

## Clustering Villages in the Mountain Areas in West Java Based on Tourism Potential Using K-Prototype Algorithm

Ainun Salsabila <sup>(1)</sup>, L. M. Risman Dwi Jumansyah <sup>(2)</sup>, Anwar Fitrianto <sup>(3)</sup>, Erfiani <sup>(4)</sup>,  
Alfa Nugraha P. <sup>(5)</sup>

<sup>1,2,3,4,5</sup>Department of Statistics and Data Science, FMIPA, IPB University

Meranti Street, IPB University Dramaga, Bogor 16680, West Javat

e-mail: [ainunsalsabila@apps.ipb.ac.id](mailto:ainunsalsabila@apps.ipb.ac.id)<sup>(1)</sup>, [rismandwijumansyah@apps.ipb.ac.id](mailto:rismandwijumansyah@apps.ipb.ac.id)<sup>(2)</sup>,  
[anwarstat@gmail.com](mailto:anwarstat@gmail.com)<sup>(3)</sup>, [erfiani@apps.ipb.ac.id](mailto:erfiani@apps.ipb.ac.id)<sup>(4)</sup>, [alfanugraha@apps.ipb.ac.id](mailto:alfanugraha@apps.ipb.ac.id)<sup>(5)</sup>

### ABSTRAK

Analisis klaster adalah salah satu metode analisis multivariat yang digunakan untuk mengelompokkan suatu objek berdasarkan kesamaan karakteristiknya. Secara umum, data yang digunakan pada proses klasterisasi adalah data numerik atau kategorik saja. Tetapi, terkadang juga banyak ditemui kasus yang menggunakan data campuran numerik dan kategorik. Oleh karena itu, algoritma yang dapat digunakan adalah *K-Prototype*. *K-Prototype* adalah pengembangan dari *K-Means* yang dapat digunakan pada data berjumlah besar bertipe numerik dan kategorik. Dasar pengembangan *K-Prototype* adalah mengukur jarak antara objek dan pusat *prototype*. Banyaknya *prototype* tergantung dengan jumlah *cluster* yang dibentuk. Pada penelitian ini, peneliti memanfaatkan algoritma tersebut untuk mengklasterkan desa-desa wilayah pegunungan di Jawa Barat berdasarkan potensi wisatanya, agar mengetahui potensi-potensi yang perlu dikembangkan berdasarkan lima komponen, yaitu Daya Tarik, Keterjangkauan, Akomodasi, Fasilitas Pendukung, dan Kesadaran. Berdasarkan *Silhouette* dan *McClain Index*, diperoleh bahwa jumlah klaster optimal adalah dua. Klaster 1 terdiri dari 103 desa dan klaster 2 terdiri dari 703 desa. Klaster 1 adalah desa-desa yang secara umum sudah lebih unggul pada komponen Keterjangkauan, Kesadaran, dan Fasilitas Pendukung, tetapi masih kurang pada komponen Daya Tarik dan Akomodasi dibandingkan desa-desa di klaster 2.

**Kata kunci:** Analisis Klaster; *K-Prototype*; Potensi Wisata

### ABSTRACT

*Cluster analysis is a multivariate analysis method used to group objects based on their similar characteristics. In general, in the clustering process only use numerical or categorical data. But, sometimes we also encounter cases that use both numerical and categorical data. Therefore, the algorithm that can be used is K-Prototype. K-Prototype is a development of K-Means that can be used on large data with numerical and categorical types. The basis of K-Prototype development is to measure the distance between the object and its centroid prototype. The number of prototypes depends on the number of clusters formed. In this study, researchers use this algorithm to group mountainous villages in West Java based on their tourism potential, in order to find out the potential that needs to be developed based on five components, namely Attractions, Access, Accommodation, Amenities and Awareness. Based on the Silhouette and McClain Index, the optimal number of clusters is two. Cluster 1 consists of 103 villages and cluster 2 consists of 703 villages. Cluster 1 are villages that are generally better in the Access, Awareness and Amenities, but are still lacking in the Attraction and Accommodation components compared to villages in cluster 2.*

**Keywords:** Cluster Analysis; *K-Prototype*; Tourism Potential

Ainun Salsabila<sup>1</sup>, L. M. Risman Dwi Jumansyah<sup>2</sup>, Anwar Fitrianto<sup>3</sup>,  
Erfiani<sup>4</sup>, Alfa Nugraha P.<sup>5</sup>/

## INTRODUCTION

Cluster analysis is grouping objects who have the same characteristics in one cluster [1, pp. 342–353]. Cluster analysis requires a distance measure that is defined for each pair of objects to be grouped. Smaller the distance, the more similar one object is to another object. The distances commonly used are distance for numerical and categorical data [2]. So, if the data is a combination of numerical and categorical, the method that is used is a method that combines distances for numerical and categorical data, such as K-Prototype. K-Prototype is a development of K-Means, introduced by Huang that can be used on large data with numerical and categorical types. The basis of K-Prototype development is to measure the distance between the object and its centroid prototype [3]. Cluster analysis and K-Prototype can be used to cluster areas like villages. So, in this research, researcher interest to cluster villages in the mountain areas in West Java based on their tourism potential which consists of five components, Attraction, Access, Accommodation, Amenities, and Awareness, in order to help the parties concerned to find out the potential, because research regarding exploring village tourism potential in mountain areas is still very difficult to find. Researcher also interest to proposed the use of K-Prototype because the data that is used in this research are mix numerical and categorical.

Research using cluster analysis was carried out by [4] with the title “Pengembangan Desa Wisata dengan Pendekatan Klaster” which applies components for developing the potential of tourist villages consisting of *Attraction, Accessibility, Amenity, Ancillary Services, Activity, and Accommodation*, which results in 6 clusters to arrange strategies to develop tourism villages. Meanwhile, the application of K-Prototype algorithm for grouping villages was carried out by [5] in his research entitled “Penerapan Algoritma *K-Prototypes* untuk Pengelompokan Desa-Desa di Provinsi Jawa Barat Berdasarkan Indikator Indeks Desa Membangun Tahun 2020”. This research produced 5 clusters with the lowest development achievement criteria for cluster 1 to the best development achievement for cluster 5.

One of the ways in which cluster analysis can be used to cluster villages is to help explore the village’s potential. One of the interesting potentials to be explored is tourism potential, especially tourism potential mountain villages. Tourism potential needs to be explored, because it is one of the sectors driving growth and improvement of the national economy. Tourism is a combination of activities, services, and industries that provide travel experiences including culinary, accommodation, transportation, entertainment, activity facilities, and hospitality services for individuals or groups traveling far from their home [6]. According to Badan Pusat Statistik, the tourism sector in 2021 gave a contribution to Gross Domestic Product (GDP) of 4.97% and increase every year, with state foreign exchange earnings of US\$ 15 billion per year [7]. Currently, the trend in tourism development in Indonesia is a tourism development model based on tourist villages. This happened because the interest of tourists began to shift who wanted entertainment with beautiful views, far from the city, and cheaper costs. Apart from that, the phenomenon of the increasing number of development of tourist villages is in line with the enactment of the Law on Villages which gives villages development authority to each village government and provides multiple villages funds [8]. A tourist villages itself is a rural area that offers an overall atmosphere that reflects the authenticity of the village, both in terms of social-economic life, social-culture, daily life, unique of village structures, interesting

economic activities, and has the tourism potential to develop, such as attractions, accommodation, food and drinks, and other tourist facilities [9].

Exploring the tourism potential of mountain villages is interesting because mountain villages have very different characteristics from valley and mainland villages. Mountain villages have limited basic facilities and employment opportunities. The geographical location of mountain villages is one of the reasons why the social and economic conditions of mountain village residents are lower. Thus, additional efforts are needed to solve existing limitations, so the economic disparities can be solved [10]. One of the efforts that can be made is to develop tourist villages in mountain areas, because not only beautiful mountain views can be offered, but the use of existing resources around the mountain such as forests, rivers or lakes as well as the sale of superior products can also be used as an additional attraction for tourists. Tourism potential needs to be developed considering that the West Java area consists of many mountains with interesting resources.

## METHOD

This research uses secondary data from Badan Pusat Statistik, namely PODES (Potensi Desa) data for 2021. The data used only data for villages in mountain areas (top and valley) where there are settlements in West Java Province. The amount of data is used consist of 806 observations with 16 variables which are divided into five components as in the Table 1 below:

**Table 1.** Description of Research Variables

| Component  | Variable | Description  | Explanation  | Scale       |
|------------|----------|--|--|-------------|
| Attraction | $X_1$    | Location of village in forest areas                                    | 1: In the forest area<br>2: Around the forest area<br>3: Outside the forest area | Categorical |
|            | $X_2$    | Use of rivers and lakes/reservoirs/dams for tourism                    | 1: River<br>2: Lake/reservoir/dam<br>3: River and lake<br>4: Neither             | Categorical |
|            | $X_3$    | Main product:<br>a. Main products                                      | 1: Food<br>2: Non Food<br>3: Food and Non Food<br>4: Neither                     | Categorical |
|            | $X_4$    | b. Main products exported to other countries                           | 1: Yes, mostly<br>2: Yes, a small part<br>3: None                                | Categorical |
| Amenities  | $X_5$    | Tourist attractions/public baths                                       | 1: Exist<br>2: None  | Categorical |
|            | $X_6$    | Total number of places to eat:<br>a. Restaurant<br>b. Stall/food stall | -  | Numeric     |
|            | $X_7$    | Number of shopping places (mini market/ supermarket)                   | -  | Numeric     |

|                      |          |  |   |             |
|----------------------|----------|--|---|-------------|
|                      | $X_8$    | Number of shops/ grocery stalls  | -   | Numeric     |
|                      | $X_9$    | Total number of health facilities:<br>a. Hospital<br>b. Community health center with inpatient care<br>c. Polyclinic/treatment center<br>d. Doctor's practice place<br>e. Village Health Post<br>f. Pharmacy | -   | Numeric     |
|                      | $X_{10}$ | Total number of worship place:<br>a. Mosque<br>b. Musholla   | -   | Numeric     |
|                      | $X_{11}$ | The number of police posts (including police stations) used  | -   | Numeric     |
|                      | $X_{12}$ | Total number of operating financial institutions:<br>a. Government Commercial Banks (BRI, BNI, Mandiri, BPD, BTN)<br>b. Private Commercial Banks (BCA, Permata, Sinarmas, CIMB, etc)<br>c. Rural Bank (BPR)  | -   | Numeric     |
| <i>Access</i>        | $X_{13}$ | Land roads between villages can be passed by motorized vehicles with 4 or more wheels  | 1: Throughout the year<br>2: Throughout the year, except certain moment (when rain, etc)<br>3: During dry season<br>4: Can't passed throughout the year | Categorical |
|                      | $X_{14}$ | The existence of public transportation   | 1: Exist, with fixed route<br>2: Exist, with unfixed route<br>3: None   | Categorical |
| <i>Accommodation</i> | $X_{15}$ | Total number of hotels and homestays   | -   | Numeric     |
| <i>Awareness</i>     | $X_{16}$ | Build, maintenance or normalization of rivers, canals, embankments, ditches, drainage, reservoirs, etc   | 1: Yes<br>2: No   | Categorical |

The stages of data analysis in this research are as follows:

1. Describe the research data in general such as the number of villages according to the topography of region in percentage.
2. Determine the optimal number of clusters using the Silhouette and McClain Index using equations: Silhouette Index is a combination of cohesion (average distance from object  $i$  with all objects that are in one cluster, denoted  $a_i$ ) and separation (average distance from object  $i$  with all objects that are in other cluster, denoted  $b_i$ ) [11]:

$$S_i = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (1)$$

McClain Index is ratio of variety of within cluster distance ( $S_w$ ) average with variety of between cluster distance ( $S_b$ ) average [12]:

$$McClain = \frac{\bar{S}_w}{\bar{S}_b} \quad (2)$$

3. Initialization of the prototype

Random selection of  $k$  prototypes from the data according to the specified number of clusters. The limit of the number of clusters  $k$  to be formed is a minimum of 2 and a maximum of  $n$  or  $n/2$ , where  $n$  is the number of samples [13].

4. Measure the distance of all objects to all prototypes

$$d(X, Y) = \sum_{j=1}^p (x_{jn} - y_{jn})^2 + \gamma \sum_{j=p+1}^m d(x_{jc}, y_{jc}) \quad (3)$$

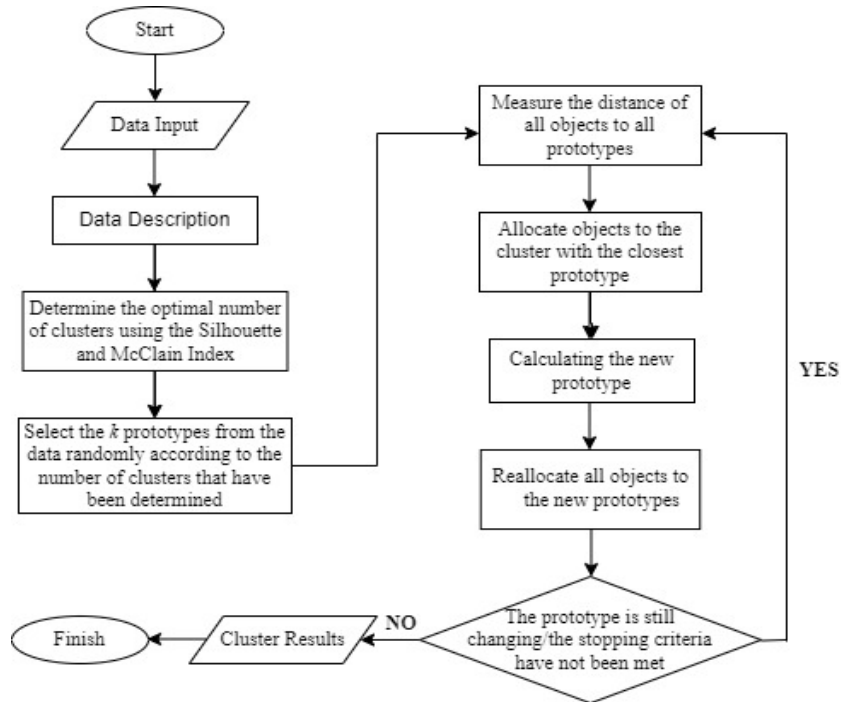
$$d(x_{jc}, y_{jc}) = \begin{cases} 0, & x_{jc} = y_{jc} \\ 1, & x_{jc} \neq y_{jc} \end{cases} \quad (4)$$

5. Allocate objects to the cluster with the closest prototype.
6. Calculating the new prototype.
7. Reallocate all objects to the new prototype

After all objects have been allocated, the distance between all objects and all existing prototypes will be re-measured. If there is an object that turns out to be closer to another prototype, membership will be transferred and then updates will be made to the old and new cluster prototypes.

8. If the prototype still changes significantly or the stopping criteria have not been met, then return to step 4 until no object movement or the maximum iteration is reached.

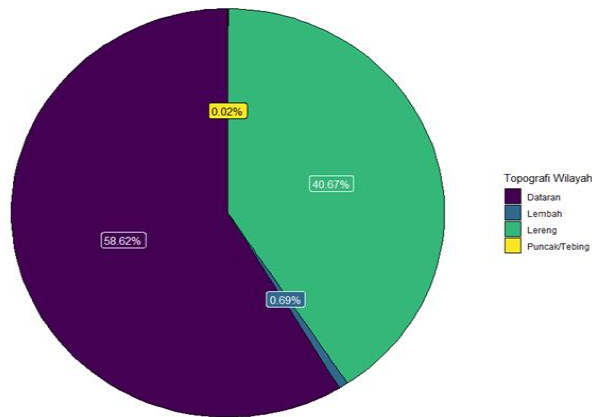
The analysis stage can be described with a flowchart as in Figure 1 below:



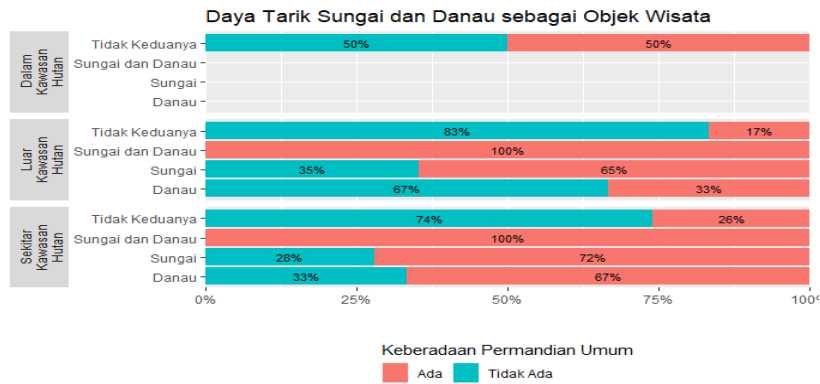
**Figure 1.** Flowchart of Data Analysis

**RESULT AND DISCUSSION**

West Java is a province that offers a very interesting topography. More than half, approximately 58.62% of the province's territory is land that is often the center of settlement and agricultural activity. Furthermore, about 40.67% of the villages are located downhill. The villages in the region have natural beauty and attractive resources that can be used as tourist destinations to support the economy of the villagers. But, to realize that, there is a lot of tourism potential that needs to be explored and developed.



(a)



(b)

Figure 2. (a) Topographic Description of Village in West Java; (b) Description of Rivers and Lakes as Tourist Attractions

Tourist potential components such as attractions, forest areas, and the presence of rivers and lakes are the main magnet for tourists both from the local area and from outside the area, as when the river and lake are used as public bathing places. As in the study conducted by [14], the Lau Seruwai River in North Sumatra is used as a place for bathing or swimming by tourists. For west Java (see Figure 2(b)), about 65% of the rivers outside the forest area are used as public bathing facilities. Interestingly, some villages around the Forest Area have rivers and lakes that are entirely or 100% allocated as public bathing facilities. Based on this, it can be known that the people or tourists in the village are already aware of the potential of rivers and lakes if used as tourist attractions.

Another potential that can be a major attraction for tourists is the products produced by local villagers, whether food or non-food products, which have been successfully exported to various countries. A description of this in the village around downhill in West Java can be seen in Figure 3. The forested territory, about 50% of the products produced, both food and non-food, have expanded the international market. In the villages outside and around the forest, the food and non-food products exported are still very small. Based on this, the villages in the outer territory and around the forest area should be more keen in marketing their superior products, to be a tourist attraction and attract tourists to come.

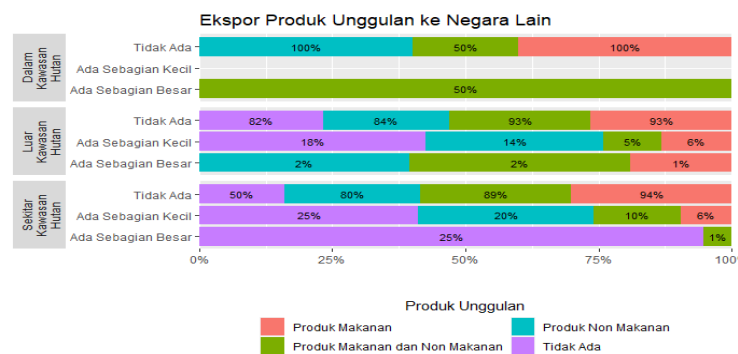


Figure 3. Description of Exported Villages Main Product

The clustering of villages begins by searching the optimal number of clusters because the K-Prototype algorithm is one of the non-hierarchical grouping methods, in which the cluster number must be determined at an early stage. The selection of the optimal number of clusters is done using the method of Silhouette and McClain coefficients.

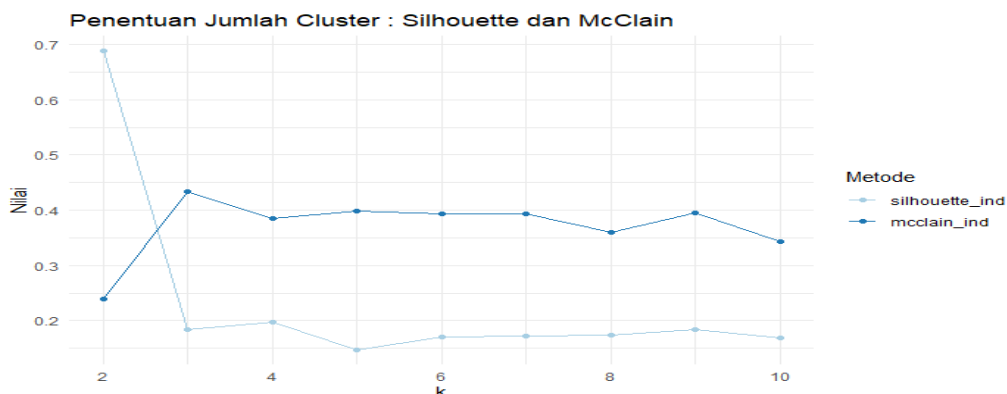


Figure 4. Silhouette and McClain Coefficient Graph

Figure 4 shows a graph of the values of the Silhouette and McClain coefficients for each number of clusters. From the graph, it can be concluded that the optimum number of Clusters is 2 because it has the highest value of Silhouettes and the lowest of McClain. The range of selection of the number of clusters is limited from 2 to 10 so that the interpretation of the characteristics of the cluster is easier.

The optimum number of clusters obtained was then used to cluster with the K-Prototype algorithm. The result of a cluster using a K-prototype with a value of  $k = 2$  mapped as many as 103 villages into cluster 1 and 703 villages in cluster 2. The district of Bandung has the most members in cluster 1 of 25 villages and the city of Tasikmalaya has no villages gathered in cluster 1. Whereas in the second cluster, Garut district has the highest number of members, 106 villages, and the town of Tasikmalaya has the lowest membership of 3 villages. But in general, mountain villages in West Java are much more in Cluster 2 than Cluster 1. The cluster results obtained are as follows:

Table 2. Cluster Results

| Region        | Number of Cluster Members |     | Region           | Number of Cluster Members |    |
|---------------|---------------------------|-----|------------------|---------------------------|----|
|               | 1                         | 2   |                  | 1                         | 2  |
| Bandung       | 25                        | 40  | Kota Tasikmalaya | 0                         | 3  |
| Bandung Barat | 7                         | 20  | Kuningan         | 1                         | 58 |
| Bogor         | 11                        | 30  | Majalengka       | 4                         | 31 |
| Ciamis        | 4                         | 86  | Pangandaran      | 0                         | 27 |
| Cianjur       | 12                        | 74  | Purwakarta       | 2                         | 29 |
| Cirebon       | 1                         | 21  | Subang           | 1                         | 24 |
| Garut         | 7                         | 106 | Sumedang         | 12                        | 75 |
| Kota Bandung  | 7                         | 5   | Tasikmalaya      | 5                         | 67 |
| Kota Bogor    | 4                         | 7   |                  |                           |    |

The characteristics of each cluster based on categorical variables are as follows:

**Table 3.** Characteristics of Cluster Results for Categorical Variable

| Component       | Variable                                     | Category                | Percentage (%) |      |
|-----------------|--|-------------------------|----------------|------|
|                 |  |                         | 1              | 2    |
| Attraction      | X <sub>1</sub>                               | In the Forest Area      | 0              | 0.6  |
|                 |  | Around the Forest Area  | 71.8           | 68.8 |
|                 |  | Outside the Forest Area | 28.2           | 30.6 |
|                 | X <sub>2</sub>                               | Lake                    | 1.0            | 1.6  |
|                 |  | River                   | 4.9            | 7.7  |
|                 |  | River and Lake          | 1.9            | 0.6  |
|                 | X <sub>3</sub>                               | Neither                 | 92.2           | 90.2 |
|                 |  | Food Product            | 47.6           | 55.3 |
|                 |  | Non Food Product        | 20.4           | 30.6 |
|                 |  | Both                    | 30.1           | 12.2 |
|                 |  | Neither                 | 1.9            | 1.8  |
|                 | X <sub>4</sub>                               | Yes, Mostly             | 1.9            | 1.4  |
|                 |  | Yes, a Small Part       | 11.7           | 6.7  |
|                 |  | None                    | 86.4           | 91.9 |
|                 | X <sub>5</sub>                               | Exist                   | 20.4           | 24.8 |
|                 |  | None                    | 79.6           | 75.2 |
| X <sub>13</sub> | During the Dry Season                        | 0                       | 0.1            |      |
|                 | Throughout Year                              | 98.1                    | 97.3           |      |
|                 | Throughout Year<br>(Depending on Conditions) | 1.9                     | 2.6            |      |
|                 | Exist (Fixed Route)                          | 60.2                    | 51.2           |      |
| Access          | X <sub>14</sub>                              | Exist (Unfixed Route)   | 37.9           | 46.4 |
|                 |  | None                    | 1.9            | 2.4  |
|                 | Neither                                      | 92.2                    | 90.2           |      |
| Awareness       | X <sub>16</sub>                              | Yes                     | 55.3           | 43.5 |
|                 |  | No                      | 44.7           | 56.5 |

Meanwhile, the characteristics of each cluster based on numerical variables are as follows:

**Table 4.** Characteristics of Cluster Results for Numerical Variable

| Component | Variable        | Minimum |   | Average |       | Maximum |     |
|-----------|-----------------|---------|---|---------|-------|---------|-----|
|           |                 | 1       | 2 | 1       | 2     | 1       | 2   |
| Amenities | X <sub>6</sub>  | 0       | 0 | 37.05   | 14.43 | 211     | 177 |
|           | X <sub>7</sub>  | 0       | 0 | 1.62    | 0.47  | 12      | 10  |
|           | X <sub>8</sub>  | 105     | 0 | 181.26  | 41.82 | 411     | 117 |
|           | X <sub>9</sub>  | 1       | 0 | 6.46    | 3.44  | 24      | 24  |
|           | X <sub>10</sub> | 11      | 1 | 46.66   | 32.85 | 127     | 121 |

|               |          |   |   |      |      |    |     |
|---------------|----------|---|---|------|------|----|-----|
|               | $X_{11}$ | 0 | 0 | 0.12 | 0.07 | 3  | 2   |
|               | $X_{12}$ | 0 | 0 | 0.73 | 0.21 | 12 | 17  |
| Accommodation | $X_{15}$ | 0 | 0 | 1.09 | 1.31 | 30 | 105 |

So, the characteristics based on Tables 3 and 4 can be summarized as follows:

**Table 5.** Summary of Cluster Characteristics

| Cluster | Characteristics  |
|---------|--|
| 1       | The location of the village tends to be far away/outside the forest, have easy access and have lots of public transportation. Has more adequate dining facilities, shopping places, health facilities, places of worship and police posts. However, the use of rivers and lakes as tourist attractions is still lacking and the number of accommodations is a little, even though the population is more aware of cleanliness and environmental maintenance, as well as exporting main products, even though the numbers are still little. |
| 2       | The location of the village is around a forest area, some even inside the forest, so many rivers and lakes are used as tourist attractions such as public baths, and have many main products, both food and non-food, but not many are exported. Access to roads and public transportation, awareness of the environment, and facilities such as places to eat, shopping, health, places of worship, and police posts are still lacking, but there are more accommodations.  |

## CONCLUSION

Mountain villages in West Java are divided into two clusters, namely cluster 1 comprises 103 villages and Cluster 2 comprises 703 villages. Cluster 1 are villages that are generally superior in the components of Accessibility, Awareness, and Supporting Facilities, but still less in the elements of Attraction and Accommodation compared to the villages of Cluster 2. Therefore, villages within Cluster 1 are given more attention and dig into the existing tourist attractions, such as using rivers and lakes as their main tourist attraction. In addition, the village in Cluster 1, must also create more excellent products to attract tourists to come. Whereas for villages inside cluster 2, it is preferable to be the focus of government attention in improving and improving the accessibility and infrastructure of villages for tourist convenience. So, when these potential tourist components can be improved, it is expected to make the village an attractive tourist village that can bring many visitors, to help and boost the economy of the village people.

## REFERENCE

- [1] A. Y. Kartini and A. M. Jamiluddin, "Application of Agglomerative Hierarchical Clustering Method for Grouping Non-Cash Food Assistance Recipients in Ngambon Bojonegoro," 2023. Accessed: Dec. 10, 2023. [Online]. Available: [www.unipasby.ac.id](http://www.unipasby.ac.id)
- [2] R. Wijayati and D. R. S. Saputro, "Clustering Data Numerik dan Kategorik Menggunakan Algoritme K-Prototype. PRISMA, Prosiding Seminar Nasional Matematika 6," pp. 702–706, 2023, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>

- [3] S. Annas, I. Irwan, R. H. Safei, and Z. Rais, “K-Prototypes Algorithm for Clustering The Tectonic Earthquake in Sulawesi Island,” *Jurnal Varian*, vol. 5, no. 2, pp. 191–198, May 2022, doi: 10.30812/varian.v5i2.1908.
- [4] I. Kurniawan and L. Fitriani, “PENGEMBANGAN DESA WISATA DENGAN PENDEKATAN KLASTER,” *Indonesian Journal of Economics, Entrepreneurship and Innovation*, vol. 2, no. 2, pp. 52–59, 2021, doi: 10.31960/ijoei.v2i2.1477.
- [5] M. Ganmanah, A. Kudus, P. Statistika, F. Matematika, D. Ilmu, and P. Alam, “Penerapan Algoritme K-Prototypes untuk Pengelompokan Desa-Desa di Provinsi Jawa Barat Berdasarkan Indikator Indeks Desa Membangun Tahun 2020”, doi: 10.29313/v0i0.28974.
- [6] L. A. Birahmatika and E. Ahyudanari, “Jurnal Aplikasi Teknik Sipil Analisis Clustering Objek Daya Tarik Wisata Kabupaten Banyuwangi.”
- [7] K. Putri and F. Sari, “PENGARUH SEKTOR PARIWISATA TERHADAP PERTUMBUHAN EKONOMI DI KABUPATEN LOMBOK TENGAH DAN KABUPATEN MALANG Oleh,” *Jurnal Ekonomi dan Bisnis*, vol. 11, no. 2, 2022.
- [8] A. Y. Asmoro, T. B. Bachri, and A. Detmuliati, “ANALISIS POTENSI WISATA DESA DENGAN KERANGKA 6A STUDI KASUS DESA NGAJUM, MALANG,” vol. 18, no. 2, 2020, doi: 10.36275/mws.
- [9] D. Olivia, W. Tisno Atmojo, A. Guna, P. Studi, and P. Wilayah Dan Kota, “Analisis Potensi Desa Wisata Sebagai Upaya Peningkatan Aksesibilitas Dan Konektivitas Di Desa Wisata Cikolelet.” [Online]. Available: <https://journals.upi-yai.ac.id/index.php/ikraith-teknologi/issue/archive>
- [10] A. Y. Saputra, E. Rustiadi, and W. Rindayati, “Perkembangan dan Karakterisasi Desa-desa Pegunungan Jawa Tengah,” *Journal of Regional and Rural Development Planning*, vol. 6, no. 1, pp. 1–13, Feb. 2022, doi: 10.29244/jp2wd.2022.6.1.1-13.
- [11] L. Qadrini *et al.*, “Metode K-Means dan DBSCAN pada Pengelompokan Data Dasar Kompetensi Laboratorium ITS Tahun 2017,” 2020. [Online]. Available: [www.unipasby.ac.id](http://www.unipasby.ac.id)
- [12] R. Aschenbruck and G. Szepannek, “Cluster Validation for Mixed-Type Data”, doi: 10.5445/KSP/1000098011/02.
- [13] R. Nooraeni, J. Suprijadi, and P. Statistika STIS, “K-PROTOTYPE UNTUK PENGELOMPOKAN DATA CAMPURAN,” *Jurnal Statistika Teori dan Aplikasi: Biomedics, Industry & Business And Social Statistics*, vol. 13, no. 1.
- [14] J. S. P. Pakpahan, O. H. Syahputra, and B. Slamet, “Jurnal Persepsi Masyarakat,” *Jurnal Penelitian Kehutanan Bonita*, vol. 4, no. 1, pp. 20–29, 2022.