International Journal of Industrial Engineering & Production Research December 2024 Vol. 35, No. 4: 1-16
DOI: 10.22068/ijiepr.35.4.2110

RESEARCH PAPER

Agglomeration Agricultural Zone Analysis Based on the Hybrid Hierarchical Clustering Method in Madura, Indonesia

Iffan Maflahah[1*], Dian Farida Asfan[2], Selamet Joko Utomo[3], Fathor AS[4] & Raden Arief Firmansyah[5]

Received 6 July 2024; Revised 20 September 2024; Accepted 20 October 2024;

ABSTRACT
Madura Island, comprising four regencies, exhibits a diverse array of agricultural resource potential, particularly in paddy, maize, cassava, and soybeans. Althought the Gross Regional Domestic Product assesses economic progress. it inadequately reflects the whole spectrum of potential within each region. A comprehensive observation of this diversity is required to facilitate a more focused development approach. This study aims to employ a hybrid hierarchical clustering method to delineate and classify the geographical regions of Madura Island according to their agricultural potential. K-means clustering, that part of hybrid hierarchical clustering approach was used to achieve aims of research. Number of farmers, land area, and commodities production were variable that used to classify regional based on its potentials. First, hierarchical method was performed to determine the appropriate number of clusters then K-means clustering was applied to classify the regions based on agricultural commodities. The results show effectively determined Madura Island's agricultural potential using the hybrid hierarchical clustering method, which categorizes locations based on characteristics of agricultural production. The research reveals six clusters, each characterized by a unique profile of primary commodity production, including paddy, corn, soybeans, and cassava. Implication of this result is offering insights into regional development of Madura based on agricultural potential.

KEYWORDS: Agricultural zone; Hybrid hierarchical clustering; K-means; Madura.

1. Introduction

Regional development is connected to leveraging local natural resources to drive economic growth and improve community well-being [1,2]. Economic metrics such as Gross Regional Domestic Product measure prosperity and expansion but fail to capture the diversity of regional potentials [3]. Analytical methods for recognizing diversity are necessary to facilitate more targeted development strategies since larger administrative regions show a variety of resource potentials [4–7]. Madura Island, consisting of Sumenep, Pamekasan, Sampang, and Bangkalan regency, is a prime example of the diversity. The primary source of income is agriculture, which includes paddy, maize, legumes, and cassava. Regional diversity should be understood to create more precisely targeted strategies since production data suggests significant disparities across the regencies. Corn production increased from 81,932 to 935,842 tonnes in Pamekasan, while paddy production ranged from 108,480 to 217,810 tonnes [8]. Strategic resource management, funding for infrastructure, sustainable practices, and improved access to capital are implemented to optimize agricultural potential.

The potential for agricultural development has the potential to augment the income of the community. According to [9,10], paddy and maize are the commodities with the most agricultural potential due to widespread production. Therefore, targeted development strategies could be informed

* **Corresponding author: Iffan Maflahah**
iffanmaflahah@trunojoyo.ac.id

1. ***Department of Agroindustrial Technology, Faculty of Agriculture, Universitas Trunojoyo Madura, Bangkalan 69162, Indonesia***.
2. ***Department of Agroindustrial Technology, Faculty of Agriculture, Universitas Trunojoyo Madura, Bangkalan 69162, Indonesia.***
3. ***Departement of Development Economic, Faculty of Economic and Business, Universitas Trunojoyo Madura, Bangkalan 69162, Indonesia.***
4. ***Departement of Management, Faculty of Economics and Bussines, Universitas Trunojoyo Madura, Bangkalan 69162, Indonesia.***
5. ***Departement of Management, Faculty of Economics and Bussines, Universitas Trunojoyo Madura, Bangkalan 69162, Indonesia.***

by regional aggregation based on these key commodities [11–13].

Cluster analysis is a flexible unsupervised learning method used to categorize related instances [2,14,15]. Hierarchical and non-hierarchical methods were used in cluster analysis [15,16]. Limitations of hierarchical method include the inability to handle non-spherical or unevenly dense data. To mitigate this, the hybrid methods manage diverse cluster densities and configurations [17,18]. Hierarchical clustering uses the agglomerative or divisive method. Agglomerative is the most commonly used due to ease of implementation and ability to capture naturally occurring data patterns.

Combining the advantages of methods can obtain superior results in the hybrid clustering. In addition to delivering interpretable clustering, the hybrid cluster supports a wide range of densities and topologies [19–21]. A combination of hierarchical and non-hierarchical methods can improve regional classification accuracy and inform development decisions [22].

In agriculture, cluster analysis using the hybrid method is extensively used to efficiently classify commodities according to common characteristics [2,23–25]. A comprehensive understanding of market dynamics is essential for effective decision-making in agriculture sector [7,9,26,27]. This research used the hybrid hierarchical clustering method on agricultural dataset to unearth valuable information used to inform the development of strategies and policies [23,28–31]. Dependable and efficient hybrid solution is developed for managing the intricate diversity of agricultural data by integrating hierarchical and partitional methods [32]. The diversity of regional potential on Madura Island is thoroughly investigated by using the hybrid hierarchical clustering to facilitate the development of suitable strategies.

The hybrid method of hierarchical clustering is used with K-means [2,23–25]. The analysis commences with hierarchical method to determine the optimal number of clusters, and K-means is adopted to finalize the classification [9,26]. K-means is frequently used to aggregate data according to the potential agricultural commodities acquired [2,23,24,30,31]. Hierarchical clustering and K-means differ fundamentally in clustering method. In K-means, the number of clusters is determined in advance [32].

Agricultural commodity clustering is important for understanding market trends and making appropriate decisions. Previous research showed the use of K-means algorithm for analyzing agricultural commodities [33–35]. Smart agriculture uses K-means to report the location of wireless sensor networks [36]. This method can evaluate the influence of climatic change on the cultivation of root and tuber crops [37]. K-means method can evaluate leaf photos to identify plant diseases [38]. The method is used to identify clustering of paddy, soybeans, and hybrid coconut within a given region [39–41]. In selecting the appropriate algorithm, the characteristics of the data and the objectives of the analysis should be considered. Every algorithm comes with an individual set of advantages and disadvantages. Therefore, this research aims to evaluate the efficacy of K-means algorithm in analyzing a wide range of agricultural commodities while considering the distinctive characteristics and inherent complexity. The application of the algorithm for data analysis and identification of trends and patterns allows informed decisions. The K-means algorithm enables stakeholders to optimize plans, reduce risk, and discover new opportunities. However, the method is susceptible to certain limitations, such as the need to select an appropriate number of clusters and the high sensitivity to outliers. Further research is necessary to explore the potential for developing and implementing best practices in applying K-means to agricultural commodity data. The comprehension of effectively implementing K-means is enhanced to tackle complex problems.

K-means and hierarchical clustering offer significant insights for comprehending and examining data on agricultural commodities, consumers, energy systems, and tourism. In addition, K-means is a prevalent and effective clustering method that assigns similar objects based on Euclidean distance from the centroids [42–44]. Numerous research investigated the dispersion of small and medium companies with a variety of attributes. The K-means algorithm is also used to cluster consumer data and hybrid clustering applications extend to the healthcare field. Data clustering has the potential to show concealed patterns and enhance decision-making efforts [33,35,45]. The method significantly improves comprehension and data analysis to address practical problems. Despite the underutilization of the hybrid clustering with K-means in agricultural commodities, the potential is explored in analyzing market prices for vegetables, fruits, and food crop production data. Empirical support is offered for the notion that K-means show a degree of efficacy in identifying patterns and trends in agricultural commodity pricing and output [35,44,46,47]. Therefore, this research aims to use the hybrid hierarchical clustering to locate and group geographic areas in Madura Island according to agricultural potential. The expected outcomes include an improved understanding of agricultural commodities market, facilitation of more efficient marketing tactics, and optimization of resource allocation in the supply chain. Clustering can assist industry stakeholders, government, and policymakers in making well-informed decisions and effectively adapting to the dynamics of the commodity market.

2. Material and Methods

2.1.Material

The geographical coordinates of Madura Island are 7°0'S 113°20'E, with an approximate area of 5,379 square kilometers. The research includes four regencies on Madura Island (fig 1.)
and data per sub-district is used for each regency.

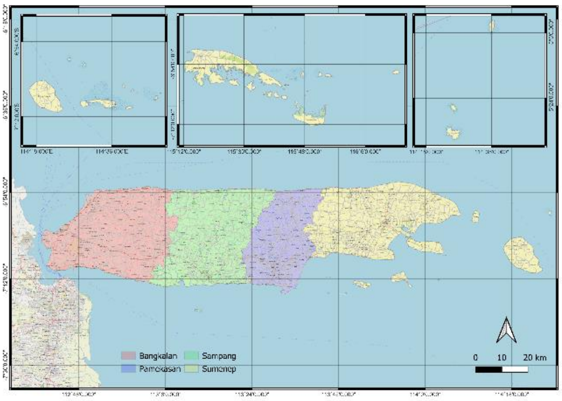

Fig. 1. Research area

The research area in Sumenep Regency consists of the sub-districts of Pragaan, Bluto, Saronggi, Giligenting, Talango, Kalianget, Sumenep, Batuan, Lenteng, Ganding, Guluk – Guluk, Pasongsongan, Ambunten, Rubaru, Dasuk, Manding, Batuputih, Gapura, Batang – Batang, Dungkek, Nonggunong, Gayam, Raas, Sapeken, Arjasa, Kangayan, and Masalembu. In addition, the research area in Pamekasan Regency comprises the sub-districts of Tlanakan, Pademawu, Galis, Larangan, Pamekasan, Proppo, Palengaan, Pengantenan, Kadur, Pakong, Waru, Batumarmar, and Pasean. Sampang Regency includes Sreseh, Torjun, Pangarengan, Sampang, Camplong, Omben, Kedungdung, Jrengik, Tambelangan, Banyuates, Robatal, Karang Penang, Ketapang, and Sokobanah subdistricts. Meanwhile, Bangkalan Regency consists of Kamal, Labang, Kwanyar, Modung, Blega, Konang, Galis, Tanah Merah, Tragah, Socah, Bangkalan, Burneh, Arosbaya, Geger, Kokop, Tanjungbumi, Sepulu, and Klampis subdistricts.
The data used is related to East Java Province, Sumenep, Pamekasan, Sampang, and Bangkalan Regency in Figures 2023 [48–51]. The data include the number of people using farming as a source of income, land area, as well as paddy, corn, soybeans, and cassava production. The data prepared as the basis for determining clustering of agricultural areas are shown in Tab 1.

Tab. 1. Agricultural commodity production data per sub-district

No	Sub District	Farmers (Person)	Fields (10,000 m^2)	Production (Tons)			
				Paddy	Corn	Soybeans	Cassava
1	Pragaan	3,554	1,387	97,299	40,476	-	12,513
2	Bluto	4,195	1,485	68,634	31,513	-	12,111
3	Saronggi	5,837	2,829	78,997	72,035	9,208	1,065
4	Giligenting	9,171	5,136	139,297	104,965	-	6,441
5	Talango	8,751	5,975	208,098	74,411	4,710	3,210
6	Kalianget	5,204	4,746	139,297	85,603	2,082	6,216
7	Sumenep	12,689	7,558	86,187	101,009	12,976	39,736
8	Batuan	5,557	2,663	229,987	20,588	-	2,804
9	Lenteng	4,564	1,143	105,208	41,231	-	1,423
10	Ganding	5,034	2,041	159,896	60,481	-	7,140
11	Guluk Guluk	1,011	325	97,537	9,459	-	-
12	Pasongsongan	5,660	1,392	338,894	6,811	-	10,493

13	Ambunten	3,824	1,258	139,914	36,264	-	7,964
14	Rubaru	17,772	6,095	215,834	77,378	-	29,865
15	Dasuk	13,456	7,080	110,374	35,195	-	711
16	Manding	6,116	3,565	67,465	67,240	-	45,323
17	Batuputih	5,155	3,842	105,440	34,551	-	29,681
18	Gapura	7,828	3,935	82,281	36,634	-	3,103
19	Batang Batang	405	5,825	1,677	17,755	-	585
20	Dungkek	187	4,160	158	13,572	-	316
21	Nonggunong	294	5,163	5,432	12,893	-	36,408
22	Gayam	209	1,995	-	4,717	-	1,443
23	Raas	204	5,504	-	15,689	-	5,020
24	Sapeken	30	464	264	1,102	-	684
25	Arjasa	121	2,580	7,461	2,509	-	134
26	Kangayan	120	3,068	9,751	3,969	-	-
27	Masalembu	371	7,854	6,388	21,707	50	2,839
28	Tlanakan	327	5,712	4,317	18,236	4	1,402
29	Pademawu	510	6,682	9,381	15,306	29	587
30	Galis	336	7,583	2,405	23,653	-	2,871
31	Larangan	357	5,264	4,310	12,956	-	1,454
32	Pamekasan	319	6,730	3,497	16,600	-	4,854
33	Proppo	342	4,608	3,412	13,807	-	3,469
34	Palengaan	292	5,953	9,902	9,214	-	2,904
35	Pengantenan	462	8,964	1,930	25,529	-	5,305
36	Kadur	527	6,489	16,587	10,509	-	11,795
37	Pakong	347	6,460	8,045	14,253	-	2,638
38	Waru	543	6,061	4,044	14,431	-	9,461
39	Batumarmar	129	3,266	1,692	8,060	-	18,964
40	Pasean	387	7,438	2,888	19,764	-	5,818
41	Sreseh	400	2,785	1,318	8,185	-	-
42	Torjun	370	6,549	8,079	14,013	-	-
43	Pangarengan	753	19,079	36,452	24,724	-	12,389
44	Sampang	259	11,832	8,429	27,153	-	4,938
45	Camplong	109	3,200	-	7,789	-	2,465
46	Omben	12,369	1,406	7,369	3,420	-	874
47	Kedungdung	14,375	3,828	21,642	1,408	-	176
48	Jrengik	5,187	937	5,037	137	-	-
49	Tambelangan	11,969	498	2,651	7,947	-	1,234
50	Banyuates	8,362	1,066	5,540	381	-	-
51	Robatal	16,773	2,204	11,853	1,809	-	2,675
52	Karang Penang	23,983	537	2,781	11,308	-	4,971
53	Ketapang	17,233	473	2,478	5,217	-	5,718
54	Sokobanah	12,583	1,135	6,131	11,071	-	739
55	Kamal	8,286	1,582	8,317	621	-	4,364
56	Labang	14,063	1,157	6,393	10,866	-	-
57	Kwanyar	19,374	1,728	8,963	14,718	-	3,151
58	Modung	11,146	1,163	6,046	13,029	-	1,212
59	Blega	6,399	1,429	10,603	3,802	-	464
60	Konang	7,325	3,368	22,881	2,060	8	404
61	Galis	4,471	824	5,330	232	-	-
62	Tanah Merah	12,581	4,607	28,560	1,345	-	2,158
63	Tragah	17,306	2,072	3,856	8,379	-	3,411
64	Socah	13,748	-	-	10,594	-	43,711
65	Bangkalan	15,019	3,074	17,723	5,424	522	10,086
66	Burneh	7,776	5,463	33,646	2,407	57	476
67	Arosbaya	9,721	1,454	6	7,477	-	17,446
68	Geger	17,361	1,752	11,777	12,763	50	7,782
69	Kokop	12,914	1,040	4,083	6,454	12,586	5,479
70	Tanjungbumi	16,694	200	1,267	4,786	18,474	8,593
71	Sepulu	24,664	364	2,656	21,141	1,623	25,990
72	Klampis	14,832	448	3,009	11,466	8,366	3,915

2.2. Methods

Data analysis to cluster agriculture-based areas uses the hybrid hierarchical clustering method with K-means, and the stages include:

1. Collecting data on production quantities per commodity, land area, and the number of farmers.
2. Normalizing the data since the scales are different using the min-max method.
3. Performing hybrid hierarchical clustering.

   The hybrid hierarchical clustering combines top-down and bottom-up methods. Bottom-up algorithms are good at grouping small sample sizes, while top-down methods handle large sample sizes well. The algorithms merge smaller groups, while top-down grouping splits larger groups. Mutual clustering uses the largest distance between elements within the smallest group.

   The single-linkage method initiates the hybrid hierarchical clustering process, initializing each data point as a cluster. This research describes the calculation of Euclidean distances between clusters, as well as the iterative merging with minimum distances. Finally, the K-means clustering method uses the centroids of the resulting clusters as the initial centroids. The mutual cluster uses the largest distance between data elements in S, which is smaller than the smallest distance for excluded data formulated as follows [52]:

   $$x \in S, y \in S, d(x,y) > diameter\ (S) = \max_{w \in S, z \in S} d(w,z) \qquad (1)$$

   where d is the distance function between two data objects, S is a subset of data x, while y is another element excluded in S, w, and z. The distance between clusters is calculated using the Euclidean Distance formula [19,53,54]. In this research, the Euclidean Distance formula is as follows:

   $$d(x,y) = \sqrt{\sum_{i=1}^{p} \left(x_{ik} - y_{jk}\right)^2} \qquad (2)$$

   The data points are combined with the minimum distance and assigned to a single cluster after calculating the distance. The process continues with an analysis of the resulting clusters when the iteration reaches the final data point or a specified number of k. Subsequently, the average data values are calculated within each formed cluster to determine the centroid. The centroid values are adopted as the initial for the K-means clustering method grouping the regions based on agricultural commodities.
4. The subsequent step is to calculate the K-means clustering method using generated centroid values. The initial step is to initialize the centroids obtained from the Single Linkage method as the starting centroids. The number of formed centroids determines the number of clusters (k). The Euclidean Distance formula is used to show the distance between each data point and centroids. Each data point is assigned to the cluster with the closest centroids. In addition, new centroids are calculated using the newly grouped data members. This repeated procedure persists until the cluster memberships reach a point of stasis [55].
5. The hybrid hierarchical clustering method categorizes agricultural regions based on the numerical value of k.
6. The cluster membership results are assessed using the Silhouette Coefficient method [56].
7. The aim of evaluating the grouped arrangement for agricultural regions is to assess the effectiveness of the specific method. The silhouette coefficient is an evaluation method for determining the accuracy of the clusters. This method uses a fusion of separation and cohesion, with the computation procedures outlined as follows [52]:

a. The following formula determines the average distance between an object and every other object in the same cluster.

$$(a)i = \frac{1}{[A]-1} \sum j \in_{A_{j \neq i}} \quad d(i,j) \qquad (3)$$

Let a(i) represent the mean difference of an item to other objects in cluster A, d(i,j) denote the distance between data i and j, while A represents the cluster collectively.

b. The minimum value is determined by calculating the average distance of the object to others and applying the equation:

$$d(i,C) = \frac{1}{[c]} \sum j \in C \quad d(i,j) \qquad (4)$$

where d (i, C) is the average difference of the object (i) and C is a cluster other than A.

c. After calculating d(i, C) for every C, the following equation is used to determine the smallest value:

$$b(i) = \min c \ \neq a\ d_i, C \qquad (5)$$

Object (i) considers Cluster B, reaching the minimum (d(i, B)) and this is the second-best cluster for object (i).

d. The value of the silhouette coefficient is calculated using the following equation:

$$s(i) = \frac{b(i)-a(i)}{\max(a(i),b(i))} \tag{6}$$

The silhouette coefficient value indicates the level of structure within the clusters that have formed. The silhouette coefficient can be found in Tab 2.

Tab. 2. Silhouette value

No	Value Silhouette Coefficient	Stucture
1	$0.7 < SC \leq 1$	Strong Structure
2	$0.5 < SC \leq 0.7$	Medium Structure
3	$0.25 < SC \leq 0.5$	Weak Structure
4	$SC \leq 0.25$	NoStructure

[57,58]

The results of the silhouette coefficient value calculation vary from -1 to 1. Clustering value can be said to be good when positive, namely (ai < bi), and ai is close to 0. Therefore, the maximum silhouette coefficient value will be 1 when ai = 0. Object i is between two clusters when the value of s(i) is 0 due to unclear structure [56,59].
The process is complete.

3. Result and Discussion

According to data obtained from the BPS of Sumenep, Pamekasan, Sampang, and Bangkalan Regency, the four regencies on the island of Madura have paddy, corn, and soybean commodities in all regions. However, soybean production is very limited, and the fulfillment of soybean needs is obtained from outside the region. The average paddy, corn, and cassava production in 2023 was 39,848 tons, 20,753 tons, and 7,133 tons, respectively. This data analyzes agricultural potential on the island of Madura using the hybrid hierarchical clustering method (**Tab. 3**). The hybrid hierarchical clustering is used to group regions based on similar characteristics of agricultural potential to provide insights into the regional agricultural dynamics of Madura.

Tab. 3. Summary of production data

	Min	Max	Mean
Farmers	1.00	72	36.50
Fields	30	24,664	6,813.9
Paddy	0	338,894	39,848
Corn	137	104,965	20,753
Soybeans	0	18,474.	982.6
Cassava	0	45,323	7,133

Based on **Tab. 3,** the number of farmers included in production ranges from 1 to 72 people, with an average of 36.5. The land area used for production varies, with a minimum and maximum of 30 and 24,664 hectares, respectively. Regarding commodity production, the average paddy, corn, soybeans, and cassava production is 39,848, 20,753, 982.6, and 7,133 tons, respectively.
The data presented in Tab 1 and Tab 3 were tested for multicollinearity. The multicollinearity test is conducted to check the correlation between the variables. In addition, the variables are the number of farmers, the area of agricultural land, as well as the production of paddy, corn, soybeans, and cassava. Multicollinearity occurs when two or more independent variables are highly correlated with each other. This leads to inaccurate estimates and unstable regression models, as shown in
Fig. 2. Multicollinearity is shown when the correlation coefficient (r) of each independent variable is more than 0.8. However, the correlation coefficient associated with each independent variable below 0.8 shows the absence of substantial multicollinearity. The results show that there is no statistically significant relationship among the variables evaluated. Therefore, the variables can be regarded as independent and appropriate for additional statistical analysis. The hybrid hierarchical clustering algorithm was performed using the distance matrix obtained from the correlation matrix.

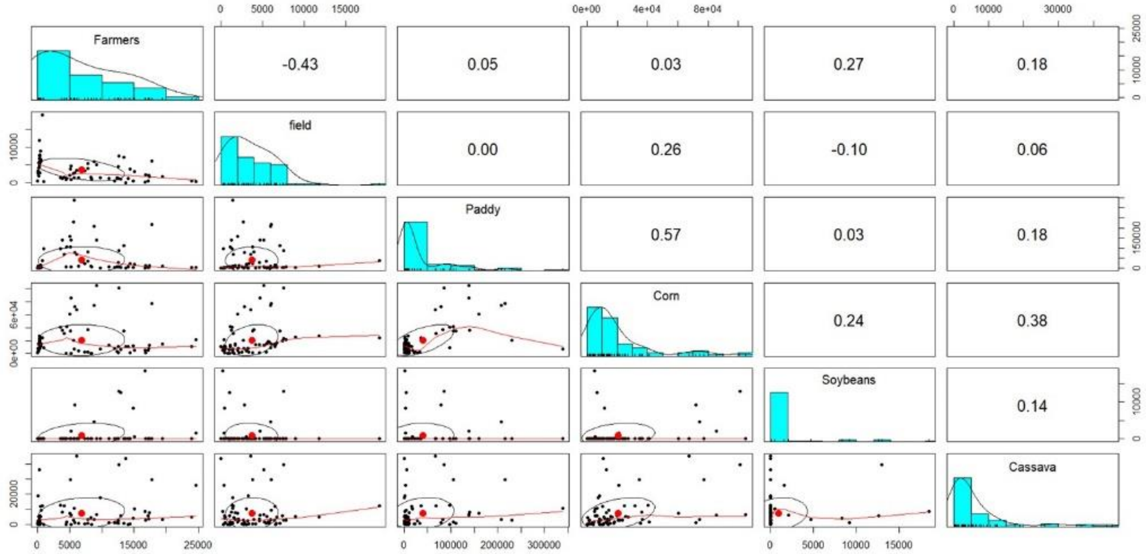

Fig. 2. Multicollinearity test

Before performing the hybrid hierarchical clustering analysis, data transformation is required to standardize the scale. This is necessary because the number of farmers, the area of agricultural land, as well as the production of paddy, corn, soybeans, and cassava have different measurement scales. The data is transformed using the scale method to standardize the variables by subtracting the mean and dividing by the standard deviation. Therefore, the variables are on a common scale, allowing for more meaningful comparison and analysis during clustering process. Data transformations are presented in **Tab. 4** to show the standardized values for each variable across the different regions. The hybrid hierarchical clustering algorithm is applied to the transformed data. The algorithm starts by considering each region, and iteratively merges the two closest clusters.

Tab. 4. Data transformation results

No	Farmers	Land Area	Paddy	Corn	Soybeans	Cassava
1	-0.4844	-0.7355	0.8602	0.8212	-0.2959	0.5091
2	-0.3892	-0.7045	0.4310	0.4480	-0.2959	0.4711
3	-0.1452	-0.2800	0.5861	2.1353	2.4773	-0.5742
4	0.3503	0.4485	1.4889	3.5065	-0.2959	-0.0655
5	0.2879	0.7135	2.5190	2.2343	1.1226	-0.3712
6	-0.2392	0.3254	1.4889	2.7003	0.3311	-0.0868
7	0.8731	1.2135	0.6938	3.3418	3.6121	3.0853
8	-0.1868	-0.3325	2.8467	-0.0069	-0.2959	-0.4097
9	-0.3343	-0.8125	0.9786	0.8527	-0.2959	-0.5403
10	-0.2645	-0.5289	1.7973	1.6542	-0.2959	0.0007
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
71	2.6526	-1.0585	-0.5568	0.0162	0.1929	1.7845
72	1.1915	-1.0320	-0.5516	-0.3867	2.2237	-0.3045

Hierarchical clustering analysis is performed to group the sub-districts based on the amount of food crop production. This analysis aims to cluster the sub-districts according to the potential of the main agricultural commodities, such as paddy, corn, soybeans, and cassava. Ward and Euclidean distance are used as the linkage method for hierarchical clustering. A comparison is made using the Pearson correlation test before calculating the distance between the data (**Tab. 5**).

The Pearson value ranges from -1.00 to 1.00, where a value of -1.00 and 1 shows a perfect negative linear and stronger positive relationship. The Pearson values between Gayam Sub-district and Lenteng, Ganding, Guluk-Guluk, Pasongsongan, Ambunten, Rubaru, Dasuk, and Manding Sub-districts are 0.038, 0.031, -0.221, -0.289, -0.075, 0.010, -0.037, and 0.492, respectively. This suggests a generally weak relationship between the sub-districts in terms of food crop production profiles.

Tab. 5. Pairwase matrix

	1	2	3	4	5	6	7	8	9	10	……
1	1.000	0.998	0.901	0.950	0.991	0.978	0.823	0.945	0.993	0.996	……
2	0.998	1.000	0.907	0.955	0.983	0.978	0.847	0.929	0.987	0.990	……
3	0.901	0.907	1.000	0.988	0.891	0.970	0.934	0.741	0.906	0.898	……
4	0.950	0.955	0.988	1.000	0.937	0.993	0.922	0.815	0.951	0.945	……
5	0.991	0.983	0.891	0.937	1.000	0.971	0.768	0.965	0.999	0.999	……
6	0.978	0.978	0.970	0.993	0.971	1.000	0.888	0.875	0.979	0.976	……
7	0.823	0.847	0.934	0.922	0.768	0.888	1.000	0.595	0.791	0.791	……
8	0.945	0.929	0.741	0.815	0.965	0.875	0.595	1.000	0.954	0.959	……
:	:	:	:	:	:	:	:	:	:	:	:

The sub-districts were grouped based on the potential of the main agricultural commodities. The linkage method was Ward and Euclidean distance used to minimize the total within-cluster variance. The grouping was based on a bottom-up algorithm using a single linkage. The initial grouping adopts data from each sub-district with several variables, including the number of farmers, the area of agricultural land, as well as the production of paddy, corn, soybeans, and cassava. Meanwhile, the Euclidean Distance algorithm obtains the straight-line distance between two places in the variable-filled three-dimensional space. This is used to determine the distance between the data points to obtain the cluster variables A dendrogram is hierarchical diagram used to show the structure of a group in examining the distances between data points. This diagram can be created by using the single linkage method**. Fig.** 3 shows the graphical structure of the sub-districts according to agricultural production characteristics.

Tab. 6. Euclidean distance calculation

	1	2	3	4	5	6	7	8	9	10	11	……
2	0.109											……
3	0.956	0.875										……
4	0.704	0.620	0.283									……
5	0.226	0.332	1.094	0.859								……
6	0.438	0.363	0.546	0.277	0.593							……
7	1.508	1.413	0.696	0.910	1.675	1.160						……
8	0.850	0.948	1.676	1.481	0.674	1.238	2.184					……
9	0.159	0.257	1.012	0.771	0.092	0.503	1.599	0.763				
10	0.140	0.159	0.257	1.012	0.771	0.092	0.503	1.599	0.763			……
11	0.140	0.246	1.035	0.792	0.087	0.523	1.610	0.746	0.045	0.822		……
12	0.140	0.246	1.035	0.792	0.087	0.523	1.610	0.746	0.045	0.822	0.920	……
:	:	:	:	:	:	:	:	:	:	:	:	……

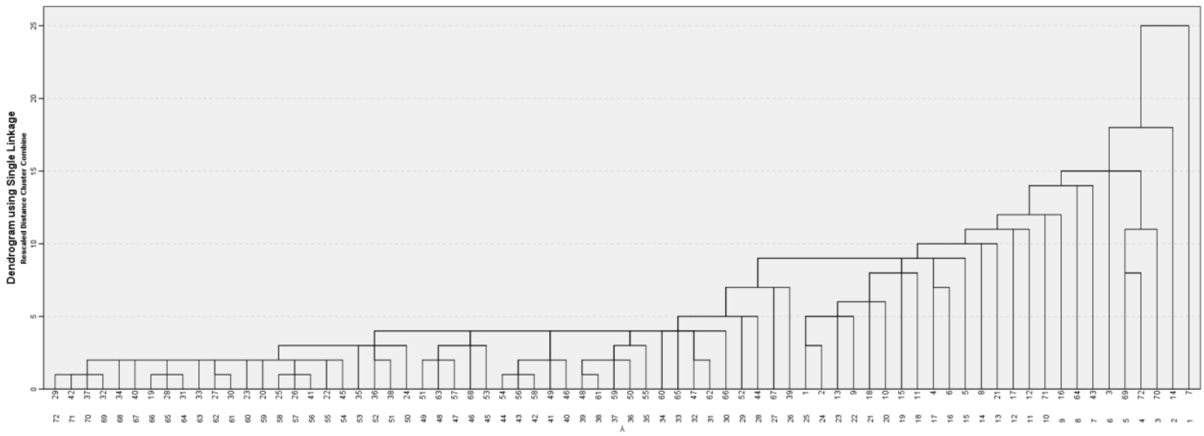

Fig. 3. Dendrogram cluster

The result of the dendrogram evaluation shows the existence of numerous distinct groupings, which report the degree of similarity in agricultural production between the sub-districts. The sub-districts within each group report comparable patterns and trends in agricultural outputs, including the production levels of critical crops such as paddy, maize, soybeans, and cassava. In the identification of these clusters, valuable knowledge is obtained on the regional dynamics and possible synergies within agricultural sector of the Madura region. This knowledge is used to make informed policy decisions and allocate resources optimally to promote sustainable agricultural development.

The number of clusters is established by estimating the distances between cluster mergers sufficiently large. The ideal number can be ascertained by vertically dividing the dendrogram at a specific elevation. According to the data in

Tab. 6Tab 6 and **Fig.** 3**,** The Euclidean distance of 0.177 is the shortest between Sapeken and Arjasa Subdistrict.

The Silhouette Coefficient method determines the ideal number of clusters. This evaluates the quality and robustness of the groupings by incorporating the concepts of separation and cohesion. Different cluster counts, such as 3, 4, 5, 6, 7, and 8, are used to determine the Silhouette Coefficient. The optimum number of clusters is determined by the value of k, which obtains the highest average Silhouette statistic. Clustering solution is determined to accurately match the data and generate the most significant sub-district groups according to agricultural production profiles. Error! Not a valid bookmark self-reference. presents the Silhouette Coefficient research for various numbers of clusters.

Tab. 7. The silhouette coefficient based on the number of clusters

Silhouette Coefficient	
K = 8	0.455
K = 7	0.462
K = 6	0.530
K = 5	0.472
K = 4	0.425
K = 3	0.525

Based on the data in**,** the silhouette analysis suggests that a 6-cluster grouping is a reasonably suitable match for the data. The value of 0.53 is within the range of 0.5 to 1, considered a strong and effective clustering solution. Therefore, clustering has successfully identified significant sub-district groupings with similar agricultural production profiles. A silhouette value of less than 0.2 is regarded as a suboptimal clustering result, suggesting a feeble and potentially unreliable clustering structure. Meanwhile, K = 4 records the lowest silhouette coefficient, with a value of 0.425, showing that dividing the data into 4 clusters is less optimal. In this context, K = 6 provides the most optimal clustering result among the available choices.

The six-cluster solution obtained by hierarchical clustering analysis using Ward significantly classifies agricultural regions according to production patterns of important commodities, including paddy, maize, legumes, and cassava. This regional clustering analysis provides a new understanding of the regional dynamics and possible synergies in agriculture industry (**Fig. 4**). These results can potentially inform policy choices and resource distribution to attain sustainable development.

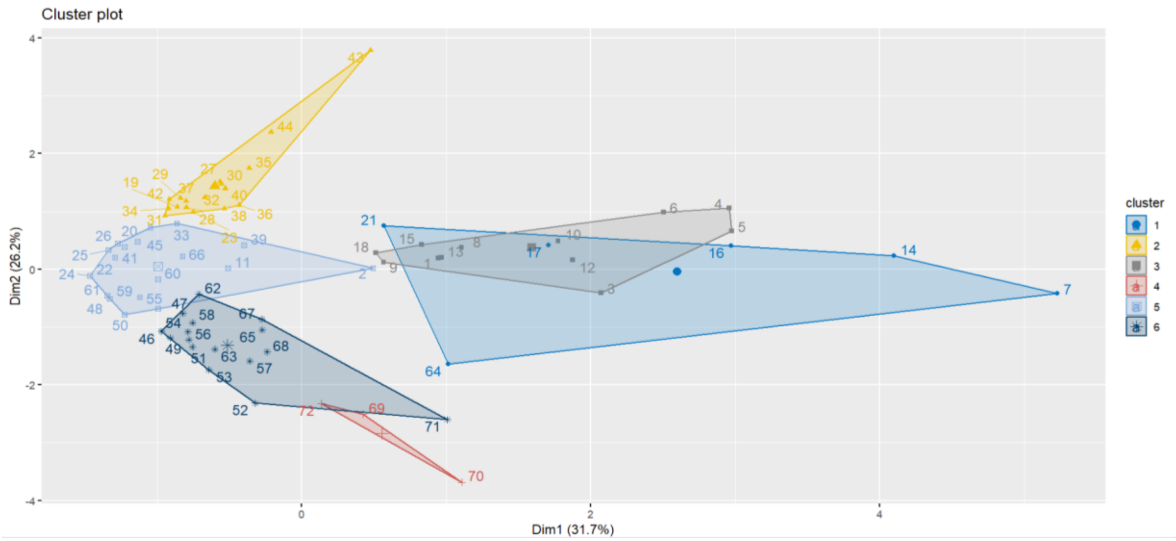

Fig. 4. Results of clustering of agricultural zones in madura

The investigation identified six clusters of sub-districts based on agricultural output patterns. The production characteristics of these regions are substantially consistent, where Cluster 1 includes the greatest group of 26 sub-districts. Clusters 2 and 6 contain five sub-districts, suggesting that there are smaller concentrations of sub-districts with comparable agricultural outputs. Cluster 3 comprises 11 sub-districts, while Cluster 4 is the smallest, with 2 sub-districts, showing more distinctive agricultural production patterns. Cluster 5 is the second largest, consisting of 23 sub-districts. This size further emphasizes the wide range of agricultural production. In detailing the specific members of the cluster, **Tab. 8** provides a comprehensive understanding of the regional dynamics and potential synergies within agricultural sector. This information informs policy decisions and resource allocation to support sustainable agricultural development.

Tab. 8. Members of each cluster

Cluster	Sub - District
1	Pragaan, Bluto, Giligenting, Talango, Sumenep, Batuan, Lenteng, Ganding, Guluk Guluk, Pasongsongan, Ambunten, Rubaru, Dasuk, Manding, Batuputih, Gapura, Batang Batang, Dungkek, Nonggunong, Gayam, Raas, Sapeken, Arjasa, Kangayan, Masalembu, dan Jrengik
2	Tanjungbumi, Sepulu, Saronggi, Omben, Ketapang
3	Kamal, Labang, Modung, Blega, Konang, Tragah, Socah, Arosbaya, Geger, Kokop, dan Klampis
4	Tanah Merah, Burneh
5	Bangkalan, Kalianget, Tlanakan, Pademawu, Galis, Larangan, Pamekasan, Proppo, Palengaan, Pengantenan, Kadur, Pakong, Waru, Batumarmar, Pasean, Sreseh, Torjun, Pangarengan, Sampang, Camplong, Kedungdung, Tambelangan, dan Banyuates
6	Kwanyar, Galis, Robatal, Karang Penang, Sokobanah

Based on the clustering results, additional analysis was conducted to profile the characteristics of each found cluster, revealing the distinct agricultural production patterns and trends that differentiate the various sub-district clusters. To enhance understanding of regional dynamics and potential synergies within the agricultural industry, the findings of this complete profiling analysis are presented in **Tab. 9**,

The geographical profiles derived from the clusters are:

1) Cluster 1 has the largest geographical area, and does not balance with the comparatively modest production levels of important crops such as paddy, soybeans, and cassava. This observation shows that the substantial land area

in the cluster is not fully exploited for agricultural yield.

2) Cluster 2 has a substantial farmer population and significant cassava output, surpassing the production levels of other key commodities such as paddy, corn, and soybeans. This implies a concentration on the production of cassava within the cluster.
3) Cluster 3 shows high paddy and corn production with moderate land area criteria. Therefore, the sub-districts in the cluster have succeeded in achieving superior productivity with limited land.
4) Cluster 4 has the greatest paddy production, reporting that the sub-districts possess ideal circumstances and methods for growing paddy as the main agricultural product.
5) Cluster 5 has the largest number of farmers and low agricultural yields. The condition shows the low quality of labor in the area.
6) Cluster 6 has a large farmer population and significantly high corn and soybean production. This profile shows that the sub-districts have developed expertise in cultivating the two commodities by optimizing labor and available land.

Tab. 9. Cluster profiling

Cluster	Farmers	field	Paddy	Corn	Soybeans	Cassava
1	622	6,196	7,145	14,096	5.40	3,928
2	9,995	2,587	36,199	29,284	325	36,222
3	7,578	3,662	133,285	56,741	617	8,245
4	5,608	2,028	284,440	13,699	0	6,648
5	11,606	1,612	12,771	5,825	25.2	2,937
6	12,593	2,415	34,709	39,150	12,322	11,758

: High
: Medium
: Low

Tab. 9 provides detailed cluster profiling of farmers, fields, and crop production, showing variations across six different clusters. Cluster 1 is characterized by 622 farmers managing 6,196 hectares to produce 7,145, 14,096, 5.40, and 3,928 tons of paddy, corn, soybeans, and cassava, respectively. Cluster 2, with 9,995 farmers, has significantly less land producing 36,199, 29,284, 325, and 36,222 tons of paddy, corn, soybeans, and cassava. Cluster 3 consists of 7,578 farmers cultivating 3,662 hectares to obtain the highest paddy production of 133,285 tons with 56,741, 617, and 8,245 tons of corn, soybeans, and cassava, respectively.

Cluster 4, with 5,608 farmers and 2,028 hectares, shows the highest paddy production of 284,440 tons. Cluster 5, with 11,606 farmers and 1,612 hectares, produces 12,771, 5,825, 25.2, and 2,937 tons of paddy, corn, soybeans, and cassava, respectively. Similarly, Cluster 6, the largest with 12,593 farmers, covers 2,415 hectares and produces 34,709, 39,150, 12,322, and 11,758 tons of paddy, corn, soybeans, and cassava, respectively. Each cluster shows different levels of production across crops, reporting variability in land use and farming output across regions.

Clustering analysis produces many useful insights into regional dynamics increasing the difficulty of exploiting the potential. Regional and sustainable development are significantly important in improving community welfare. Based on the analysis data obtained, policymakers will make the right decisions to improve agricultural welfare and long-term food security in the Madura region.

Fig. 5. displays a graphical depiction of the clusters discovered based on the clustering results. The area visualization provides a clear picture of the geographic composition and attributes of six different sub-district groups. This description of regional potential can be used as a consideration in industrial development based on similarities in regional potential.

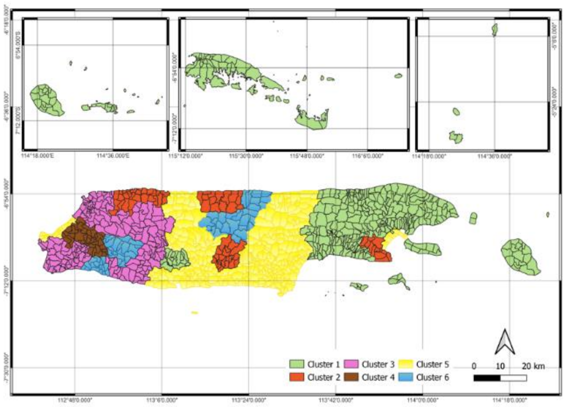

Fig. 5. Visualization of clustering results

The implications of this research include:

1. Examining the factors shaping distinct agricultural production patterns across clusters to inform targeted policy interventions.
2. Assessing climate change and environmental impacts on regional dynamics, and developing adaptive strategies to enhance sector resilience.
3. Exploring cross-cluster collaboration, knowledge sharing, and resource optimization to leverage production synergies and promote sustainable, equitable agricultural development.
4. Incorporating additional data sources to refine the analytical framework and provide a more comprehensive understanding of the regional agricultural system.

The results of clustering have identified the potential that focuses on resource allocation to promote sustainable agricultural growth. The geographical arrangement of clusters strengthens the understanding of regional groupings to enable broader decision-making and policy development. The hybrid hierarchical clustering methods are particularly effective for uncovering the complexity of complex regional agricultural systems. This method allows the development of strategies based on improving agricultural welfare and long-term food security in the Madura region.

4. Conclusion

The application of the hybrid hierarchical clustering method provided an overview of agricultural production trends and regional dynamics in the Madura region, Indonesia. Grouping analysis found six groups with varying characteristics. The results of this clustering showed heterogeneity in developing several regions with synergy in agricultural sector. Grouping each cluster provided information on the unique features and factors influencing agricultural productivity.

5. Acknowledgments

The author would like to thank the Directorate of Research, Technology and Community Service (DRTPM), Ministry of Education, Culture, Research and Technology for providing financial support through the Fundamental Research Schemes 2024 contract number 023/UN46.4.1/PT.01.03 /2024

References

[1] Kuswardhani N, Soni P, Shivakoti GP. Cluster analysis for classification of farm households based on socio-economic characteristics for technology adoption in agriculture: A case study of West Java province, Indonesia. J Food, Agric Environ. Vol. 12, No. 1, (2014), pp. 238-47.

[2] Sudarwati W, Hardjomidjojo H, Machfud, Setyaningsih D. K-means cluster approach in grouping of seaweed agroindustry development areas. IOP Conf Ser Earth Environ Sci. Vol. 1063, No. 1, (2022).

[3] Ramadhan A, Sutandi A, Pravitasari AE. Analysis of the Regional Development Level in Aceh Province. IOP Conf Ser Earth Environ Sci. Vol. 556, No. 1, (2020).

[4] Barmuta K, Matjonov B, Shichiyakh R, Dymchenko O. Cluster initiatives in agricultural sector: current state, problems and prospects of development. BIO Web Conf. Vol. 93, (2024), pp. 1-6.

[5] Wandel J. the Cluster-Based Development Strategy in Kazakhstan'S Agro-Food Sector: a Critical Assessment From an "Austrian" Perspective. Institute of Agricultural Development in Central and Eastern Europe (IAMO). (2010), pp. 7-25.

[6] Stovba E, Stovba A, Abdrashitova A, Baygildina A. Use of methods of cluster analysis in designing strategy of region's agro-food complex. Adv Econ Bus Manag Res. Vol. 38, No. 16, (2017), pp. 648-52.

[7] Rasulov TS, Makhmasobirova NU, Adizov AJ. Agricultural Clusters : Theory , World Practice And Development Opportunities In Uzbekistan. Solid State Technol. Vol. 63, No. 6, (2020), pp. 15530-52.

[8] BPS. Provinsi Jawa Timur dalam Angka (East Java Province in Number). Jakarta: Badan Pusat Statistik; (2023).

[9] Markos D, Mammo G, Worku W. Principal component and cluster analyses based characterization of maize fields in southern central Rift Valley of Ethiopia. Open Agric. Vol. 7, No. 1, (2022), pp. 504-19.

[10] Dini SK, Fauzan A. Clustering Provinces in Indonesia based on Community Welfare Indicators. EKSAKTA J Sci Data Anal. Vol. 1, No. 1, (2020), pp. 56-63.

[11] Iammarino S, Rodriguez-Pose A, Storper M. Regional inequality in Europe: Evidence, theory and policy implications. J Econ Geogr. Vol. 19, No. 2, (2019), pp. 273-98.

[12] Ruiga IR, Kovzunova ES, Teterin YA, Bugaeva S V., Burmenko TA. Assessment of the cluster structures functioning effectiveness in the agro-industrial complex: A methodological approach, practical implementation. IOP Conf Ser Earth Environ Sci. Vol. 677, No. 2, (2021).

[13] Mulawarman A, Sudrajat A, Hendri N, Kamar K, Mulyadi D, Budiyanto G, et al. FMADM for determining superior commodity at agroindustry area. Int J Eng Technol. Vol. 7, No. 4, (2018), pp. 4667-4673 International.

[14] Yousefzadeh R, Bodaghi H, Noubar K, Matin YA, Farahmand NF hi. Factors affecting the identification and prioritization of export industrial clusters of Iran to the eurasian economic union (case study : Gilan province). Int J Nonlinear Anal Appl. Vol. 14, (2023), pp. 1027-45.

[15] Fauziyah WM, Achmad AI. Penerapan Analisis Cluster Hybrid untuk Pengelompokan Kabupaten/Kota di Provinsi Jawa Barat Berdasarkan Indikator Kemiskinan Tahun 2022. Bandung Conf Ser Stat. Vol. 3, No. 2, (2023), pp. 566-74.

[16] Al Radhwani AMN, Algamal ZY. Improving K-means clustering based on firefly algorithm. J Phys Conf Ser. Vol. 1897, No. 1, (2021).

[17] Shetty P, Singh S. Hierarchical Clustering: A Survey. Int J Appl Res.Vol. 7, No. 4, (2021), pp. 178-81.

[18] Azzahrah F, Annas S, Rais Z. Hybrid Hierarchical Clustering dalam Pengelompokan Daerah Rawan Bencana Tanah Longsor di Sulawesi Selatan. VARIANSI J Stat Its Appl Teach Res. Vol. 4, No. 3, (2022), pp. 153-61.

[19] Ahmed M, Seraj R, Islam SMS. Review The k-means Algorithm: A Comprehensive Survey and Performance Evaluation. Electronics. Vol. 9, No. 8, (2020), p. 1295.

[20] Mittal M, Goyal LM, Hemanth DJ, Sethi JK. Clustering approaches for high-dimensional databases: A review. Wiley Interdiscip Rev Data Min Knowl Discov. Vol. 9, No. 3, (2019), pp. 1-14.

[21] Cohen-Addad V, Kanade V, Mallmann-Trenn F, Mathieu C. Hierarchical Clustering: Objective Functions and

Algorithms. J ACM. Vol. 14, No. 4, (2019), pp. 1-42.

[22] Chulkova G V., Lukasheva OL, Novikova NE, Trofimenkova E V., Podolnikova EM. Cluster approach for the development of the agro-industrial complex in the region. IOP Conf Ser Earth Environ Sci. Vol. 677, No. 2, (2021).

[23] Yuan C, Yang H. Research on K-Value Selection Method of K-Means Clustering Algorithm. Multidiscip Sci J. Vol. 2, No. 2, (2019), pp. 226-35.

[24] Xie H, Zhang L, Lim CP, Yu Y, Liu C, Liu H, et al. Improving K-means clustering with enhanced Firefly Algorithms. Appl Soft Comput J [Internet]. Vol. 84, (2019), p. 105763.

[25] Sofiyessi E, Marimin, Eriyatno, Sutrisno. The shallot agro-industrial cluster based on regional characteristic with soft system methodology approach: A conceptual design. IOP Conf Ser Earth Environ Sci. Vol. 443, No. 1, (2020).

[26] Руйга ИР, Kovzunova ES, Teterin YA, Bugaeva S V, Burmenko T. Assessment of the cluster structures functioning effectiveness in the agro-industrial complex: a methodological approach, practical implementation. IOP Publ. Vol. 677, No. 2, (2021), pp. 022057-022057.

[27] Okutaeva ST, Yesmagulova ND, Kishko N V. Cluster Approach to The Development of Regional Agro-Industrial Complex Based on The Materials of Akmola Region. (2018), p. 1400010134.

[28] Sumaryanti L, Widjastuti R, Tempola F, Ismanto H. Classification of Agriculture Area based on Superior Commodities in Geographic Information System. Int J Adv Comput Sci Appl. Vol. 13, No. 10, (2022), pp. 115-21.

[29] Akhmadeev R, Redkin A, Glubokova N, Bykanova O, Malakhova L, Rogov A. Agro-industrial cluster: Supporting the food security of the developing market economy. Entrep Sustain Issues. Vol. 7, No. 2, (2019), pp. 1149-70.

[30] Feizi YA, Sharbatdar MK, Mahjoub R, Raftari M. A hybrid unsupervised learning method for structural health monitoring by artificial neural networks and k-means clustering. Sci Iran. (2023).

[31] Nacife JM, Soares F, Castoldi G. Cluster Analysis Referring to Rural Enterprises of Sugarcane Local Productive Arrangement (LPA) in Quirinópolis, Brazil. Int J Bus Adm. Vol. 10, No. 5, (2019), p. 1.

[32] Laurenso J, Jiustian D, Fernando F, Suhandi V, Rochadiani TH, K-means A. Implementation of K-Means , Hierarchical , and BIRCH Clustering Algorithms to Determine Marketing Targets for Vape Sales in Indonesia. J Appl Informatics Comput. Vol. 8, No. 1, (2024), pp. 62-70.

[33] Fira A, Rozikin C, Garno G. Komparasi Algoritma K-Means dan K-Medoids Untuk Pengelompokkan Penyebaran Covid-19 di Indonesia. J Appl Informatics Comput [Internet]. Vol. 5, No. 2, (2021), pp. 133-138.

[34] Fadhillah DAF, Faisol A, Vendyansyah N. Penerapan Metode K-Means Clustering Pada Pemetaan Lahan Kopi Di Kabupaten Malang. J Mhs Tek Inform [Internet]. Vol. 6, No. 1, (2022), pp. 162-170.

[35] Puntoriza P, Fibriani C. Analisis Persebaran UMKM Kota Malang Menggunakan Cluster K-means. J Inf Syst [Internet]. Vol. 5, No. 1, (2020), pp. 86-94.

[36] Correia FP, Silva SR da, Carvalho FBS de, Alencar MS de, Assis KDR, Bacurau RM. LoRaWAN Gateway Placement in Smart Agriculture: An Analysis of Clustering Algorithms and Performance Metrics. Energies. Vol. 16, No. 5, (2023), pp. 1-21.

[37] Babatunde G, Abidemi Emmanuel A,

Oluwaseun OR, Bukola Bunmi O, Ehiedu Precious A. Impact of Climatic Change on Agricultural Product Yield Using K-Means and Multiple Linear Regressions. Int J Educ Manag Eng. Vol. 9, No. 3, (2019), pp. 16-26.

[38] Devi N, Leela Rani P, Guru Gokul AR, Kannadasan R, Alsharif MH, Jahid A, et al. Categorizing diseases from leaf images using a hybrid learning model. Symmetry (Basel). 13(11): 2021, pp. 1-14.

[39] Supriyatna A, Carolina I, Widiati W, Nuraeni C. Rice Productivity Analysis by Province Using K-Means Cluster Algorithm. IOP Conf Ser Mater Sci Eng. Vol. 771, No. 1, (2020).

[40] Andema H, Defit S, Putra Indonesia YPTK Padang U. Optimalisasi Penggunaan Lahan Perkebunan Kelapa Hibrida Menggunakan K-Means Clustering. J Inform Ekon Bisnis [Internet]. Vol. 2, No. 2, (2020), pp. 32-8.

[41] Liu D, Tang Z, Cai Y. A Hybrid Model for China's Soybean Spot Price Prediction by Integrating CEEMDAN with Fuzzy Entropy Clustering and CNN-GRU-Attention. Sustain. Vol. 14, No. 23, (2022).

[42] Chhabra G, Vashisht V, Ranjan J. Crime Prediction Patterns Using Hybrid K-Means Hierarchical Clustering. Jour Adv Res Dyn Control Syst. Vol. 11, No. 06, (2019), pp. 1249-1258.

[43] Ezugwu AE, Ikotun AM, Oyelade OO, Abualigah L, Agushaka JO, Eke CI, et al. A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. Eng Appl Artif Intell [Internet]. 110(January), (2022), p. 104743.

[44] Rahmadhan R, Wasesa M. Segmentation using Customers Lifetime Value: Hybrid K- means Clustering and Analytic Hierarchy Process. J Inf Syst Eng Bus Intell. Vol. 8, No. 2, (2022).

[45] Rochman EMS, Khozaimi A, Suzanti IO, Husni H, Jannah R, Khotimah BK, et al. A combination of algorithm agglomerative hierarchical cluster (AHC) and K-means for clustering tourism in Madura-Indonesia. J Math Comput Sci. (2022), 1-19.

[46] Falakhi A. Pengolahan Data Pelanggan Dengan Tenik Clustering K-Means Di Aplikasi Weka. J Comput Sci Inf Syst [Internet]. Vol. 3, No. 2, (2023), pp. 54-60.

[47] Rathore V, Kushwaha RS. Applications of Hybrid K Means Clustering and Genetic Algorithm (HKCGA) in Applications of Hybrid K - Means Clustering and Genetic Algorithm (HKCGA) in the healthcare domain : A Review. Int J Eng Appl Sci Manag. Vol. 5, No. 6, (2024).

[48] BPS. Kabupaten Sumenep dalam Angka. Kabupaten Sumenep; (2024).

[49] BPS. Provinsi Jawa Timur dalam Angka [Internet]. Surabaya: Badan Pusat Statistik; (2024).

[50] BPS. Kabupaten Pamekasan dalam Angka. Pamekasan; (2021).

[51] Badan Pusat Statistik Kabupaten Bangkalan. Bangkalan Dalam Angka. Katalog BPS 14033526. (2020), p. 313.

[52] Tango T. Statistical methods for disease clustering. Springer Science & Business Media; (2010).

[53] Dalmaijer ES, Nord CL, Astle DE. Statistical power for cluster analysis. BMC Bioinformatics [Internet]. Vol. 23, No. 1, (2022), p. 205.

[54] Ultsch A, Lötsch J. Euclidean distance-optimized data transformation for cluster analysis in biomedical data (EDOtrans). BMC Bioinformatics [Internet]. Vol. 23, No. 1, (2022), p. 233.

[55] Bouhmala N. How good is the euclidean distance metric for the clustering problem.

In: 5th IIAI international congress on advanced applied informatics (IIAI-AAI). IEEE; (2016), p. 312-5.

[56] Dudek A. Silhouette index as clustering evaluation tool. In: Classification and Data Analysis: Theory and Applications 28. Springer International Publishing,; (2020).

[57] Wang F, Franco-Penya HH, Kelleher JD, Pugh J, Ross R. An analysis of the application of simplified silhouette to the evaluation of k-means clustering validity. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 10358 LNAI: (2017), pp. 291-305.

[58] Shutaywi M, Kachouie NN. Silhouette Analysis for Performance Evaluation in Machine Learning with Applications to Clustering. Vol. 23, Entropy. (2021).

[59] Batool F, Hennig C. Clustering with the Average Silhouette Width. Comput Stat Data Anal. (2021), p. 158.

Follow this article at the following site:

Iffan Maflahah, Dian Farida Asfan, Selamet Joko Utomo, Fathor AS & Raden Arief Firmansyah. Agglomeration Agricultural Zone Analysis Based on the Hybrid Hierarchical Clustering Method in Madura, Indonesia. IJIEPR 2024; 35 (4) :1-16
URL: http://ijiepr.iust.ac.ir/article-1-2110-en.html