Bouldin Index (DBI) method to determine the quality level of data grouping based on within-cluster homogeneity and between-cluster heterogeneity.

The Davies Bouldin Index value is used as the basis for determining the optimal number of clusters, where the smaller the DBI value, the better the quality of the clusters produced.

Table 1.3 Davies Bouldin Index Test Results

(K)	DBI Value
2	-0.447
3	-0.609
4	-0.549
5	-0.406

Source: Data processed by author using RapidMiner Studio (2026)



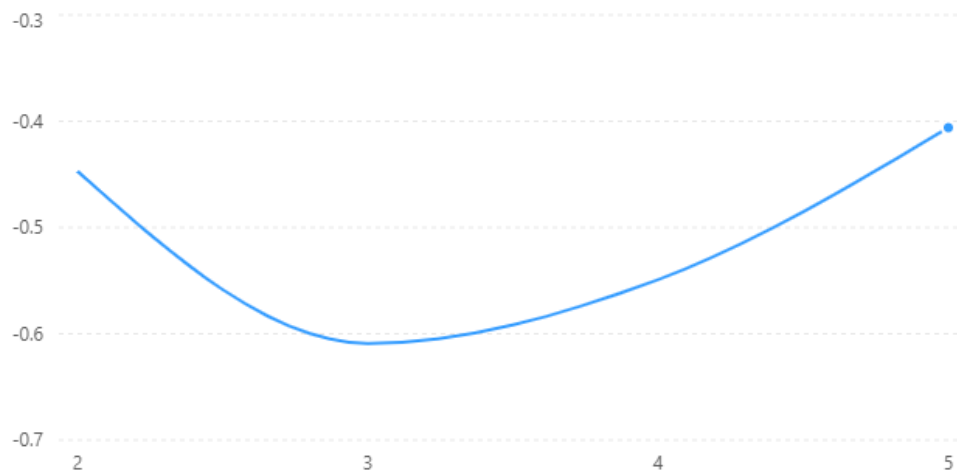


Figure 1.4 Davies Bouldin Index Graph

Based on the Davies Bouldin Index (DBI) test results for several cluster numbers, DBI values were obtained for K=2 at -0.447, K=3 at -0.609, K=4 at -0.549, and K=5 at -0.406. The test results show that the best number of clusters is at K=3 because it produces the most optimal DBI value compared to other cluster numbers.

The Davies Bouldin Index graph shows that the DBI value decreases from K=2 to K=3, then increases again at K=4 and K=5. This condition indicates that the use of three clusters can produce a good level of homogeneity in each cluster and a sufficiently clear level of heterogeneity between clusters.

Theoretically, the smaller the Davies Bouldin Index value, the better the clustering quality produced, as it indicates a high degree of data similarity within one cluster and clearer differences between clusters.

Clustering Results Using the K-Means Algorithm

The data grouping process in this study was carried out using the K-Means algorithm with the variables of number of inspections, number of P2TL findings, and officer performance assessments. Based on the evaluation results using the Davies Bouldin Index (DBI), the optimal number of clusters was three, with the best DBI value of -0.609. The clustering results show that the P2TL officer data was successfully grouped into three main categories: high-performance cluster, medium-performance cluster, and low-performance cluster.

Table 1.4 Cluster Grouping Results

Cluster	Number of Members	Cluster Category	Cluster Characteristics
Cluster 0	7	Medium Performance	Has medium levels of number of inspections, P2TL findings, and performance assessment values.
Cluster 1	11	High Performance	Has the highest number of inspections, P2TL findings, and performance assessment values compared to other clusters.
Cluster 2	12	Low Performance	Has relatively low number of inspections, P2TL findings, and performance assessment values compared to other clusters.

Source: Data processed by author using RapidMiner Studio (2026)

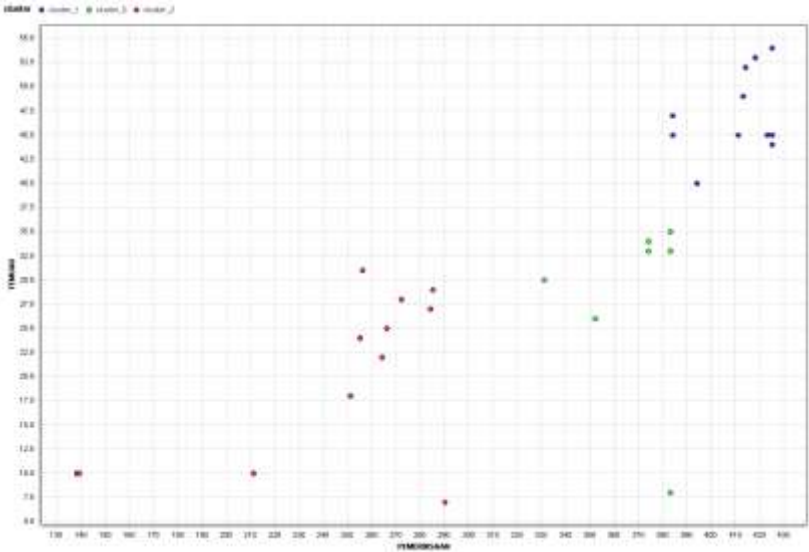


Figure 1.5 Scatter Plot of Clustering Results

The scatter plot visualization shows that the P2TL officer data was successfully segmented into several clusters based on the similarity of operational characteristics of the officers. The distribution of data in each cluster shows differences in performance

patterns based on the variables of number of inspections, P2TL findings, and performance assessments, so that the clustering process was able to produce a fairly clear separation of data between groups of officers.

Cluster Centroid Analysis

Centroid analysis was conducted to determine the characteristics of each cluster based on the average values of the research variables, which include the number of inspections, number of P2TL findings, and officer performance assessments.

Table 1.5 Centroid Value of Each Cluster

Variable	Cluster 0	Cluster 1	Cluster 2	Category
Number of Inspections	368.57	410.55	242.58	
P2TL Findings	28.43	47.18	20.08	
Performance Assessment	63.00	85.91	43.00	
Category	Medium	High	Low	

Source: Data processed by author using RapidMiner Studio (2026)

Based on Table 5.6, it is known that cluster 1 has the highest centroid values for all research variables: number of inspections at 410.55, P2TL findings at 47.18, and performance assessment at 85.91. These results indicate that cluster 1 represents the group of officers with high performance levels. Meanwhile, cluster 0 has centroid values at an intermediate level with an average number of inspections of 368.57, P2TL findings of 28.43, and performance assessment of 63.00, so it is categorized as a cluster with medium performance levels. Cluster 2 has the lowest centroid values for all research variables: number of inspections at 242.58, P2TL findings at 20.08, and performance assessment at 43.00, so it is categorized as a cluster with low performance levels.

Overall, the centroid results show that the K-Means Clustering algorithm was able to group P2TL officers based on the similarity of operational characteristics quite clearly. The grouping results can be used as the basis for performance evaluation and more objective, data-driven managerial decision-making.



Cluster Characteristic Analysis**High Performance Cluster**

The high-performance cluster is a group of officers who have the highest average values of number of inspections, number of P2TL findings, and performance assessments compared to other clusters.

Table 1.6 Characteristics of High Performance Cluster

Variable	Centroid Value (C1)
Number of Inspections	410.545
P2TL Findings	47.182
Performance Assessment	85.909

Source: Data processed by author using RapidMiner Studio (2026)

Based on Table 5.7, the high-performance cluster has a centroid value for number of inspections of 410.545, number of P2TL findings of 47.182, and performance assessment of 85.909. These values are the highest compared to other clusters, indicating that officers in this cluster have good levels of productivity and work effectiveness in implementing Electricity Usage Control (P2TL) activities.

The high number of inspections and P2TL findings indicates that officers in this cluster have better technical skills, job mastery, discipline, and effectiveness in carrying out their duties compared to other groups. In addition, the high performance assessment values indicate that officers in this cluster are able to achieve operational targets optimally and make significant contributions to P2TL activities.

The segmentation results show that the group of officers in the high-performance cluster has good performance consistency and has the potential to become role models in operational activities. From a managerial perspective, this group can serve as the basis for formulating reward policies, promotions, and advanced competency development to maintain and sustainably improve work productivity.

Medium Performance Cluster

The medium-performance cluster indicates officers with moderate levels of inspections, P2TL findings, and performance assessment values. This group shows a sufficiently good level of productivity, but still has the potential to be improved in order to reach the high-performance category.

515



Table 1.7 Characteristics of Medium Performance Cluster

Variable	Centroid Value (C0)
Number of Inspections	368.571
P2TL Findings	28.429
Performance Assessment	63.000

Source: Data processed by author using RapidMiner Studio (2026)

Based on Table 5.8, the medium-performance cluster has a centroid value for number of inspections of 368.571, number of P2TL findings of 28.429, and performance assessment of 63.000. These values are between the high-performance cluster and the low-performance cluster, indicating that officers in this cluster have been able to carry out operational tasks fairly well. However, there is still an opportunity to improve the effectiveness of inspections so as to produce more optimal findings and support the achievement of company operational targets.

The group of officers in the medium-performance cluster is a group with great potential to be elevated to the high-performance cluster. Therefore, management needs to conduct coaching, improve technical competencies, and strengthen operational supervision on an ongoing basis. This strategy is expected to improve work productivity and encourage more optimal performance achievement in Electricity Usage Control (P2TL) activities.

Low Performance Cluster

The low-performance cluster consists of officers with relatively low numbers of inspections, P2TL findings, and performance assessment values compared to other clusters. This condition indicates that the effectiveness of operational activity implementation is still not optimal.

Table 1.8 Characteristics of Low Performance Cluster

Variable	Centroid Value (C2)
Number of Inspections	242.583
P2TL Findings	20.083



Performance Assessment	43.000
------------------------	--------

Source: Data processed by author using RapidMiner Studio (2026)

Based on Table 5.9, the low-performance cluster has a centroid value for number of inspections of 242.583, number of P2TL findings of 20.083, and performance assessment of 43.000. These values are the lowest compared to other clusters, indicating that officers in this group have levels of productivity and work effectiveness that still need to be improved. This condition suggests that there are factors that potentially affect officer performance achievement. These factors can become the focus of future research to obtain a more comprehensive understanding of the determinants of P2TL officer performance.

These segmentation results provide important information for management in determining priorities for coaching and human resource development, including training programs, work mentoring, technical competency improvement, and strengthening of performance monitoring and evaluation mechanisms.

Discussion of Research Results

Based on the data processing results using the K-Means algorithm, three main clusters were obtained representing the performance levels of P2TL officers: high, medium, and low performance clusters. The number of clusters was determined through Davies Bouldin Index (DBI) evaluation, where the best DBI value of -0.609 was obtained at K=3. This finding shows that the use of three clusters can produce optimal data grouping based on the similarity of officer performance characteristics.

The clustering results show fairly clear characteristic differences between clusters. The high-performance cluster has higher average values of number of inspections, P2TL findings, and performance assessments compared to other clusters. Conversely, the low-performance cluster shows the lowest average values for all research variables. This condition indicates that the levels of productivity and work effectiveness of P2TL officers have diverse characteristics, thus requiring different human resource management approaches for each group.

These research findings reinforce the concept of performance management that emphasizes the importance of performance measurement and evaluation as the basis for human resource management. Through the clustering approach, performance evaluation can be carried out more objectively because it is based on actual data patterns that reflect



the operational achievements of officers. This approach enables organizations to obtain more accurate information in the planning, supervision, and human resource development processes (Aguinis, 2019; Armstrong, 2021).

The results of this study are also consistent with the research by Regina, Sutinah, and Agustina (2021), which shows that the K-Means algorithm can group employee performance quality based on their characteristics, thereby facilitating the organizational performance evaluation process. In this study, the K-Means algorithm successfully grouped P2TL officers into three performance groups: high, medium, and low performance clusters with different characteristics based on number of inspections, P2TL findings, and performance assessments.

These research findings also support the results of Pramitasari and Nataliani (2021), which states that clustering methods can be used to identify groups of employees based on their performance levels. This similarity in results indicates that the K-Means algorithm has good capability in grouping data based on the similarity of characteristics between objects, thus producing more objective information compared to conventionally conducted performance evaluations.

These segmentation results provide more objective information about officer performance conditions based on operational data. This information enables management to identify groups of officers with optimal performance, groups that still have potential to be improved, and groups that require more attention and coaching. Thus, the performance evaluation process is not only based on individual assessments but also takes into account performance patterns formed from data analysis results.

The research findings show that the productivity and effectiveness of P2TL officers are not evenly distributed across all officers. This condition indicates the need for different human resource management strategies for each performance group. Officers in the high-performance cluster require strategies to maintain performance through awards and advanced competency development, while officers in the medium and low-performance clusters require more intensive coaching programs, training, and work mentoring. This approach enables the organization to allocate development resources more precisely, thereby improving the overall effectiveness of performance management.

From a managerial perspective, the results of this study can be used as the basis for formulating human resource development strategies, determining training needs, providing rewards, and improving operational supervision effectiveness. Officers in the high-performance group can serve as role models in P2TL activities, while the medium



and low-performance groups can be prioritized in coaching and competency improvement programs.

With officer performance segmentation, management can identify groups of officers who need coaching, groups with potential for development, and groups worthy of being performance benchmarks. This information serves as an important basis for formulating human resource management policies and sustainably improving operational performance.

CONCLUSION

This study successfully applied the K-Means algorithm to segment the performance of Electricity Usage Control (P2TL) officers at PT PLN (Persero) UP3 Kebon Jeruk. Evaluation results using the Davies Bouldin Index (DBI) show that the optimal number of clusters was obtained at K=3 with a DBI value of -0.609.

The clustering results produced three officer performance groups: high-performance cluster, medium-performance cluster, and low-performance cluster. The characteristic differences in each cluster indicate that officer performance has different levels of productivity and effectiveness based on number of inspections, P2TL findings, and performance assessments.

The segmentation results obtained can be used as the basis for supporting the performance evaluation process, coaching, competency development, and more objective, data-driven managerial decision-making at PT PLN (Persero) UP3 Kebon Jeruk.

RESEARCH LIMITATIONS

This study still has several limitations that need to be considered in interpreting the research results. First, the variables used in the clustering process are still limited to number of inspections, number of P2TL findings, and officer performance assessments. Other variables that potentially affect performance, such as work experience, operational area, education level, workload, or assignment period, have not been included in the research model. Therefore, future research is expected to add more diverse variables to produce more comprehensive segmentation.

Second, the data used in this study is still limited to officer operational data within a certain period, so the segmentation results obtained do not fully represent the overall officer performance conditions in the long term. Future research can use data with a longer observation period to obtain more representative segmentation results.

519



Third, this study only uses the K-Means algorithm as the clustering method. Although this algorithm was able to produce fairly good data grouping, future research can make comparisons with other clustering methods, such as Hierarchical Clustering, DBSCAN, or Fuzzy C-Means to obtain more optimal segmentation results and provide a broader analytical perspective.

REFERENCES

- Aguinis, H. (2019). *Performance Management* (4th ed.). Chicago Business Press.
- Agustina, N., & Prihandoko, P. (2018). Comparison of K-Means and Fuzzy C-Means algorithms for clustering employee discipline performance levels. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 2(3), 621–626.
<https://doi.org/10.29207/resti.v2i3.492>
- Armstrong, M. (2021). *Armstrong's Handbook of Performance Management* (7th ed.). Kogan Page.
- Berry, M. J. A., & Linoff, G. S. (2011). *Data Mining Techniques for Marketing, Sales, and Customer Relationship Management* (3rd ed.). Wiley Publishing.
- Dessler, G. (2020). *Human Resource Management* (16th ed.). Pearson.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
- Intan, M., Priyatna, B., Nurapriani, F., & Hilabi, S. S. (2023). Implementation of K-Means clustering method on employee performance assessment at PT Kopetri Citra Abadi. *Jurnal Informasi Interaktif*, 8(1).
- Kurniawan, H., Defit, S., & Sumijan. (2020). Data mining using K-Means clustering method to determine tuition fees. *Journal of Applied Computer Science and Technology (JACOST)*, 1(2).
- Kusrini, & Luthfi, E. T. (2019). *Data Mining Algorithms*. Andi.
- Larose, D. T. (2014). *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley & Sons.
- Mathis, R. L., Jackson, J. H., & Valentine, S. R. (2021). *Human Resource Management* (16th ed.). Cengage Learning.
- Moeheriono. (2018). *Competency-Based Performance Measurement*. RajaGrafindo Persada.
- Pramitasari, A. E., & Nataliani, Y. (2021). Comparison of employee clustering based on performance values with K-Means and Fuzzy C-Means algorithms. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 8(3), 1119–1132.



- Prasetyo, E. (2018). Data Mining: Concepts and Applications Using MATLAB. Andi Offset.
- RapidMiner. (2024). RapidMiner Studio Documentation. RapidMiner Inc.
- Regina, S., Sutinah, E., & Agustina, N. (2021). Clustering of employee performance quality at a chemical company using the K-Means algorithm. Media Informatika Budidarma, 5(2), 573–582.
- Santosa, B. (2017). Data Mining: Data Utilization Techniques for Business Purposes. Graha Ilmu.
- Setiawan, S. (2019). Cluster analysis using K-Means algorithm to identify employee capabilities in IT at CV Roxed Ltd. Jurnal Pelita Informatika, 18(3), 80–86.
- Sugiyono. (2022). Quantitative, Qualitative, and R&D Research Methods. Alfabeta.
- Tan, P. N., Steinbach, M., & Kumar, V. (2016). Introduction to Data Mining. Pearson.
- Witten, I. H., Frank, E., & Hall, M. A. (2017). Data Mining: Practical Machine Learning Tools and Techniques (4th ed.). Morgan Kaufmann.

