Cluster Identification of Agrotourism in Banten Province Indonesia

Suherna Suherna[1], *Weksi* Budiaji[1,*] and *Nurmayulis* Nurmayulis[2]

[1] Agribusiness Department, University of Sultan Ageng Tirtayasa, 42163 Serang, Banten, Indonesia
[2] Agroecotechnology Department, University of Sultan Ageng Tirtayasa, 42163 Serang, Banten, Indonesia

Abstract. The post-pandemic era affects the tourism sector. Before the pandemic, agrotourism was growing in Banten Province in line with the development of special economic zones. To regain the competitiveness of agrotourism in Banten Province, characterization among regions is crucial. This study aims to cluster the regions in Banten Province into groups to identify the agrotourism potency. This study employed a descriptive quantitative study by applying k-medoid clustering in the secondary data of Banten in Figures and eliciting information from the Tourism Department of Banten Province. The number of agrotourism sites, the population number, and the number of restaurants were included in the clustering process. The clustering analysis that results in three clusters was validated and visualized via relative and internal criteria. Cluster 1 was the least potential cluster of agrotourism with members of two municipals Tangerang and Tangerang Selatan. Cluster 2 was the most potential agrotourism region which has members of three regions of Pandeglang, Lebak, and Serang and two municipals of Serang and Cilegon. It is characterized by the highest number of agrotourism sites. Cluster 3 was the second potential agrotourism cluster where the regency of Tangerang was the only member possessing the highest population as a potential market.

1 Introduction

The year 2020 is well-known as the year of covid-19 pandemic. As a pandemic, the disease spread globally affecting many sectors worldwide. The most damaged suffering hard decline was the leisure and tourism sector where there was a 50% loss of revenue [1]. Indonesia suffered an unprecedented decline of inbound tourism as well. Both non-tourism and tourism-dependent regions in Indonesia were affected by the growth reduction and poverty increment significantly [2].

As a non-tourism-dependent region for its regional revenue, Banten Province was selected to be one of the Special Economic Zones (SEZ) in 2012 for the Tanjung Lesung site. Although it still requires support and improvement, it increases the mobilization of people, goods, and services [3]. The natural resources and environmental attractions also provide new business opportunities [4]. Thus, agrotourism has a high potential as the Department of

* Corresponding author: budiaji@untirta.ac.id

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

Tourism of Banten Province [5] mentioned that 80% of the land use in Banten Province is agriculture.

Although Banten Province has potency in agrotourism sites to be visited, it is contradictory to the pandemic restrictions which limited people's mobility. Then, the post-pandemic actions to attract visitors to this business are crucial. The government also has an important role in recovery strategies [6]. Clustering regions in Banten Province to characterize every region can support the management of agrotourism destinations.

Clustering has been conducted for regions [7-9] and tourist destination sites [10-11]. The efficient indicators, important attributes, inefficient behaviour, and popular destinations can be obtained by this analysis so that cooperation among regions can increase tourism competitiveness. Thus, the objective of this study was to identify regions based on the agrotourism potency in Banten Province by cluster analysis. The clustering analysis can identify the emerging clusters and clusters that require support to improve agrotourism competitiveness.

2 Method

This research was a descriptive quantitative study. The data were secondary data collected from the Tourism Department of Banten and Banten Province in Figures 2022 [12]. There were three important variables in this study, namely the number of agrotourism sites, the number of populations, and the number of restaurants. The number of agrotourism objects could change over time [13-14] so that the newest data set from the Tourism Department in 2023 was indispensable. The number of populations was also required because the local population could be competitive with agrotourism [15]. Moreover, the number of restaurants was included due to its largest contribution to the regional revenue [16].

The regions in Banten Province consisting of four municipals and four regencies were grouped via cluster analysis based on the three variables. The regions clustering [7-9] was intended to identify potential regions w.r.t agrotourism. A distance-based clustering was applied in the cluster analysis where the k-medoids method was opted for due to its distance variation [17]. Three distance variations were applied, namely Euclidean, squared Euclidean, and Manhattan distances.

Then, clustering results were validated via relative and internal criteria. The relative criterium applied a consensus matrix [18], while the internal criterium implemented silhouette width [19] and medoid-based shadow value [20-21]. The results from both criteria were visualized via bar-plot and spider-plot. These plots assisted the description of the regions in Banten Province based on the important variables of agrotourism.

3 Result and Discussion

Banten Province consists of four regencies and four municipals. The municipals have smaller areas than the regencies. The density of the municipals is high where it is more than two thousand people per km^2 (Table 1). The municipality of Tangerang has the highest population density. Meanwhile, the region of Tangerang and the region of Pandeglang have the highest number of restaurants and agrotourism sites, respectively.

To group all of the regions, a medoid-based cluster analysis, namely k-medoid, is applied via Euclidean, squared Euclidean, and Manhattan distances. The heatmaps of the consensus matrix of relative criterium validate the clustering result to three clusters only ($k = 3$) due to the small number of objects, i.e., eight regions. Figure 1 shows the heatmap of two and three clusters of all distances. A heatmap with clear-square diagonals and clear-white off-diagonals is the most stable result. Figure 1 depicts the two clusters ($k = 2$) as unstable clusters because the right-bottom square diagonals are blurred and can be two squares. The three clusters (k

= 3) have clear-white off-diagonal and the squared Euclidean has the most stable cluster due to its clear-square diagonal.

Table 1. Summary data of regions of Banten Province

Region	Area (km2)	Population	Density	Number of Restaurants	Number of Agrotourism Sites
R. Pandeglang	2,746.89	1,288,314	469.008	116	10
R. Lebak	3,426.56	1,407,857	410.866	56	4
R. Tangerang	1,011.86	3,293,533	3,254.930	907	4
R. Serang	1,734.28	1,647,790	950.129	16	12
M. Tangerang	153.93	1,911,914	12,420.672	426	1
M. Serang	175.5	441,761	2,517.157	2	2
M. Cilegon	266.71	704,618	2,641.888	46	3
M. Tangerang Selatan	147.19	1,365,688	9,278.402	477	4

* R = Regency, M = Municipal
Source: [12] and eliciting the Tourism Department of Banten Province

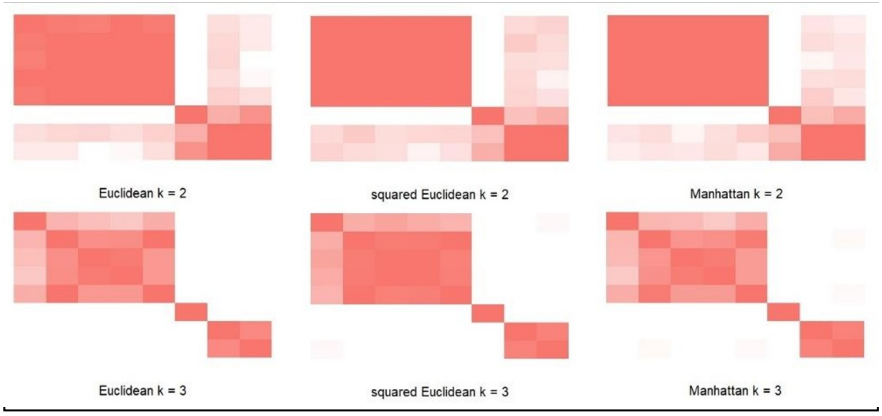

Fig. 1. Heatmaps of Euclidean, squared Euclidean, and Manhattan distances of k-medoids (k = 2, 3)

The other validations are silhouette width and medoid-based shadow value to quantify the separation and compactness among clusters. Table 1 and 2 shows silhouette width and medoid-based shadow values, respectively. The values indicate that the higher the value, the better the separation among clusters. Thus, the squared Euclidean with three clusters has the best separation in both silhouette width and medoid-based shadow values criteria.

Table 2. Silhouette width of all distances (k = 2, 3, 4, and 5)

	K = 2	**K = 3**	**K = 4**	**K = 5**
Euclidean	0.71	0.88	0.75	0.77
Squared Euclidean	0.79	**0.98**	0.88	0.89
Manhattan	0.70	0.86	0.72	0.70

Table 3. Medoid-shadow value of all distances (k = 2, 3, 4, and 5)

	K = 2	**K = 3**	**K = 4**	**K = 5**
Euclidean	0.87	0.93	0.87	0.89
Squared Euclidean	0.96	**0.99**	0.96	0.96
Manhattan	0.86	0.92	0.84	0.86

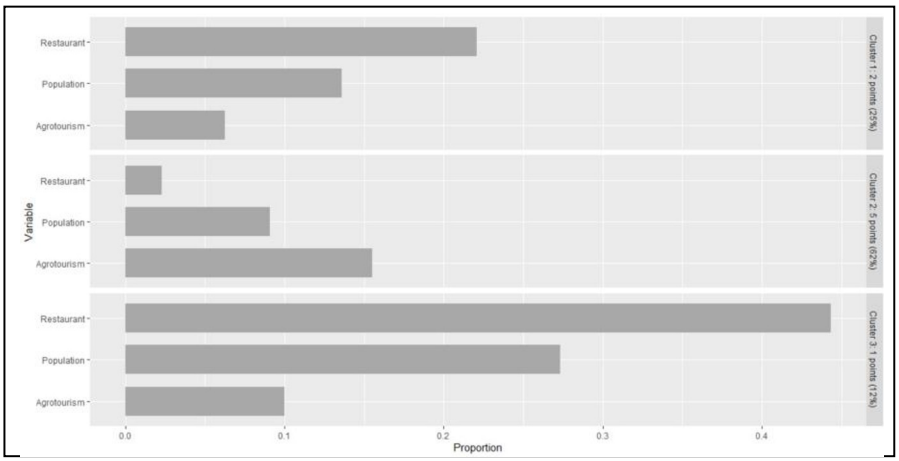

Fig. 2. Bar-plot of clusters 1, 2, and 3

Based on the relative and internal criteria, the suitable number of clusters is three. Cluster 1 has the municipals of Tangerang and Tangerang Selatan as members (2 members), while Cluster 2 consists of regencies of Pandeglang, Lebak, and Serang, and municipals of Serang and Cilegon (5 members). Cluster 3 has only the regency of Tangerang as a member. Then, the clustering result is plotted in a bar plot (Figure 2) and a spider plot (Figure 3) to describe each cluster's characteristics. Both plots apply proportion measure for comparability purpose.

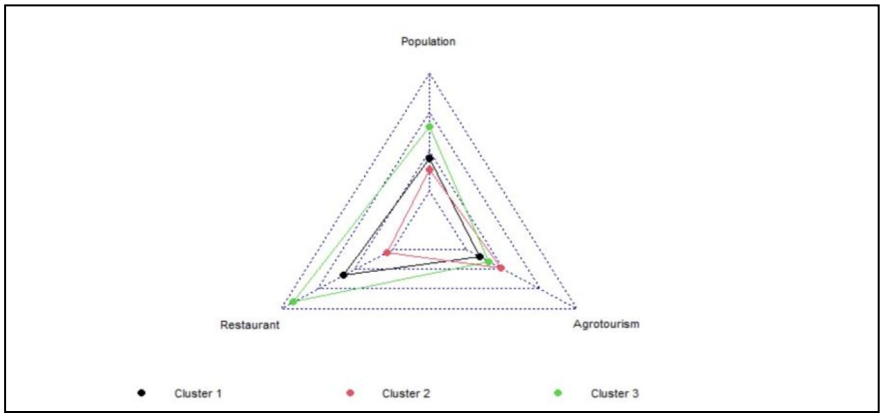

Fig. 3. Spider-plot of clusters 1, 2, and 3

Cluster 1 has the number of restaurants as the dominant variable (Figures 1 and 2). The restaurant is one of the important support facilities for tourism. With the municipals of Tangerang and Tangerang Selatan being the most densely populated in Banten Province, the restaurant is a promising business. Unique products can be very crucial to attracting people in densely populated regions [22]. The credit has also a growth of 18% [23]. Although its tax contributes to the regional revenue [16], the harvested area of these two municipals is the lowest [12]. Thus, cluster 1 is the least potential cluster of agrotourism.

Cluster 2, on the other hand, has superiority in the number of agrotourism sites (Figures 1 and 2). Cluster 2 has the highest members, i.e., 5 out of 8 regions in Banten Province (62%). It consists of three regencies and two municipals. They also have the most harvested area among regencies and municipals [12], respectively. The people's mobility index from 2020 to 2022 in Banten Province to the park/ garden sites increased by 48.2 [23], which was the highest. It indicates that agrotourism has very potential in this cluster. When the number of

sites also offers its uniqueness and novelty, the tourism experience increases visitors' interest and positive emotions [24].

Similar to cluster 1, cluster 3 has a strength in the number of restaurants (Figure 1). However, it has the highest population (Figure 2) and becomes the most densely populated regencies. With the existence of restaurants as a support facility and a high market of customers, agrotourism is also potential in cluster 2. However, the level of apathy where commercialism relations between tourists and local people [25] has to be preserved in the agrotourism development [26]. Thus, cluster 3 is the second potential for agrotourism.

4 Conclusion

The clustering analysis of municipals and regencies in Banten Province based on their agrotourism potency resulted in three clusters. Cluster 1 indicated by its strength in the number of restaurants was the least potential cluster with the municipalities of Tangerang and Tangerang Selatan as members. On the other hand, the regency of Tangerang could be the second potential region for agrotourism, indicated by a potential market and the support facility, i.e., restaurants. Last, there were three regencies of Pandeglang, Lebak, and Serang, and two municipals of Serang and Cilegon being the most potential regions. They have strength in the number of agrotourism sites. The novelty and uniqueness of agrotourism sites have to be varied to maintain visitor enjoyment.

The authors would like to thank LPPM University of Sultan Ageng Tirtayasa for the financial support via the PPGB grant.

References

1. J. Abbas, R. Mubeen, P. T. Iorember, S. Raza, and G. Mamirkulova. Curr. Res. in Behav. Sci. **2** 1, (2021)
2. T. Pham. A. Nugroho. Ann. of Tour. Res. Empir. Insight. **3** 1, (2022)
3. M. I. A. Hamudy, and M. S. Rifki. Policy and Gov. Rev. **5** 1, (2021)
4. R. Amanda, R. P. Tambunan, and T. Waryono. *IOP Conf. Series: Earth and Environ. Sci.* **561**, (2020)
5. Dept. of Tourism. *Tour. in Number Banten* (2019)
6. E. Susanti, and D. Emilia. Econ.: J. of Econ. and Econ. Educ. **10** 85-91, (2021)
7. G. Kolvekova, E. Liptakova, L. Strba, B. Krsak, C. Sidor, M. Cehlar, S. Khouri, and M. Behun. Sustain. **11**, (2019)
8. N.Y. Titonova, M. A. Pervukhin, and V. E. Zyglina. Adv. in Econ. Bus. and Manag. Res. **79**, (2019)
9. M. Gomez-Vega, L. C. Herrero-Prieto, and M. V. Lopez. Tour. Econ. **28**, (2022)
10. N. Iswandhani, and M. Muhajir. *IOP Conf. Series: J. of Phys.* **974**, (2018)
11. G. Lukoseviciute, and L. N. Pereira. Tour. and Manag. Stud. **17**, (2021)
12. BPS-Statistics of Banten Prov. *Banten Prov. in Fig.* (2022)
13. M. Mamoodi, M. Roman, and P. Prus. Sustain. **14**, 4555 (2022)
14. G. Grillini, G. Sacchi, L. Chase, J. Taylor, C. C. Van-Zyl, P. v. D. Merwe, T. Streifeneder, and C. Fischer. Sustain. **14**, 7903 (2022)
15. M. Roman, and P. Grudzien. Agric. **11**, 458 (2021)
16. J. Sulaksana, D. Sudirno, and L. Suparto. Mimbar. **37**, 88-100 (2021)

17. W. Budiaji, and F. Leisch. Algorithm. **12**, 77 (2019)
18. P. Boileau, L. Kakinami, T. Barnett, M. Handerson, and L. Popovic. F1000Res. **11**, 771 (2022)
19. P. Rousseeuw. Journ. of Comput. and Appl. Math. **20**, 53-65 (1987)
20. W. Budiaji. Int. Journ. of Adv. in Intell. Inform. **5**, 76-88 (2019)
21. W. Budiaji, R. A. Riyanto, and S. Suherna. *IOP Conf. Series: J. of Phys.* **1861**, (2021)
22. T. Matherly, Z. G. Arens, and T. J. Arnold. Int. Journ. of Res. in Mark. **35**, 15-33 (2018)
23. BI Banten. Econ. Report. of Banten Prov. (2022)
24. O. Mitas, and M. Bastiaansen. Ann. of Tour. Res. **72**, 98-108 (2018)
25. E. Zaidan, J. F. Kovacs. Eur. Journ. of Sustain. Dev. **6**, 291-307 (2017)
26. I. R. Blackie, T. Tsholetso, and M. Keetile. Cog. Soc. Sci. **9**, 2200356 (2023)