International Journal Software Engineering and Computer Science (IJSECS)
6 (2), 2026, 483-495
Published Online August 2026 in IJSECS (http://www.journal.lembagakita.com/index.php/ijsecs)
P-ISSN: 2776-4869, E-ISSN: 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v6i2.7579.

RESEARCH ARTICLE **Open Access**

User Experience Evaluation of the Mobile JKN Application Using the System Usability Scale (SUS) and K-Means Clustering

Khaila Mukti Harahap [1*], **Abdul Halim Hasugian** [2]

[1*,2] Universitas Islam Negeri Sumatera Utara, Deli Serdang Regency, North Sumatra Province, Indonesia.

**Corresponding author*: muktiharahap321@gmail.com.

Received: June 16, 2026; Accepted: July 3, 2026; Published: August 1, 2026.

Abstract: The Mobile JKN application developed by BPJS Kesehatan provides digital access to healthcare administration services in Indonesia. However, users continue to report difficulties when interacting with the application. This study evaluated the usability of Mobile JKN and examined user experience patterns by combining the System Usability Scale (SUS) with K-means clustering. Data were collected through an online questionnaire completed by 550 respondents in Medan who reported experience using the application. The mean SUS score was 50.13, which was classified as Poor and Grade D under the interpretation framework adopted in this study. Among the four UX dimensions, Efficiency received the lowest score at 53.15 out of 100, followed by Ease of Use at 55.14, System Reliability at 56.38, and User Satisfaction at 57.42. The three-cluster solution produced different score profiles: Cluster 0 (Fairly Satisfied) had a mean SUS score of 60.09 and comprised 168 respondents; Cluster 1 (Somewhat Dissatisfied) had a mean score of 47.61 and comprised 201 respondents; and Cluster 2 (Dissatisfied) had a mean score of 43.69 and comprised 181 respondents. Efficiency and ease of use were therefore the dimensions receiving the least favorable evaluations. Combining SUS with K-means revealed variations across the overall usability score and the four UX dimensions that were not represented by the aggregate SUS score alone. However, the low Silhouette Score indicated weak separation among the clusters; therefore, the resulting profiles should be interpreted as exploratory patterns that can guide further usability evaluation rather than as definitive user segments.

Keywords: User Experience; Mobile JKN; System Usability Scale; K-Means Clustering; Usability.

1. Introduction

The development of information technology has changed the way public services are delivered, including health administration services. In Indonesia, BPJS Kesehatan provides Mobile JKN as a digital service that allows National Health Insurance participants to access various administrative and informational services through a mobile application. The application is intended to simplify access to BPJS Kesehatan services and reduce the need for users to visit service offices directly (Akbar *et al.*, 2025). Mobile JKN provides several services, including checking membership information, updating personal data, registering at healthcare facilities, and submitting complaints. The application was launched by BPJS Kesehatan on November 15, 2017, as part of its efforts to provide easier access to information and services for participants (Ulum *et al.*, 2024).

The large number of Mobile JKN users indicates that the application has been widely adopted. However, the number of users alone does not necessarily indicate that an application provides a satisfactory user experience. Users may continue to encounter difficulties even when an application is widely used. Initial interviews with BPJS Kesehatan staff at the Medan Branch indicated that some participants still visited the office to complete services despite having used Mobile JKN. Reported difficulties included problems accessing services through the application, navigation issues, and slow system responses. These conditions indicate that the availability of digital services does not necessarily ensure that users can complete their tasks easily and efficiently. Therefore, an evaluation of the usability and user experience of Mobile JKN is needed to identify aspects that may affect users' interaction with the application.

One commonly used instrument for usability evaluation is the System Usability Scale (SUS), developed by John Brooke in 1986. SUS consists of 10 statements with a five-point Likert response scale and produces an overall usability score ranging from 0 to 100. The instrument has been widely applied in usability evaluation because it provides a relatively simple way to obtain an overall assessment of perceived system usability (I Komang Adyanata *et al.*, 2024; Putri & Liu, 2024). Previous research has also applied SUS to evaluate Mobile JKN and other information systems. For example, Aziza and Ristriani (2023) evaluated Mobile JKN using usability and heuristic evaluation methods and reported a usability score of 62.25%. However, an overall usability score may not fully describe differences in perceptions among individual users. Respondents with different experiences may produce substantially different evaluations even when their responses are summarized by a single average score.

To obtain a more detailed view of these differences, clustering techniques can be applied to group respondents based on similarities in their evaluations. K-Means is a partition-based clustering algorithm that assigns observations to clusters according to their distance from cluster centroids and seeks to minimize the Within-Cluster Sum of Squares (WCSS). In usability research, this approach can be used to identify groups of respondents with similar evaluation patterns. Nurzanah *et al.* (2026), for example, applied X-Means clustering with the SUS instrument to evaluate the CBT Candy application and identified two user groups with a Silhouette Score of 0.781. The study demonstrates the potential of combining usability measurement with clustering to obtain user segmentation beyond an overall usability score. However, the object and clustering method used in that study differ from those used in the present research.

Previous studies on Mobile JKN have generally focused on measuring usability, user experience, or user satisfaction using a single evaluation approach. Aziza and Ristriani (2023), for example, reported usability findings from usability and heuristic evaluation, while other studies have examined Mobile JKN using instruments such as UEQ and EUCS (Irwan *et al.*, 2023; Ulum *et al.*, 2024; Zulpa Salsabila *et al.*, 2024). These studies provide useful information regarding users' perceptions of Mobile JKN, but further analysis is needed to examine whether users can be grouped according to differences in their usability evaluations. This gap provides an opportunity to apply clustering to Mobile JKN usability data.

Based on this background, this study evaluates the usability of the Mobile JKN application using the System Usability Scale and applies K-Means Clustering to identify groups of users with similar evaluation patterns. The analysis considers four aspects used in the study: Ease of Use, Efficiency, System Reliability, and User Satisfaction. Data were obtained from 550 respondents in Medan who had experience using the Mobile JKN application. The clustering results are used to examine differences among user groups and identify aspects that may require further attention. The findings are expected to provide more specific information for evaluating Mobile JKN and support recommendations related to usability, service efficiency, system reliability, and user satisfaction. The remainder of this paper is organized as follows. Section 2 presents the related literature and theoretical foundations. Section 3 describes the research methodology. Section 4 presents the results and clustering analysis. Section 5 discusses the findings and limitations of the study, while Section 6 presents the conclusions and recommendations for future research.

2. Related Work

2.1 User Experience and Usability

User Experience (UX) refers to the perceptions, responses, and experiences that users develop when interacting with a digital product or system. UX is not limited to system functionality but also involves aspects such as ease of use, efficiency, comfort, and user satisfaction. In public service applications such as Mobile JKN, usability is an important aspect because users need to access information and services through the application efficiently and with minimal difficulty (Aziza & Ristriani, 2023). Usability refers to the extent to which a system can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction within a specified context of use (Ernawati & Rahayu, 2023). In this study, usability evaluation considers four aspects: Ease of Use, Efficiency, System Reliability, and User Satisfaction. Ease of Use refers to the extent to which users can understand and operate the application's features. Efficiency refers to the extent to which the application supports users in completing their tasks within a reasonable amount of time and effort. System Reliability refers to the perceived stability of the application and the occurrence of technical disruptions during use. User Satisfaction refers to users' overall evaluation of their experience with the application. These four aspects are used as evaluation variables in this study and should be distinguished from the three usability components defined in ISO 9241-11.

2.2 System Usability Scale (SUS)

The System Usability Scale (SUS) is a usability evaluation questionnaire developed by John Brooke in 1986. The instrument consists of 10 statements with response options ranging from 1 to 5 on a Likert scale. Odd-numbered statements are positively worded, whereas even-numbered statements are negatively worded (I Komang Adyanata *et al.*, 2024). The SUS provides a standardized overall usability score that can be used to summarize users' perceptions of a system's usability. The SUS score is calculated by converting the responses to the 10 items. For odd-numbered items, the response score is reduced by 1, whereas for even-numbered items, the response score is subtracted from 5 (Saputra *et al.*, 2022). The resulting item contributions are summed and multiplied by 2.5 to obtain a final score ranging from 0 to 100. The interpretation of SUS scores should follow one classification system consistently throughout the study. In this study, SUS is used to measure the perceived usability of the Mobile JKN application based on user responses (Irwan *et al.*, 2023). The resulting usability data are subsequently analyzed using K-Means Clustering to identify groups of respondents with similar evaluation patterns.

2.3 K-Means Clustering

K-means is a partition-based clustering algorithm that divides a dataset into $k$clusters by assigning each data point to the nearest cluster centroid. The algorithm iteratively updates cluster assignments and centroid positions to minimize the within-cluster sum of squares (WCSS), defined as the sum of the squared distances between each data point and its corresponding centroid (Prayitno *et al.*, 2025). The number of clusters can be selected using the Elbow Method by identifying the point at which further increases in k produce progressively smaller reductions in WCSS (Maori & Evanita, 2023). Table 1 compares relevant studies according to their research object, evaluation method, main finding, and methodological limitation.

Table 1. Comparison of Relevant Studies

Study	Research object	Method	Main finding	Research limitation
Aziza and Ristriani (2023)	Mobile JKN	Usability evaluation and heuristic evaluation	Usability score of 62.25%	Did not segment users according to their usability perceptions
Ulum *et al.* (2024)	Mobile JKN	UEQ	Reported the user experience evaluation of Mobile JKN	Did not include clustering analysis
Zulpa Salsabila *et al.* (2024)	Mobile JKN	UEQ+	Reported positive evaluations across the measured UX scales	Did not include user segmentation
Irwan *et al.* (2023)	Mobile JKN	EUCS	Reported the level of user satisfaction with Mobile JKN	Did not analyze usability-based user segments
Nurzanah *et al.* (2026)	Candy CBT	SUS and X-means	Identified user segments based on usability evaluations	Examined a different application domain
Current study	Mobile JKN	SUS and K-means	Evaluates usability and identifies user segments based on SUS and UX dimensions	—

The studies summarized in Table 1 evaluated Mobile JKN using usability, user experience, or user satisfaction instruments, but they did not segment users according to differences in their usability perceptions. Nurzanah *et al.* (2026) combined SUS with clustering, although the study examined a computer-based testing application rather than a public health service application. The current study therefore combines SUS and K-means to evaluate the usability of Mobile JKN and identify user groups based on the overall SUS score and four UX dimensions: ease of use, efficiency, system reliability, and user satisfaction.

2.4 Previous Studies

Previous studies related to this research have examined Mobile JKN usability, user experience, and satisfaction using different evaluation methods. Aziza and Ristriani (2023) evaluated Mobile JKN using usability and heuristic evaluation methods and reported a usability score of 62.25%. Ulum *et al.* (2024) evaluated a government application using SUS and reported a moderate level of usability. Zulpa Salsabila *et al.* (2024) evaluated Mobile JKN using UEQ+, while Irwan *et al.* (2023) examined Mobile JKN using the EUCS approach. These studies provide information about users' perceptions of the application, but their analyses do not focus on grouping users according to similarities in their evaluation scores. Nurzanah *et al.* (2026) combined SUS with X-Means Clustering to evaluate the CBT Candy application and identified user groups based on SUS-related data. The study demonstrates the potential of combining usability measurement with clustering analysis. However, its research object differs from Mobile JKN, and it uses X-Means rather than K-Means. Based on the reviewed studies, previous research has provided measurements of usability, user experience, or satisfaction for Mobile JKN, but limited attention has been given to identifying groups of users with similar usability evaluation patterns. Therefore, this study combines the SUS method with K-Means Clustering to evaluate the usability of Mobile JKN and identify user groups based on their evaluation characteristics. The approach is intended to provide a more specific description of user evaluation patterns than an overall usability score alone.

3. Methodology

3.1 Research Design and Data Collection

This study employed a descriptive quantitative design to evaluate users' perceptions of the usability and user experience of the Mobile JKN application. Data were collected in February 2026 through an online questionnaire distributed to individuals in Medan who had experience using the application. Participation was voluntary. After data screening, all 550 responses were considered complete and valid and were included in the analysis. The research procedure comprised problem identification, questionnaire development, data collection, data preprocessing, SUS scoring, UX dimension scoring, feature construction, data normalization, K-means clustering, and cluster evaluation. The sequence of these stages is shown in Figure 1.

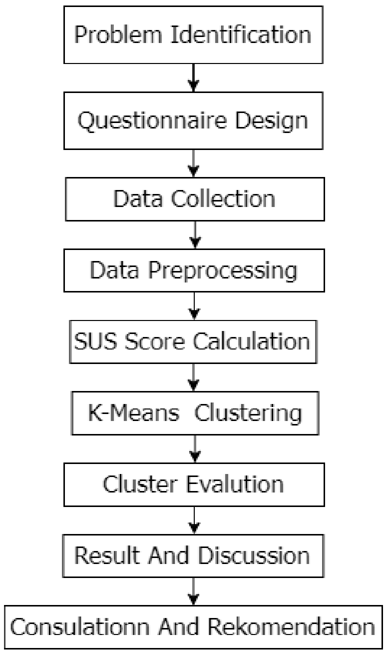

Figure 1. Research Stages

The research problem was identified based on reported difficulties in using the Mobile JKN application, including navigation problems, slow service processes, and low levels of user satisfaction. These issues provided the basis for developing the research instrument and evaluating the application's usability from the users' perspective.

3.2 Research Instrument

The research instrument comprised 26 questionnaire items divided into two components. The UX questionnaire contained 16 items grouped into four dimensions: Ease of Use, Efficiency, System Reliability, and User Satisfaction. Each dimension was represented by four items, and its score was calculated as the mean of the corresponding item responses. The second component comprised the 10 standard System Usability Scale items, which were used to calculate an overall usability score ranging from 0 to 100. The structure of the instrument is presented in Table 2.

Table 2. Structure of the Research Instrument

Instrument	Items	Construct or Dimension	Number of Items	Output
UX questionnaire	D1–D4	Ease of Use	4	Mean dimension score
UX questionnaire	D5–D8	Efficiency	4	Mean dimension score
UX questionnaire	D9–D12	System Reliability	4	Mean dimension score
UX questionnaire	D13–D16	User Satisfaction	4	Mean dimension score
SUS	Q1–Q10	Overall Usability	10	SUS score ranging from 0 to 100

3.3 Data Preprocessing and Feature Construction

All 550 questionnaire responses were screened for missing values and compliance with the response requirements. No incomplete or invalid responses were identified; therefore, no records were excluded from the analysis. The SUS score for each respondent was calculated using the standard scoring procedure described by Brooke (1986). For the positively worded odd-numbered items—Q1, Q3, Q5, Q7, and Q9—the contribution was calculated by subtracting 1 from the response value. For the negatively worded even-numbered items—Q2, Q4, Q6, Q8, and Q10—the contribution was calculated by subtracting the response value from 5. The 10 item contributions were then summed and multiplied by 2.5, producing a score ranging from 0 to 100. The 16 UX items were analyzed separately from the SUS items. Ease of Use was calculated as the mean of D1–D4, Efficiency as the mean of D5–D8, System Reliability as the mean of D9–D12, and User Satisfaction as the mean of D13–D16. Each respondent therefore received four UX dimension scores on a scale ranging from 1 to 5. The questionnaire responses were subsequently transformed into five analytical features for clustering: SUS Score, Ease of Use, Efficiency, System Reliability, and User Satisfaction. The 26 individual questionnaire items were not used directly as clustering variables. Instead, the items were aggregated according to their respective constructs to produce interpretable indicators and reduce the dimensionality of the clustering input. Because the five features had different measurement ranges, Min–Max normalization was applied before clustering. The SUS score ranged from 0 to 100, whereas the four UX dimensions ranged from 1 to 5. Normalization transformed each feature to a common range between 0 and 1, preventing the SUS score from disproportionately influencing the distance calculations because of its larger numerical scale. The normalization was calculated as follows:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

Where X represents the original feature value, $X_{\min}$ represents the minimum value of the feature, $X_{\max}$ represents its maximum value, and X' represents the normalized value.

3.4 K-Means Clustering and Cluster Evaluation

K-means clustering was applied to the five normalized features to identify groups of respondents with similar usability and user experience perceptions. The procedure began by specifying the number of clusters and initializing their centroids. Each respondent was assigned to the cluster with the nearest centroid based on the distance across the five normalized features. The centroids were then recalculated from the observations assigned to each cluster. Assignment and centroid updating continued iteratively until the cluster memberships converged. The clustering workflow is shown in Figure 2.

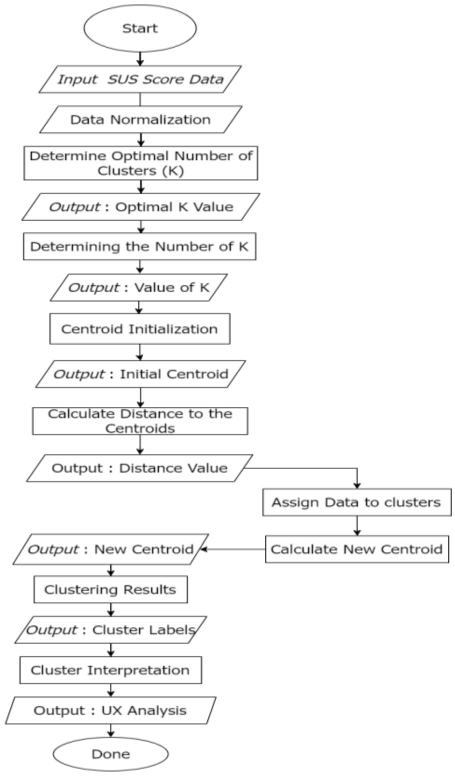

Figure 2. K-Means Clustering Workflow

The number of clusters was evaluated using the Elbow Method and the Silhouette Score. The Elbow Method was used to examine reductions in the within-cluster sum of squares as the number of clusters increased. The Silhouette Score was used to assess within-cluster cohesion and between-cluster separation. The final number of clusters was selected by considering these evaluation results together with the interpretability of the resulting user profiles.

4. Result and Discussion

4.1 Results

4.1.1 SUS Results

SUS scores were calculated for all 550 respondents. Table 3 presents the mean response and mean adjusted contribution for each SUS item.

Table 3. Mean Responses and Adjusted Contributions for the SUS Items

No.	SUS Item	Mean Response	Mean Adjusted Contribution	Item Polarity
1	I think that I would like to use the Mobile JKN application frequently.	3.000	2.000	Positive
2	I found the Mobile JKN application unnecessarily complex.	2.833	2.167	Negative
3	I thought the Mobile JKN application was easy to use.	2.835	1.835	Positive
4	I think that I would need the support of another person to use the Mobile JKN application.	3.389	1.611	Negative
5	I found that the various functions in the Mobile JKN application were well integrated.	3.140	2.140	Positive
6	I thought there was too much inconsistency in the Mobile JKN application.	3.465	1.535	Negative
7	I imagine that most people would learn to use the Mobile JKN application very quickly.	3.905	2.905	Positive
8	I found the Mobile JKN application very cumbersome to use.	3.407	1.593	Negative

9	I felt very confident using the Mobile JKN application.	3.362	2.362	Positive
10	I needed to learn many things before I could use the Mobile JKN application.	3.095	1.905	Negative
Total mean adjusted contribution			20.053	
Mean SUS score			50.13	

The item-level results provide additional context for the aggregate SUS score. The relatively high response to Q7 suggests that respondents perceived the application as potentially learnable. However, the lower response to Q3 indicates that their direct experience of ease of use was less favorable. The relatively high responses to the negatively worded Q6 and Q8 also indicate perceived inconsistency and difficulty in using the application. The descriptive statistics of the SUS scores are presented in Table 4.

Table 4. *Descriptive Statistics of the SUS Scores*

Statistic	Value
Number of respondents	550
Mean	50.132
Median	50.0
Standard deviation	9.265
Minimum	22.5
Maximum	72.5

The respondents obtained a mean SUS score of 50.13 with a standard deviation of 9.27. The scores ranged from 22.5 to 72.5, with a median of 50.0. Table 5 presents the distribution of respondents across the adopted SUS categories, while Figure 3 shows the score distribution and category frequencies.

Table 5. *Distribution of SUS Score Categories*

Category	Score Range	Count	Percentage	Grade
Excellent	≥85	0	0.0%	A
Good	71–84	5	0.9%	B
OK	51–70	231	42.0%	C
Poor	25–50	312	56.7%	D
Awful	<25	2	0.4%	F
Total	—	550	100.0%	—

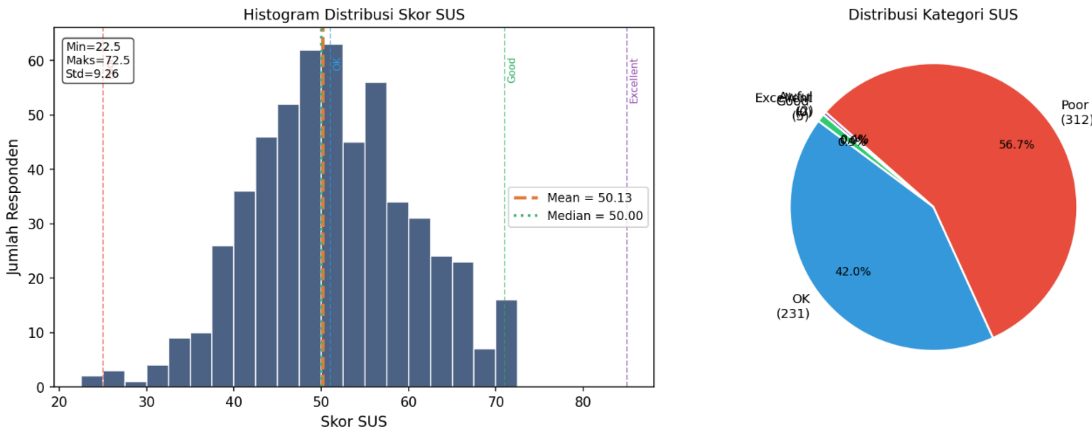

Figure 3. Distribution of SUS Scores and Categories

Of the 550 respondents, 312 respondents, or 56.7%, were classified in the Poor category. A further 231 respondents, or 42.0%, were classified in the OK category. Five respondents, or 0.9%, were classified in the Good category, and two respondents, or 0.4%, were classified in the Awful category. No respondents were classified in the Excellent category.

4.1.2 UX Dimension Results

The 16 UX items were aggregated into four dimensions: Ease of Use, Efficiency, System Reliability, and User Satisfaction. Table 6 presents the mean, standard deviation, observed range, and converted score for each dimension. The converted scores were calculated on a scale ranging from 0 to 100.

Table 6. Descriptive Statistics of the UX Dimensions

UX Dimension	Items	Mean, 1–5 Scale	Standard Deviation	Minimum	Maximum	Converted Score, 0–100
Ease of Use	D1–D4	3.205	0.590	1.25	4.75	55.14
Efficiency	D5–D8	3.126	0.575	1.50	4.75	53.15
System Reliability	D9–D12	3.255	0.586	1.50	5.00	56.38
User Satisfaction	D13–D16	3.297	0.618	1.50	5.00	57.42
Mean across dimensions	—	3.221	—	—	—	55.52

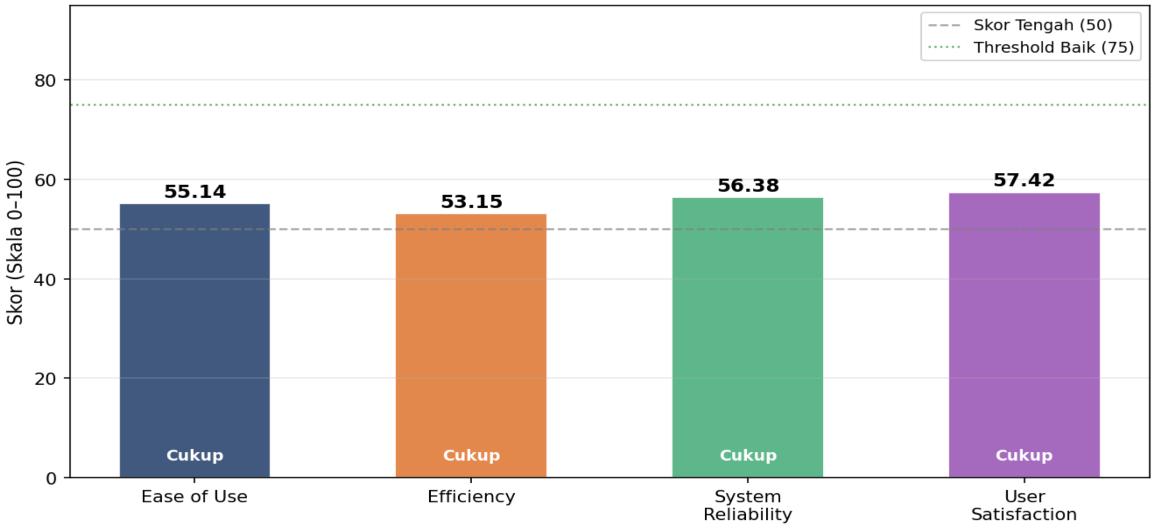

Figure 4. Mean Scores of the Four UX Dimensions

As shown in Table 6 and Figure 4, User Satisfaction had the highest converted score at 57.42, followed by System Reliability at 56.38 and Ease of Use at 55.14. Efficiency had the lowest score at 53.15. The mean score across the four dimensions was 55.52.

4.1.3 Clustering Results

The clustering analysis used five normalized features: SUS Score, Ease of Use, Efficiency, System Reliability, and User Satisfaction. The correlation coefficients among the five features ranged from −0.104 to 0.039, indicating that no strong linear relationships were observed among the features. The correlation matrix is presented in Figure 5.

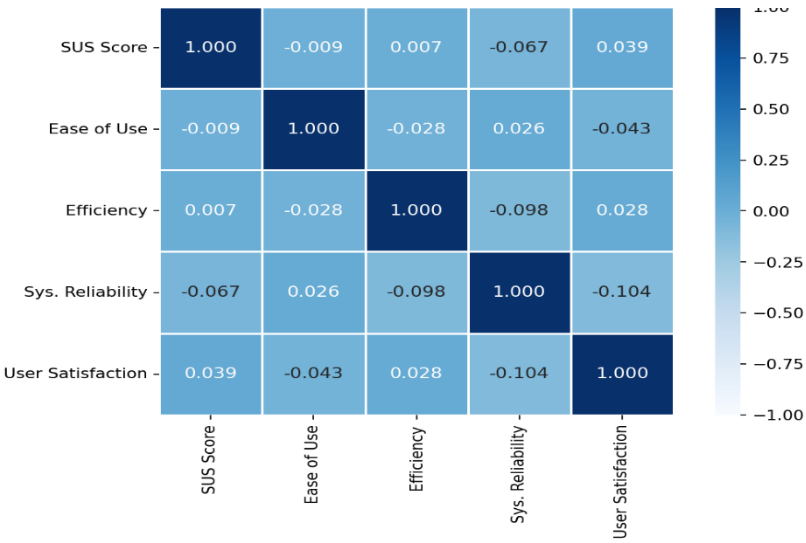

Figure 5. Correlation Heatmap of the Clustering Features

The number of clusters was assessed using the Elbow Method and the Silhouette Score. The evaluation results for values of K ranging from 2 to 10 are presented in Table 7 and Figure 6.

Table 7. Cluster Evaluation Results

K	WCSS	Silhouette Score
2	70.5354	0.1603
3	62.5475	0.1454
4	56.3172	0.1450
5	51.7889	0.1494
6	48.1696	0.1458
7	45.7459	0.1458
8	43.4586	0.1478
9	41.0093	0.1495
10	39.5902	0.1420

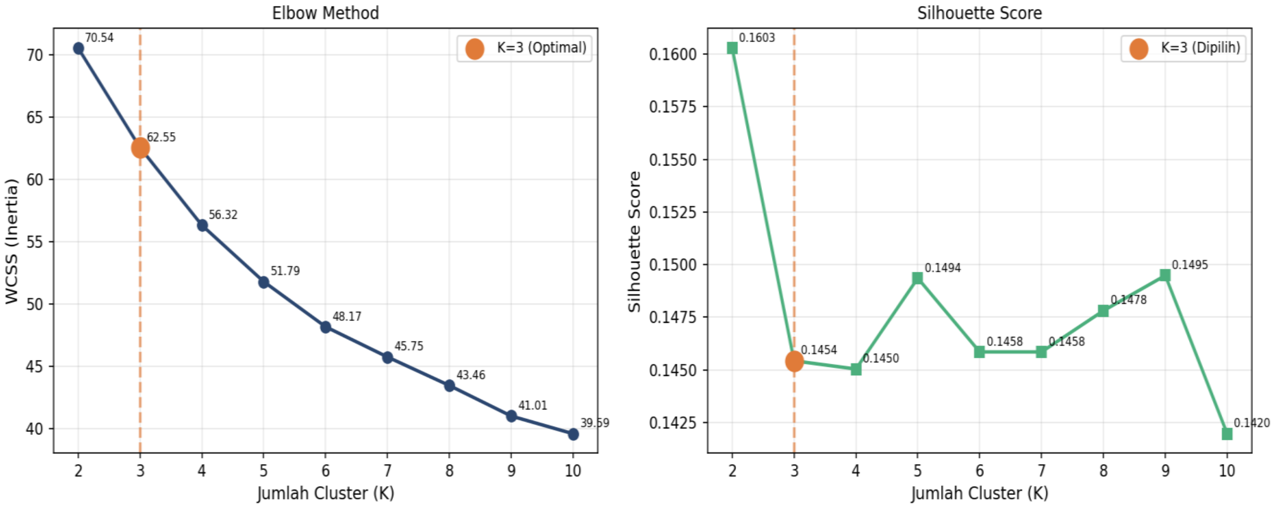

Figure 6. Elbow and Silhouette Evaluation Results

The WCSS decreased from 70.5354 at $K = 2$ to 62.5475 at $K = 3$, a reduction of approximately 7.99. The reductions became progressively smaller as additional clusters were introduced. The highest Silhouette Score was 0.1603 at $K = 2$, whereas the score at $K = 3$ was 0.1454. Based primarily on the elbow pattern and the interpretability of the resulting profiles, three clusters were selected for the final analysis. The relatively low Silhouette Score indicated limited separation among the resulting clusters. The K-means algorithm was executed with random_state=42 and converged after nine iterations. The final cluster profiles are presented in Table 8 and Figure 7.

Table 8. Final K-Means Cluster Profiles for K = 3

Cluster	Study-Assigned Label	n	Percentage	SUS	Ease of Use	Efficiency	System Reliability	User Satisfaction
C0	Fairly Satisfied	168	30.5%	60.09	50.63	56.47	52.72	63.06
C1	Somewhat Dissatisfied	201	36.5%	47.61	55.44	51.40	62.84	43.31
C2	Dissatisfied	181	32.9%	43.69	58.98	52.00	52.59	67.85

Note. All profile scores are presented on a scale ranging from 0 to 100. Percentages may not sum to exactly 100% because of rounding.

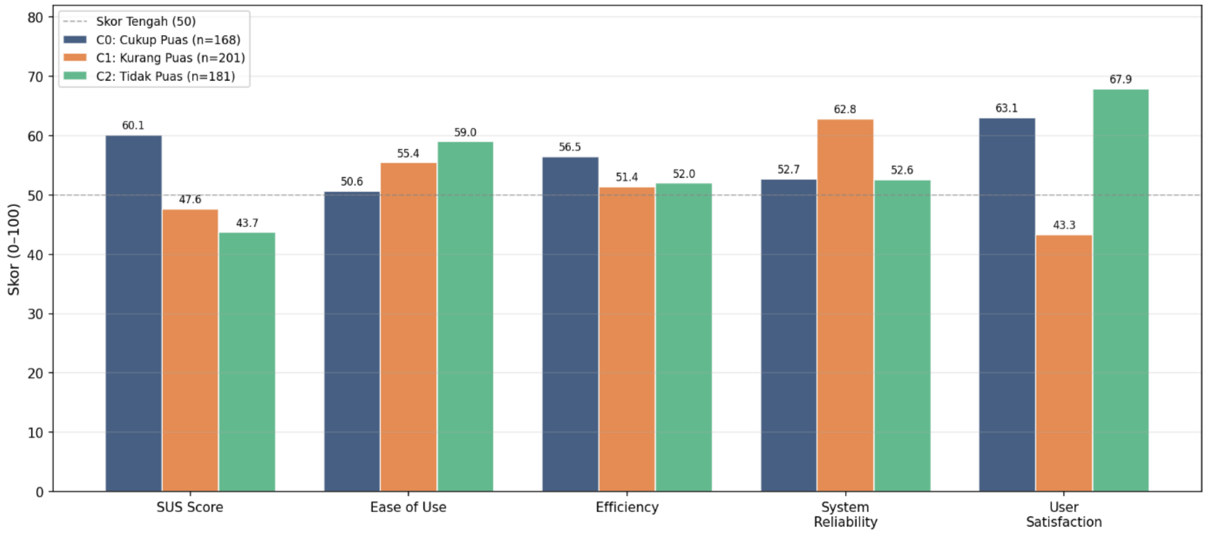

Figure 7. Comparison of SUS and UX Dimension Scores Across Clusters

Cluster 0 comprised 168 respondents and had the highest mean SUS score at 60.09. It also had the highest Efficiency score at 56.47 among the three clusters. Cluster 1 was the largest group, comprising 201 respondents, and had the highest System Reliability score at 62.84 but the lowest User Satisfaction score at 43.31. Cluster 2 comprised 181 respondents and had the lowest mean SUS score at 43.69. However, it had the highest Ease of Use score at 58.98 and the highest User Satisfaction score at 67.85. Figure 8 presents a two-dimensional PCA visualization of the cluster assignments. The first principal component explained 24.1% of the variance, while the second explained 21.5%. Together, the two components represented 45.6% of the variance in the five clustering features.

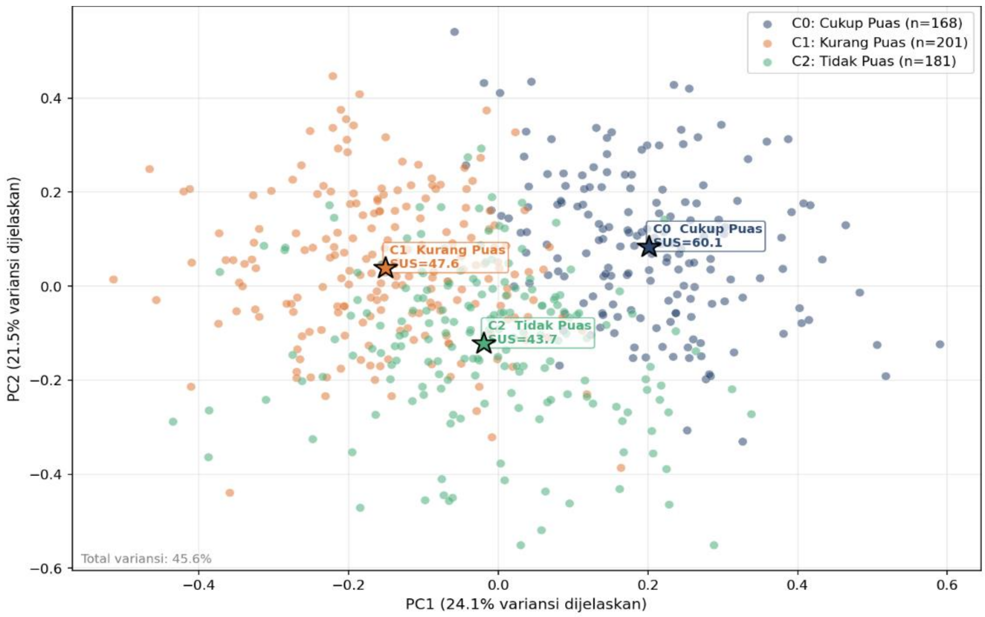

Figure 8. PCA Visualization of the K-Means Clustering Results

Each point in Figure 8 represents one respondent, and the colors indicate cluster membership. The three clusters overlapped substantially, particularly in the central area of the plot. Cluster 0 tended to appear toward the right side of the projection, whereas Clusters 1 and 2 were more concentrated in the center-left area. The observed overlap was consistent with the relatively low Silhouette Score. Because the two principal components represented 45.6% of the total variance, the visualization provides only a partial two-dimensional representation of the original five-dimensional data.

4.2 Discussion

The mean SUS score of 50.13 indicates that respondents generally reported a low level of usability for the Mobile JKN application under the interpretation framework adopted in this study. More than half of the respondents were classified in the Poor category, while only five respondents reached the Good category and none reached the Excellent category. These results suggest that the application did not consistently provide an acceptable usability experience for the surveyed users. The item-level results provide additional context for the aggregate score. Q7 obtained the highest mean response among the positive items, suggesting that respondents perceived the application as potentially learnable. However, the lower response to Q3 indicates that respondents did not evaluate their direct experience of ease of use as positively. The relatively high responses to the negatively worded Q6 and Q8 further indicate perceived inconsistency and difficulty in using the application. These response patterns should be interpreted as differences among related aspects of usability rather than as a direct contradiction between items.

Among the four UX dimensions, Efficiency received the lowest converted score at 53.15. Ease of Use also remained relatively low at 55.14. These findings indicate that task completion and interaction with application features were among the areas receiving the least favorable evaluations. System Reliability and User Satisfaction received somewhat higher scores, although neither dimension substantially exceeded the midpoint of the converted scale. The findings therefore suggest that the usability problems reported by respondents were not limited to one dimension. The three-cluster solution revealed variation in how the SUS and UX dimensions were combined across respondents. Cluster 0 had the highest mean SUS score and the highest Efficiency score, supporting its designation as the relatively more favorable group. However, its Ease of Use and System Reliability scores remained lower than some scores observed in the other clusters, indicating that a higher SUS score did not correspond to uniformly high scores across all UX dimensions. Cluster 1 had the highest System Reliability score but the lowest User Satisfaction score. This pattern indicates that a relatively favorable perception of reliability did not necessarily coincide with satisfaction. The available data do not establish why this difference occurred. It may reflect aspects of service value, feature suitability, task outcomes, or other factors that were not directly measured. Additional qualitative evidence would be required to determine the reasons for this profile.

Cluster 2 had the lowest mean SUS score but the highest Ease of Use and User Satisfaction scores. Its Efficiency and System Reliability scores were lower than its Ease of Use and User Satisfaction scores. This mixed profile shows that the study-assigned label "Dissatisfied" does not fully represent all dimensions within the cluster. The label should therefore be interpreted primarily in relation to its low SUS score rather than as evidence that respondents rated every UX dimension negatively. The profile indicates that cluster interpretation should be based on the complete set of features rather than on a single label. The low Silhouette Score and overlap in the PCA visualization indicate that the boundaries among the three clusters were weak. The Elbow Method suggested diminishing WCSS reductions after three clusters, but the highest Silhouette Score was obtained at $K = 2$. The selection of three clusters was therefore based on the elbow pattern and interpretability rather than superior separation according to the Silhouette Score. The resulting profiles can be used as exploratory descriptions of respondent patterns, but they should not be interpreted as sharply separated or definitive user types.

Aziza and Ristriani (2023) reported a usability result of 62.25% using usability and heuristic evaluation methods, whereas the present study obtained a mean SUS score of 50.13. These values should not be compared as equivalent measurements because the studies used different instruments, samples, and evaluation periods. Nevertheless, both studies identified efficiency-related issues as an area requiring attention. The present results also indicate perceived inconsistency and difficulty of use, which correspond to concerns reported in the earlier evaluation. Nurzanah *et al.* (2026) obtained a Silhouette Score of 0.781 when applying SUS-based X-means clustering to the Candy CBT application. The score is substantially higher than the value of 0.1454 obtained for the three-cluster solution in this study. However, the difference does not establish that one clustering approach is generally superior because the studies examined different applications, respondents, features, and data distributions. In the present study, the lower value indicates that respondents' evaluations were distributed without clearly separated group boundaries.

The findings provide several practical considerations for the Mobile JKN development team. Efficiency received the lowest dimension score, making service workflows, response times, and task-completion processes relevant areas for further evaluation. Cluster 1 indicates that perceived reliability alone may not be sufficient to produce satisfaction, whereas Cluster 2 suggests that relatively favorable evaluations of ease of use and satisfaction can coexist with lower SUS, efficiency, and reliability scores. These patterns support differentiated investigation of usability problems, although technical causes should be verified through usability testing, system performance measurements, or qualitative interviews before specific design or backend changes are implemented. Data were collected online, and the researchers could not directly control the conditions under which respondents completed the questionnaire. The study also did not verify whether respondents were using the application while answering or relying on prior experience. Demographic characteristics, frequency of use, and duration of experience with the application were not included in the clustering features. These factors may be associated with differences in usability perceptions. The low Silhouette Score limits the strength of the cluster interpretation. The identified groups should therefore be treated as exploratory profiles rather than definitive user segments. Future studies could combine questionnaire data with interviews or task-based usability testing to clarify the reasons underlying each profile. Longitudinal evaluation could also be used to examine changes in SUS and UX dimension scores following application updates. Comparisons with other clustering methods, including hierarchical clustering and DBSCAN, may help assess the robustness of the identified patterns.

5. Conclusion and Recommendations

This study evaluated the usability and user experience of the Mobile JKN application using the System Usability Scale (SUS) and K-means clustering based on responses from 550 users in Medan. The mean SUS score was 50.13, which was classified as Poor and Grade D under the interpretation framework adopted in this study. Among the four UX dimensions, Efficiency received the lowest score at 53.15, followed by Ease of Use at 55.14, System Reliability at 56.38, and User Satisfaction at 57.42. These results indicate that efficiency and ease of use were the least favorably evaluated dimensions. The three-cluster solution produced distinct combinations of SUS and UX dimension scores. Cluster 0, comprising 168 respondents, had the highest mean SUS score at 60.09 and the highest Efficiency score among the three groups. Cluster 1, comprising 201 respondents, had the highest System Reliability score but the lowest User Satisfaction score. Cluster 2, comprising 181 respondents, had the lowest mean SUS score but the highest Ease of Use and User Satisfaction scores. These mixed profiles show that respondents with similar overall SUS scores did not necessarily report the same perceptions across the four UX dimensions. The findings demonstrate that an aggregate SUS score alone does not represent all variations in respondents' evaluations. Combining SUS with K-means provided exploratory profiles that revealed different patterns across ease of use, efficiency, system reliability, and user satisfaction. However, the low Silhouette Score of 0.1454 indicates that the boundaries among the three clusters were weak. The profiles should therefore be interpreted as exploratory patterns rather than definitive or sharply separated user segments.

Based on the dimension scores, further evaluation of the Mobile JKN application should prioritize task-completion efficiency and ease of locating and using its features. Task-based usability testing could be conducted to identify stages in the service workflow that require excessive time or effort. Perceived differences in system reliability and satisfaction should also be examined through qualitative interviews and technical performance testing before specific interface or system changes are implemented. This study was limited by its online data-collection procedure, the absence of demographic variables in the clustering analysis, and the lack of verified information about respondents' frequency and duration of application use. The low separation among clusters also limits the generalizability of the resulting profiles. Future studies should incorporate demographic and usage-related variables, validate the cluster structure using additional samples, and combine questionnaires with interviews or task-based usability testing. Longitudinal evaluation could be used to assess changes in user experience following application updates. Comparisons with other clustering methods, such as hierarchical clustering and DBSCAN, may also help assess the robustness of the identified patterns.

References

Akbar Mokoginta, J., Palinggi, J., Margareta, D., Rindengan, Y. D. Y., & Ade, Y. (2025). Digitalisasi layanan BPJS Kesehatan dalam transformasi e-government. *Kohesi: Jurnal Sains dan Teknologi, 10*(6). https://cibinstitute.id/index.php/kohesi/article/view/2103

Aziza, R. F. A., & Ristriani, P. (2023). Measuring UX using usability and heuristic methods in JKN Mobile application. *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer), 9*(1), 96–101. https://doi.org/10.33480/jitk.v9i1.4070

Ernawati, S., & Rahayu, S. (2023). Analisa usability pada aplikasi Identitas Kependudukan Digital menggunakan metode usability testing. *BIOS: Jurnal Teknologi Informasi dan Rekayasa Komputer, 5*(1), 12–19. https://doi.org/10.37148/bios.v5i1.87

I Komang Adyanata, G. A. P. N., Gunawan, I. M. A. O., & Indrawan, G. (2024). Evaluasi sistem informasi SIAkad UPMI menggunakan metode System Usability Scale. *Jurnal JUPITER, 16*(1), 61–70.

Irwan, I., & Putra Rakhmadani, D. (2023). Analisis penggunaan aplikasi Mobile JKN dengan metode EUCS. *Remik: Riset dan E-Jurnal Manajemen Informatika Komputer, 7*(3), 1491–1502. https://doi.org/10.33395/remik.v7i3.12826

Maori, N. A., & Evanita, E. (2023). Metode Elbow dalam optimasi jumlah cluster pada K-Means clustering. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer, 14*(2), 277–288. https://doi.org/10.24176/simet.v14i2.9630

Nugroho, R. R., & Usman, O. (2025). Analyzing mobile app design's impact on Instagram user experience and satisfaction. *International Student Conference on Business, Education, Economics, Accounting, and Management (ISC-BEAM), 3*(1), 202–215. https://doi.org/10.21009/isc-beam.013.13

Nurzanah, L., Zufria, I., & Chit, S. C. (2026). X-Means clustering for UX evaluation of the Candy CBT application using the SUS instrument. *Inform: Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi, 11*(1), 71–79. https://doi.org/10.25139/inform.v11i1.11277

Prayitno, E., Perdana, I. J., & Nasyuha, A. H. (2025). Implementasi data mining dan machine learning untuk segmentasi pelanggan: Pendekatan hybrid menggunakan big data. *Jurnal Ilmiah FIFO, 17*(1), 57. https://doi.org/10.22441/fifo.2025.v17i1.007

Putri, S. I., & Liu, K. (2024). Assessing Ticket.com app usability through the System Usability Scale (SUS) method. *International Journal for Applied Information Management, 4*(1), 30–40. https://ijaim.net/journal/index.php/ijaim/article/view/73

Saputra, A. B., Cahyo, P. W., Habibi, M., & Priadana, A. (2022). Analysis and visualization of BPJS on Twitter using K-Means clustering. *International Journal of Health Science and Technology, 3*(3). https://doi.org/10.31101/ijhst.v3i3.2466

Shobri, M. Q., Al-Kubro, P. B., & Kawuri, G. V. (2025). Cluster analysis of BPJS Kesehatan claim data in Madiun City to identify high claim patterns and fraud indications. *Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi, 13*(3), 419–424.

Ulum, K., Wahyudi, A., & Trinovita, D. (2024). Analisis pemanfaatan sistem informasi Mobile JKN di RSU Az-Zahra menggunakan User Experience Questionnaire (UEQ). *Journal of Computer Science and Informatics (JOCSI), 2*(1), 32–37.

Yani, R., Nazhifah, I., & Pradika, M. I. (2025). System Usability Scale in information system application development using systematic mapping study. *Indonesian Journal of Applied Technology and Innovation Science, 2*(2), 130–140.

Zulpa Salsabila, Z., Halim, F., Damanik, R., & Cinta, T. P. (2024). Menggunakan metode UEQ+ dalam menganalisis dan mengevaluasi pengalaman pengguna Mobile JKN. *JSAI (Journal Scientific and Applied Informatics), 7*(3), 553–560. https://doi.org/10.36085/jsai.v7i3.7258