Available online at www.coins-indocompt.org

COINS JOURNAL
Computing and Information System Journal

COmputing and INformation Systems Journal

IndoCompt PUBLISHER

E-ISSN: 3109-3248

Vol. 2, No. 2, August, 2026

Optimization of the Promotion Strategy for Madrasah Aliyah Karya Pembangunan Al-Hidayah Islamic Boarding School through Data Analysis Using the K-Means Method

Arief Rahmat Putra[a*], Sukma Puspitorini[b]

[a,b]*Departement of Informatics, Faculty of Computer Science, Nurdin Hamzah University, Jambi 36124, Indonesia*

Abstract

The Al-Hidayah Islamic Boarding School (PKP) faces challenges in attracting new prospective students every year in the midst of competition from private and public schools, budget limitations, and negative issues of Islamic boarding schools that are rampant on social media. This research aims to optimize data-driven promotion strategies through the application of Data Mining techniques, especially K-Means algorithms, to map the potential of promotion target areas and produce more effective policies. The analysis was carried out by utilizing the data of the origin of students using Python-Scikit-Learn to group based on the region of origin. The findings show that the K-Means method is effective in identifying potential areas to be prioritized for promotion, so that promotional strategies can be expanded and become more efficient. The results of this study are expected to assist PKP Al-Hidayah in increasing the number of new registrants in a sustainable manner through more targeted and evidence-based promotion decisions.

Keywords: K-Means; Data analysis; Promotion Strategy; Data Mining

1. Introduction

Every year, both private and public schools compete to attract prospective students. This competition is increasingly fierce due to budget constraints, rapid shifts in parent and student preferences, and a diverse range of offline and digital communication channels, which often render traditional promotional approaches off-target and resource-intensive. This challenge is further compounded by the widespread circulation of negative narratives regarding the education system in Islamic boarding schools (pondok pesantren) across social media platforms. In light of these conditions, this research is conducted at Pondok Pesantren Karya Pembangunan (PKP) Al-Hidayah to optimize a data-driven promotional strategy.

Pondok Pesantren Karya Pembangunan (PKP) Al-Hidayah is an Islamic boarding school institution owned by the Jambi Provincial Government, established pursuant to the Jambi Governor's Decree Number 228 of 1983.[1] This institution aims to cultivate development cadres who are knowledgeable, charitable, pious, and skilled. However, in recent years, PKP Al-Hidayah has faced fluctuations in the number of new students (*santri*), indicating that its promotional strategies are not yet optimal. Based on admission data, the number of new students in the 2023-2024 academic year was recorded at 230, decreasing to 185 in 2024-2025, and further declining to 184 in 2025-2026. This pattern demonstrates that the success of promotional efforts remains inconsistent from year to year. One of the primary causes is a promotional strategy that continues to focus on the area immediately surrounding the boarding school, thereby failing to reach other potential regions.

In addition to budget constraints and shifting public preferences, challenges also arise from promotional methods being heavily concentrated in the school's vicinity. This renders traditional promotional strategies, such as distributing brochures and participating in educational exhibitions, frequently ineffective and inefficient in reaching the target audience. Data limitations pose a major obstacle, as PKP Al-Hidayah does not yet maintain records of the new students' previous schools or possess an adequate database, meaning that several analytical methods cannot be utilized optimally.

To address these issues, the application of Data Mining techniques presents a potential solution. Data mining is a process that utilizes statistical, mathematical, artificial intelligence, and machine learning methods to extract and identify valuable information from various large databases.[2] Data Mining is defined as the process of mining data, or the effort to extract valuable and useful information from very large databases.[3] Meanwhile, another definition describes data mining as the process of extracting large volumes of data or previously unknown information from massive databases so that it becomes comprehensible, valuable, and applicable for making critical business decisions (Sumadikarta, 2023).[4] Through this technique, the school can transform operational and demographic data into strategic insights, such as the segmentation of prospective students, response patterns to promotional media, and enrollment rate predictions, which can support the formulation of efficient and targeted promotional policies. The K-Means clustering method is known to be effective in grouping data based on feature similarity, enabling promotional strategies to be targeted according to regional potential.[5] According to Prianto & Bunyamin (2020, p. 18), the K-Means clustering algorithm is a partitional algorithm because it relies on determining the initial number of clusters by defining their initial centroid values.[6]

The researcher contributes by implementing the K-Means method to support more targeted promotional decision-making. The analysis is conducted on the students' origin data using Python and the Scikit-Learn (sklearn) library, which offers comprehensive documentation, intuitive syntax, and extensive community support.[7] Hasil pengelompokan diharapkan dapat memberikan gambaran daerah-daerah prioritas promosi.

Based on this background, this study aims to optimize the marketing strategy of Madrasah Aliyah PKP Al-Hidayah through data analysis using the K-Means method, with the hope of helping the institution sustainably improve promotional efficiency and the number of new applicants.

Table 1. Literature Review

Author & Year	Heading	Purpose	Excess	Deficiency
Latifah et al., 2024	Implementasi Algoritma K-Means Clustering untuk Strategi Promosi Kampus IBISA	Classify prospective new IBISA students based on data to determine effective promotion strategies.	The test was performed three times with random centroid values, resulting in an accurate and stable cluster. Encourage the expansion of promotions beyond the main regions.	It does not discuss the integration of external data and has not tested the results of the promotional strategy after the implementation of the algorithm.
Syahputra, 2020	Menentukan Strategi Promosi Menggunakan Algoritma Clustering K-Means di SMA dan SMK Harapan Bangsa	Identify clusters of prospective students by school/major of origin to determine promotion strategies based on admissions data.	Provide two promotion strategies based on cluster results: promotion based on your favorite major and your home school.	The data is from only one institution and has not compared the effectiveness of strategies before and after the use of K-Means clusters.
Nurul Huda University, 2024	Penerapan Metode K-Means Clustering Dalam Pengembangan Strategi Promosi Berbasis Data Penerimaan Mahasiswa Baru	Using the K-Means method to group data of prospective new students so that promotions can be directed to high-potential areas.	Using a strategic KDD and generating three different clusters that are used to effectively expand the reach of the promotion to potential areas.	A specific case study at a particular university, the results have not been tested in a high school or non-college context.

Previous studies exhibit several major limitations, including research data confined to a single educational institution, a lack of external data integration testing, and the absence of evaluations regarding the effectiveness of promotional strategies before and after K-Means implementation. Furthermore, some studies focus exclusively on specific educational tiers (such as universities), meaning their findings are less empirically verified for the context of Islamic boarding schools or secondary schools. This study also incorporates an analysis of promotional area expansion by combining demographic segmentation with regional potential mapping, rather than clustering solely based on the institution's internal data. Additionally, this

research is conducted within a regional government-owned Islamic boarding school environment, a setting that presents unique challenges and has not been widely investigated.

2. Research Methodology

This research employs a case study approach, focusing on Pondok Pesantren Karya Pembangunan (PKP) Al-Hidayah as the research object. The analytical development method applied is Data Mining based on the K-Means Clustering algorithm, as this method is considered effective for grouping student origin data through an iterative cycle. This accurately produces regional segmentation in accordance with the institution's promotional strategy needs, following these sequential steps:

a. Preprocessing data
in this phase, a data cleaning process is conducted on the new student admission data of PKP Al-Hidayah by selecting relevant variables and eliminating those unnecessary for the analysis. Subsequently, all categorical data is converted into a numerical format to enable processing during the clustering phase using the K-Means algorithm. This stage is crucial to ensure that the data utilized is thoroughly clean, consistent, and ready to form regional clusters of prospective students, which serve as an accurate foundation for the boarding school's promotional strategy recommendations.

b. Data Normalization
The subsequent step involves normalizing the student registration data at PKP Al-Hidayah from the last three years. Data normalization is the process of adjusting several variables, so they fall within the same range of values [8]. This data normalization utilizes the Min-Max formula, as shown below:

$$x_{new} = \frac{x_{old} - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Every registered student's data (such as the total count, school of origin, and other relevant variables) is scaled to a range of 0 to 1 prior to clustering [8]. This stage is essential to ensure that the implementation of the K-Means algorithm can optimally cluster the data in alignment with the research objectives, namely mapping priority promotional areas and enhancing the efficiency of the new student admission strategy at Pondok Pesantren Karya Pembangunan Al-Hidayah.

c. Determining the Number of Clusters
There are several methods for determining the number of clusters, one of which is utilizing the elbow diagram. The elbow diagram compares the inertia values (Sum of Squared Errors, or SSE) against the number of clusters to identify the optimal number of clusters.

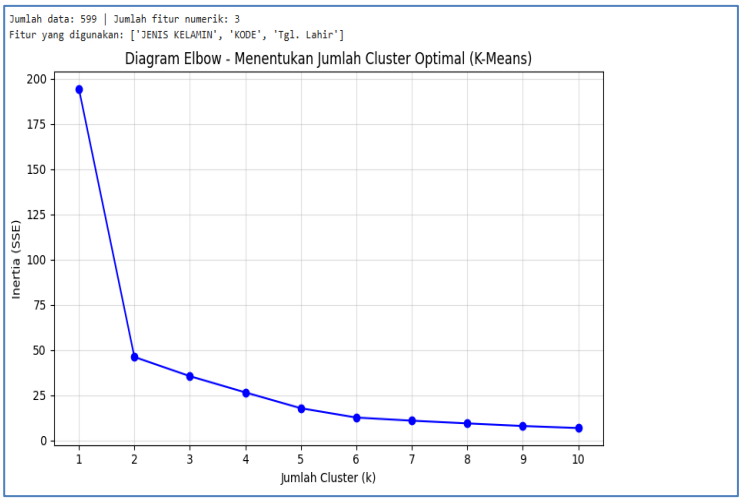

Figure 1. Elbow diagram graph

By selecting a K value that corresponds to the inflection point, we can determine the most appropriate number of clusters for our data.[9]

d. Determining the Initial Centroids
In this study, initial centroids are determined randomly to ensure an even distribution of center points and to achieve more optimal clustering results. Subsequently, the algorithm iteratively updates the centroid positions until a

stable final result is reached.

e. Initiazilation
In this study, the Euclidean distance dissimilarity metric is employed. The formula is as follows:

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (2)$$

description:
$d\ (x, y)$ = distance from data point x to data point y
xi = value of the i-th variable of data x
y_i = value of the i-th variable of data y[10]

The iterative process is carried out until no data points switch clusters or until a state of convergence is reached. Convergence refers to a condition where there are no significant changes or reassignments of data points among the clusters [10]

f. Visualization of Clustering Results
In the visualization phase of the clustering results, the grouped data obtained by applying the K-Means algorithm to the new student admissions at PKP Al-Hidayah is visualized in the form of graphs or plots. This visualization aims to illustrate the spread and distribution of data within each formed cluster, enabling the boarding school management to easily identify prospective student regions that share similar characteristics and have high potential to be prioritized for promotion. In this way, the clustering results become easier to interpret and can be directly utilized as a foundation for making targeted promotional strategy decisions.

g. Data Analysis for Each Cluster
The data analysis for each cluster is conducted by observing the characteristic profiles of the students within each group formed by the K-Means algorithm. At this stage, each cluster is analyzed based on primary variables such as region of origin, gender, and age. The analytical results provide a specific overview of areas with a high potential for applicants, regions with low applicant growth, as well as territories that have not been adequately reached by the boarding school's promotional strategies thus far. By comparing the characteristics of each cluster, the management of PKP Al-Hidayah can design more targeted promotional policies, thereby optimizing the potential to increase the number of new students.

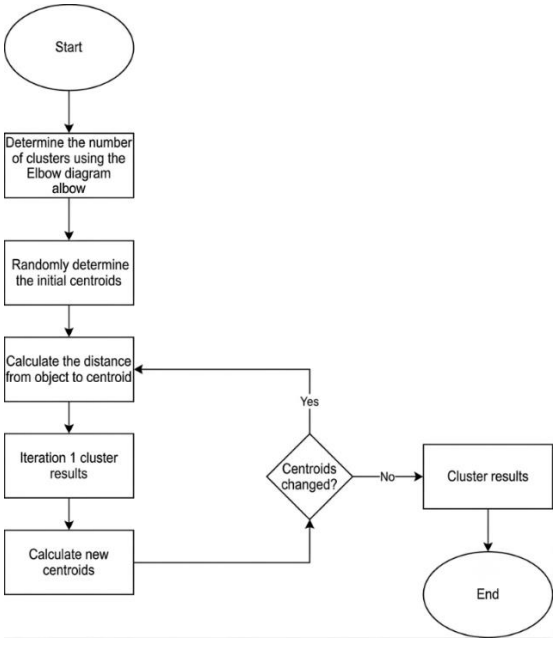

Figure 2. Research flow diagram

3. Results and Discussion

3.1. Preprosessing Data

During the data preprocessing phase, data collection and identification were carried out. The data collection was conducted during the research period from August 11 to September 30, 2025, in coordination with Ms. Ayu, the administrative staff member of PKP Al-Hidayah. The acquired dataset consists of 599 records detailing the number of new students from the 2023 to 2025 academic years. Table 2 presents the student data of PKP Al-Hidayah.

Table 2. Student data

No	Gender	Address (Region)	Age
1	Male	Musi Banyuasin	15
2	Male	Sarolangun	15
3	Male	Kota Jambi	15
4	Male	Kota Jambi	15
5	Male	Tanjung Jabung Barat	15
….	….	…..	….
595	Female	Kerinci	17
596	Female	Kota Jambi	17
597	Female	Tanjung Jabung Timur	17
598	Female	Muaro Jambi	17
599	Female	Tanjung Jabung Timur	17

Table 2 presents the student data that has been cleaned, featuring the variables selected for use. Subsequently, the data will undergo a type conversion process commonly referred to as encoding as shown in Table 3 below.

Table 3. Data transformation

No	Gender	Addres(region)		Age
1	0	22		15
2	0	2		15
3	0	3		15
4	0	3		15
5	0	4		15
….	….	…..		….
595	1	9		17
596	1	3		17
597	1	6		17
598	1	7		17
599	1	6		17

Table 4. Data normalization

No	jenis kelamin	Alamat	Umur
1	0	0,344262295	0,2000
2	0	0,016393443	0,2000
3	0	0,032786885	0,2000
4	0	0,032786885	0,2000
5	0	0,049180328	0,2000
….	….	…..	….
595	1	0,032786885	0,6000
596	1	0,081967213	0,6000
597	1	0,098360656	0,6000
598	1	0,081967213	0,6000
599	1	0,147540984	0,6000

This data encoding is conducted by assigning a value of 0 to male students and 1 to female students. Meanwhile, the address variable is transformed based on distinct codes assigned to each regency or city.

Following this transformation, the data is normalized to standardize the range of each variable using equation (1), resulting in a data range of 0 to 1 for each variable, as can be seen in the table 4.

3.2. Implementation

The initial stage in this clustering process is to determine the optimal number of clusters. This determination is carried out using the Elbow Method by analyzing the inflection point (elbow point) from the student dataset, as illustrated in the figure below:

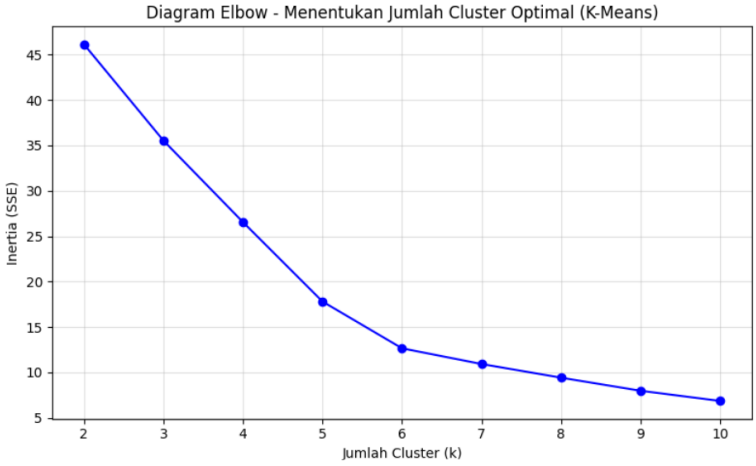

Figure 3. Elbow diagram.

Evaluation using the Elbow diagram demonstrates that the inflection point occurs at K = 6, indicating that the optimal number of clusters is 6. The subsequent step is the initialization of the initial centroids, where three points are randomly selected from the dataset to act as cluster centers in accordance with the predetermined value of K. The initial centroids for this clustering are:

Table 5. Initial centroids

Centroid 1	0	0.180327869	0.2
Centroid 2	0	0.049180328	0.2
Centroid 3	1	0.163934426	0.2
Centroid 4	1	0.081967213	0.6
Centroid 5	0	0.196721311	0.6
Centroid 6	1	0.049180328	1

3.2.1. Clustering Results

The results of the clustering process on the new student admission data at PKP Al-Hidayah utilizing the K-Means algorithm demonstrate the application of three primary numerical features, namely 'JENIS KELAMIN' (Gender), 'Alamat' (Address), and 'Usia' (Age), comprising a total of 599 data rows and 3 analyzed variables. Once the cluster centers are determined, the subsequent stage involves calculating the distance of the entire dataset to the cluster centers using equation (ii). This distance calculation is carried out continuously until each cluster reaches a stable and convergent state. This clustering process underwent 11 iterations, yielding the following results:

Table 6. Clustering results

No	Gender	Address	Age	Cluster
1	0	0,344262295	0,2000	1
2	0	0,016393443	0,2000	1
3	0	0,032786885	0,2000	1
4	0	0,032786885	0,2000	1
5	0	0,049180328	0,2000	1
….	….	…..	….	….

595	1	0,032786885	0,6000	5
596	1	0,081967213	0,6000	5
597	1	0,098360656	0,6000	5
598	1	0,081967213	0,6000	5
599	1	0,147540984	0,6000	5

The initial centroids, which were selected randomly, are located at two different points, whereas the final centroids resulting from normalization exhibit a significant positional shift, representing the distribution of data characteristics within each cluster. Further analysis of the formed clusters reveals that Cluster 1 consists of 41 data points with an average centroid value of 0.1669; Cluster 2 contains 93 data points with an average feature value of 0.1246; Cluster 3 contains 169 data points with an average feature value of 0.4684; Cluster 4 contains 26 data points with an average feature value of 0.7765; Cluster 5 contains 135 data points with an average feature value of 0.2723; and Cluster 6 contains 135 data points with an average feature value of 0.5838. This indicates the existence of distinct profile patterns among prospective new students, which can serve as a foundation for determining priority regions or target groups to execute PKP Al-Hidayah's promotional strategies more effectively and in a segmented manner.

3.2.2. Visualization of Clustering Results

Based on the clustering results, a total of 599 observations were successfully and exclusively mapped to their respective nearest centroids (C1 to C6). Proportion analysis indicates a variance in the distribution of member counts across each cluster. Cluster 3 (Blue) represents the majority group with the largest population, comprising 169 observations. Conversely, Cluster 4 (Orange) constitutes the minority group with the smallest distribution, consisting of 26 data points. Meanwhile, Cluster 5 (Purple) and Cluster 6 (Light Blue) exhibit a symmetrical membership level, each containing 135 observations, followed by Cluster 2 (Green) with 93 observations and Cluster 1 (Red) with 41 observations.

Spatially, the positions of the centroids (C1-C6) demonstrate a well-defined degree of separation among the groups within the t-SNE space. Each cluster successfully groups data points that exhibit closely related characteristics. The identification of distinct profiles between the majority group (such as Cluster 3) and specific groups (such as Cluster 4) provides a solid quantitative foundation. Mapping student characteristics through this spatial segmentation is highly crucial for formulating the educational institution's promotional strategy designs, rendering them more targeted, efficient, and customized according to the dominant trend patterns within each cluster.

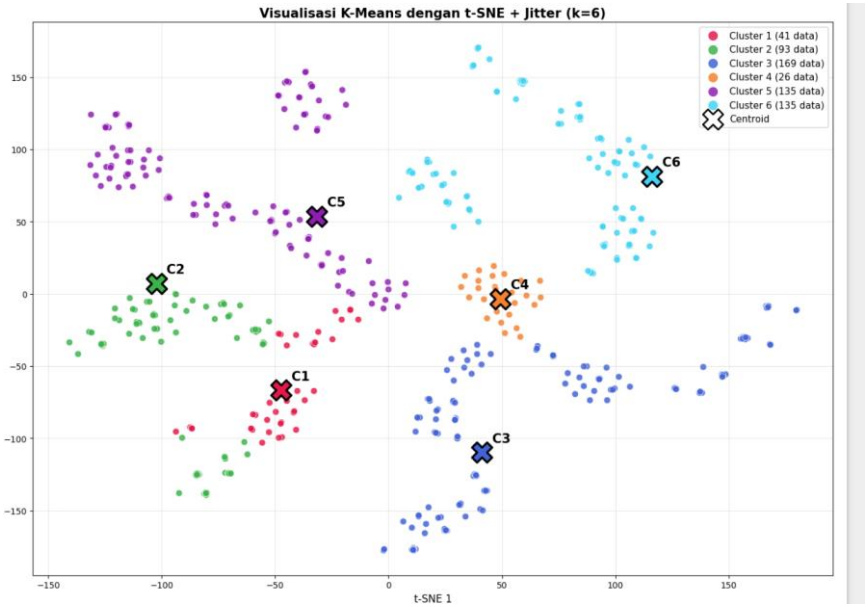

Figure 4. Clustering Visualization Results

3.2.3. Data Analysis for Each Cluster

The results of the data profile analysis for each cluster (namely Cluster 1 and Cluster 6) are presented in Table 7 below.

Table. 7 Analysis of clustering results

Cluster	Number of Individuals	Gender	Age	Predominant Regions of Origin
1	41	Male (100%)	14–16 th (avg 15,3)	Kerinci, Merangin, Batang Hari, Muaro Jambi
2	93	Male (100%)	14–16 th (avg 15,6)	Kota Jambi, Sarolangun, Tanjung Jabung Barat, Tanjung Jabung Timur
3	169	Female (100%)	15–16 th (avg 15,5)	Kota Jambi, Sarolangun, Tanjung Jabung Barat, Tebo
4	26	Female (100%)	16–18 th (avg 17,0)	Sebaran luas (22 daerah), banyak dari luar Jambi (Indragiri Hilir, Muara Enim, Rokan Hilir, Tasikmalaya, dll)
5	135	Male (100%)	17–19 th (avg 17,3)	Kota Jambi, Muaro Jambi, Sarolangun, Tanjung Jabung Barat, Kerinci
6	136	Female (100%)	17–19 th (avg 17,3)	Kota Jambi, Tanjung Jabung Barat, Tanjung Jabung Timur, Muaro Jambi, Batang Hari

The profile analysis results of Cluster 1 indicate that this group consists of 41 students, all of whom are male, in the early adolescent age range. The majority of the cluster members originate from Jambi Province, with the largest distribution in the Kerinci, Merangin, Batang Hari, and Muaro Jambi Regencies, while a small portion (approximately 29.3%) comes from outside the province, primarily from Indragiri Hilir and Sragen. The age distribution in this cluster has an average of 15.29 years, ranging from 14 to 16 years, with the majority falling in the 15-year-old age group. These characteristics illustrate a group of young male students who tend to originate from the western inland regions of Jambi, with a minor recruitment reach from outside the province.

The profile analysis of Cluster 2 reveals a group comprising 93 male students with a dominant age of around 16 years. Most of the members originate from Jambi Province, particularly Jambi City, Sarolangun, West Tanjung Jabung, and East Tanjung Jabung, while only about 4.3% come from outside the province, dominated by Musi Banyuasin. This cluster has an average age of 15.62 years, ranging from 14 to 16 years, with the highest frequency in the 16-year-old age group. These characteristics indicate a local male student group that is highly homogeneous geographically and can be identified as a core recruitment segment around Jambi City and the eastern coastal areas.

The profile analysis results of Cluster 3 show a group consisting of 169 female students with a dominant age range of 15 to 16 years. The majority of the members originate from Jambi Province, with the largest distribution in Jambi City, Sarolangun, West Tanjung Jabung, and Tebo, whereas approximately 8.9% come from outside the province, primarily Musi Banyuasin. This cluster possesses an average age of 15.49 years, ranging from 15 to 16 years, with the majority in the 15-year-old age group. These characteristics depict a group of young female students with a strong and evenly distributed recruitment base across several main regencies in Jambi Province.

The profile analysis of Cluster 4 identifies a small group comprising 26 female students with a dominant age of around 17 years. Unlike the other clusters, the vast majority of this cluster's members (approximately 96.2%) actually originate from outside Jambi Province, with a fairly even distribution across various regions such as Indragiri Hilir, Muara Enim, and Rokan Hilir. This cluster has an average age of 16.96 years, ranging from 16 to 18 years, with the highest frequency in the 17-year-old age group. These characteristics indicate a female student group with a broad cross-provincial promotional reach, identifiable as a unique segment resulting from out-of-region recruitment.

The profile analysis results of Cluster 5 reveal a group consisting of 135 male students in the middle to late adolescent age range. The majority of members originate from Jambi Province, with the largest spread across Jambi City, Muaro Jambi, Sarolangun, West Tanjung Jabung, and Kerinci, while about 13.3% come from outside the province, mainly Indragiri Hilir. This cluster has an average age of 17.31 years, ranging from 17 to 19 years, with the majority falling in the 17-year-old age group. These characteristics describe a group of senior male students distributed across many regions of Jambi, indicating a mature recruitment base that expands as age increases.

The profile analysis of Cluster 6 presents a group comprising 136 female students with a dominant age of around 17 years. The majority of its members originate from Jambi Province, particularly Jambi City, West Tanjung Jabung, East Tanjung Jabung, Muaro Jambi, and Batang Hari, whereas only about 2.2% come from outside the province, dominated by Musi Banyuasin. This cluster possesses an average age of 17.27 years, ranging from 17 to 19 years, with the highest frequency in the 17-year-old age group. These characteristics denote a geographically highly homogeneous senior female student group, reflecting the stability of local female student recruitment in the eastern and central regions of Jambi.

In general, the six formed clusters demonstrate a grouping pattern based on a combination of gender and age stages. Clusters 1, 2, and 3 represent the younger student segment (14–16 years old), while Clusters 5 and 6 represent the senior student segment (17–19 years old), which is homogeneous and strongly based in Jambi Province. Meanwhile, Cluster 4 stands out as a unique segment with a broad cross-provincial recruitment reach. The resulting profiles can serve as a vital reference for the institution's management in formulating more targeted promotional strategies, both to strengthen the local base in Jambi and to expand recruitment potential from outside the province.

4. Conclusions

Conclusion

Based on the clustering results, the six clusters formed exhibit consistent grouping patterns based on a combination of gender and age stage among the santri (students). Clusters 1 and 2 are dominated by male santri in early adolescence (14–16 years old), with their regional origins concentrated in several regencies/cities within Jambi Province, particularly Kota Jambi, Sarolangun, and Tanjung Jabung Barat. This homogeneous and geographically concentrated characteristic suggests that the young male santri segment has, so far, largely originated from local areas surrounding the pesantren (Islamic boarding school) environment.

Cluster 3 displays a similar pattern among young female santri (15–16 years old), with their regional origins likewise concentrated within Jambi Province, though covering a slightly more diverse number of regencies/cities compared to Clusters 1 and 2. Meanwhile, Clusters 5 and 6 represent late-adolescent santri (17–19 years old), both male and female, who remain predominantly from within Jambi Province, with a relatively small proportion originating from outside the province (below 15%).

In contrast to the other clusters, Cluster 4 exhibits a markedly different pattern: despite having the smallest number of members, the majority of female santri in this cluster (more than 95%) originate from outside Jambi Province, with a highly diverse spread of regional origins. This finding suggests the presence of a female santri segment with a different recruitment pattern from the other groups, possibly stemming from different channels or sources of information compared to the local segment.

Overall, these clustering results illustrate differing geographic and demographic distribution patterns across santri segments, with most clusters showing a concentration of origin within Jambi Province, while one cluster demonstrates a broader reach extending beyond the province. These patterns may serve as valuable considerations for pesantren administrators in understanding the characteristics of the current santri segments, as well as a basis for further analysis in designing promotional approaches better suited to the characteristics of each segment going forward.

5. Acknowledgements

The author would like to thank Mrs. Sukma Puspitorini, S.T., M. Kom. for the guidance and direction given, as well as to the Al-Hidayah Islamic Boarding School for the opportunity and data provided during the research.

6. Funding

This research does not have financial support or grants received from external parties, either from campus institutions or from research agencies (Pondok Pesantren Karya Pembangunan Al-Hidayah).

7. Declaration of Competing Interest

The authors state that there are no competing and known financial interests or personal relationships, which may appear to have influenced the results of the work reported in this scientific article.

8. References

[1] S. Yang, P. Pondok, and P. Al-Hidayah, "HALAMAN PERSETUJUAN PEMBIMBING."

[2] A. Sani Sekolah Tinggi Manajemen Informatika dan Komputer Widuri Jl Palmerah Barat No, G. Utara, K. Lama, and J. Selatan, "PENERAPAN METODE K-MEANS CLUSTERING PADA PERUSAHAAN."

[3] A. Kurniawan and A. Zein, "Jitu: Jurnal Informatika Utama Hal," pp. 81–96, doi: 10.55903/jitu.v2i2.xx.

[4] J. Manajemen *et al.*, "Pengaruh Data Mining, Strategi Perusahaan Terhadap Laporan Kinerja Perusahaan," vol. 3, no. 1, 2024.

[5] T. Amalina, D. Bima, A. Pramana, and B. N. Sari, "Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food," *Jurnal Ilmiah Wahana Pendidikan*, vol. 8, no. 15, pp. 574–583, 2022, doi: 10.5281/zenodo.7052276.

[6] N. M. Fithriyani, M. Martanto, A. R. Dikananda, and D. Rohman, "ALGORITMA K-MEANS UNTUK MENINGKATKAN SEGMENTASI POLA KEKERASAN," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, Jan. 2025, doi: 10.23960/jitet.v13i1.5795.

[7] N. Nur Afidah, "Penerapan Metode Clustering dengan Algoritma K-means untuk Pengelompokkan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang," *PRISMA, Prosiding Seminar Nasional Matematika*, vol. 6, pp. 729–738, 2023, [Online]. Available: https://journal.unnes.ac.id/sju/index.php/prisma/

[8] Wenny, "Bulletin of Information System Research (BIOS) Normalisasi Data Kependudukan Dengan Model Min Max Dan Algoritma K-Means Untuk Pengelompokkan Tingkat Ekonomi Masyarakat," 2024, [Online]. Available: https://journal.grahamitra.id/index.php/bios

[9] D. Amelia, T. N. Padilah, and A. Jamaludin, "Optimasi Algoritma K-Means Menggunakan Metode Elbow dalam Pengelompokan Penyakit Demam Berdarah Dengue (DBD) di Jawa Barat," *Jurnal Ilmiah Wahana Pendidikan*, vol. 8, no. 11, pp. 207–215, 2022, doi: 10.5281/zenodo.6831380.

[10] A. Rajsya and A. Rachman Manga, "Rancang Bangun Penerapan Metode Elbow Pada K-Means Untuk Clustering Data Persediaan Barang," *Literatur Informatika & Komputer*, vol. 1, no. 4, pp. 395–403, 2024, doi: 10.33096/linier.v1i4.2539.