Bulletin of Applied Mathematics and Mathematics Education
Vol. 6, No. 1, April 2026, pp. 37-52
e-ISSN: 2776-1029 p-ISSN: 2776-1002 DOI: 10.12928/bamme.v6i1.15064
APPLIED MATHEMATICS

Clustering of productivity in the food, plantation, and farm sectors in Mamasa Regency using K-Means clustering analysis

Muhammad Zulfadhli [1], Kesumaning Dyah Larasati [2,*], Rizky Fitria Ramadhanni [3], Raynaldi Anggiat Samuel Siahaan [4], Mochamad Fatih Romadlon [5]

[1] Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2] Department of Geomatics Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[3] Department of Urban and Regional Planning, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[4] Department of Information Systems, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[5] Department of Development Studies, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
*Corresponding author

Abstract
The productivity of the food, *plantation*, and *farm sectors* is the main driver of the economy in Mamasa Regency. However, there are significant disparities between subdistricts, requiring targeted development strategies. This study aims to group 17 subdistricts based on the productivity characteristics of the three main sectors using the K-Means Clustering method. The secondary data analyzed includes productivity, production, planted area, and the number of farmers/ranchers for each sector. The research stages include descriptive analysis, identification of leading commodities, and classification of subdistricts based on the productivity of each sector. The analysis results divide the subdistricts into three main clusters with different characteristics. Cluster 1 (High Productivity) is dominated by leading subdistricts, such as Pana for cattle farming, Nosu for Arabica coffee, and Tabulahan for patchouli. Cluster 2 (Medium Productivity) includes subdistricts with balanced performance in several commodities, while Cluster 3 (Low Productivity) consists of subdistricts that still face challenges in land optimization, production, and cultivation efficiency. The implication of this study is cluster-based policy recommendations. Local governments are advised to implement specific strategies, such as developing cattle breeding centers in Pana, processing Arabica coffee in Nosu, and patchouli industry in Tabulahan. For low-productivity clusters, interventions are directed at improving infrastructure, access to inputs, and technological assistance. With this evidence-based strategy, local potential can be optimized, regional disparities reduced, and economic growth in Mamasa Regency can be more inclusive and sustainable.

Keywords: Applied mathematics, ecludian distance, K-Means clustering, mamasa, productivity

How to cite: Zulfadhli, M., Larasati, K. D., Ramadhanni, R. F., Siahaan, R. A. S., & Romadlon, M. F. (2026). Clustering of productivity in the food, plantation, and farm sectors in Mamasa Regency using K-means clustering analysis. *Bulletin of Applied Mathematics and Mathematics Education, 6*(1), 37-52. https://doi.org/10.12928/bamme.v6i1.15064

Article history: Received 25/11/2025, Accepted 25/05/2026, Published 25/06/2026

Correspondence address: Department of Geomatics Engineering, Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo, Surabaya, East Java, 60111, Indonesia. E-mail: kesumaningdyahlaras@gmail.com

INTRODUCTION

Mamasa Regency is an agrarian region with substantial potential in the *food crops*, *plantation*, and livestock sectors. These three sectors play a strategic role in driving regional economic growth and supporting the livelihoods of the majority of the local community. However, agricultural development across these sectors remains uneven throughout the regency. Significant productivity disparities exist among subdistricts; while some areas demonstrate relatively high performance, others continue to experience persistent stagnation.

Preliminary data indicate that food crop productivity, particularly rice and corn commodities, varies considerably across subdistricts, as reflected by a high standard deviation of 3,990.

Subdistricts such as Mehalaan, Rantebulahan, and Tabulahan have achieved productivity levels above 5.5 tons/ha, whereas Balla, Tabang, and Tandukkalua record lower productivity levels, ranging only between 4.0 and 4.3 tons/ha. Similar disparities are also observed in the *plantation sector*, where cocoa productivity ranges widely from 0 to 1,504 kg/ha. In the livestock sector, cattle and pig populations are highly concentrated in particular areas, such as Pana and Sumarorong.

Agricultural productivity gaps have been identified as one of the fundamental barriers to local economic development in agrarian regions, particularly when spatial imbalances occur across administrative areas (Prasetyo et al., 2022). To address this issue, cluster-based spatial mapping is considered important because it can help identify regional characteristics and support the formulation of more targeted development strategies (Sari & Wijaya, 2021). In this context, the K-Means Clustering method is an appropriate analytical approach for grouping regions based on similarities in productivity characteristics. Previous studies have shown that K-Means can classify regions using multiple productivity indicators with a high level of analytical precision (Gupta & Kumar, 2023; Haryanto & Zulfadhli, 2024; Ismail et al., 2024; Zhang et al., 2020; Zulfadhli et al., 2025). In addition, clustering approaches can reveal latent spatial patterns that may not be captured through conventional descriptive analysis (Gupta & Kumar, 2023).

Despite the growing number of clustering studies in agriculture, a significant research gap remains. Most existing studies tend to focus on a single agricultural subsector, such as *food crops*, *plantations*, or *livestock*, in *isolation*. This sector-silo approach limits the ability of policymakers to understand the comprehensive agricultural potential of a region. In reality, smallholder farmers in agrarian areas such as Mamasa often operate within an interconnected agricultural system, where crop residues can be used as livestock feed, while livestock manure can support soil fertility for *plantation* and food crop production. Therefore, analyzing these sectors separately may result in fragmented policy recommendations that do not fully reflect the integrated nature of local agricultural activities.

Based on this gap, the integration of *food crops*, *plantations*, and *livestock* into a single cluster analysis is necessary to capture the comprehensive, multi-sectoral agricultural profile of each subdistrict. This integrated approach enables regional development policies to be more synchronized with the actual cross-sectoral potential of each area, thereby reducing the risk of overlapping, conflicting, or redundant interventions. Accordingly, this study aims to map the spatial clustering of *food crops*, *plantation*, and livestock sectors simultaneously using the K-Means Clustering method. The results are expected to provide a precise, data-driven foundation for targeted regional agricultural development in Mamasa Regency.

RESEARCH METHOD

Research Type

This study is a quantitative study covering 17 subdistricts in Mamasa Regency, West Sulawesi. The data was obtained from secondary data from the Mamasa Regency Central Statistics Agency and the Mamasa Regency Agriculture Office.

Analysis Steps

The analysis in this study was conducted through several steps. First, data were collected on productivity, production, planted area, and number of farmers for the food and plantation sectors, as well as productivity, quantity, and number of farmers for the livestock sector, along with data on food planted area, plantation planted area, and total area for all subdistricts in Mamasa Regency. Second, the characteristics of the data, including the mean, minimum, maximum, and standard deviation, were analyzed, while also checking for missing data and outliers. Third, the variables were normalized so that they would have a comparable scale. Fourth, the subdistricts were grouped based on similarities in productivity across subsectors.

Fifth, the characteristics of each resulting cluster were identified and categorized as high, medium, or low. Sixth, the results of the analysis were interpreted to derive meaningful insights. Finally, conclusions were drawn and recommendations were formulated based on the findings.

Research Variables

The research variables used in this study are shown in Table 1.

Table 1. Research Variables

Sector	Commodity	Variable	Description	Data Scale	Unit
Food	Rice/Corn	$X_{1,1}$	rice/corn productivity	Rasio	ton/ha
		$X_{1,2}$	rice/corn production	Rasio	ton
		$X_{1,3}$	rice/corn planting area	Rasio	ha
		$X_{1,4}$	rice/corn farmers	Rasio	household
Plantation	Cocoa	$X_{2,1}$	cocoa productivity	Rasio	ton/ha
		$X_{2,2}$	cocoa production	Rasio	ton
		$X_{2,3}$	cocoa cultivation area	Rasio	ha
		$X_{2,4}$	cocoa farmers	Rasio	household
	Arabica Coffee	$X_{3,1}$	Arabica coffee productivity	Rasio	ton/ha
		$X_{3,2}$	Arabica coffee production	Rasio	ton
		$X_{3,3}$	Arabica coffee cultivation area	Rasio	ha
		$X_{3,4}$	Arabica coffee farmers	Rasio	household
	Robusta Coffee	$X_{4,1}$	Robusta coffee productivity	Rasio	ton/ha
		$X_{4,2}$	Robusta coffee production	Rasio	ton
		$X_{4,3}$	Robusta coffee cultivation area	Rasio	ha
		$X_{4,4}$	Robusta coffee farmers	Rasio	household
	Patchouli	$X_{5,1}$	patchouli productivity	Rasio	ton/ha
		$X_{5,2}$	patchouli production	Rasio	ton
		$X_{5,3}$	patchouli cultivation area	Rasio	ha
		$X_{5,4}$	patchouli farmers	Rasio	household
Farm	Cattle	$X_{6,1}$	cattle productivity	Rasio	tail/person
		$X_{6,2}$	cattle quantity	Rasio	tail
		$X_{6,3}$	cattle farmers	Rasio	person
	Buffalo	$X_{7,1}$	buffalo productivity	Rasio	tail/person
		$X_{7,2}$	buffalo quantity	Rasio	tail
		$X_{7,3}$	buffalo farmers	Rasio	person
	Pig	$X_{8,1}$	pig productivity	Rasio	tail/person
		$X_{8,2}$	pig quantity	Rasio	tail
		$X_{8,3}$	pig farmers	Rasio	person

Based on Table 1, the variables used in this study will be the main characteristics in forming each cluster.

Productivity in the Agricultural Sector and Regional Development

Agricultural productivity is a key indicator in regional development, especially in agrarian areas. According to (Todaro & Smith, 2020), agricultural productivity not only reflects the efficiency of production factor utilization, but also is a major determinant of regional economic growth. A study by (Fuglie et al., 2020) proves that a 1% increase in agricultural productivity can contribute to a 0.5-0.7% increase in regional income in rural areas.

In the context of Indonesia, research conducted by (Arifin et al., 2021) identified that disparities in productivity between regions are influenced by factors such as technology, quality of human resources, and access to markets. These findings are reinforced by (Maryudi et al., 2022), who emphasize the importance of a location-specific approach in addressing productivity gaps.

e-ISSN: 2722-2691|p-ISSN: 2722-2683

Analysis Steps

Descriptive statistics is the process of converting research data into tabular form so that it is easier to understand and interpret. According to (Anderson et al., 2020), descriptive statistics plays a role in studying methods of collecting, recording, compiling, and presenting research data in the form of frequency tables or graphs. Furthermore, statistical values are measured, such as mean/average, standard deviation, median, mode, and others. According to (Triola, 2021), descriptive statistics only describe and analyze specific data groups without drawing conclusions for generalization to larger data groups.

Because the variables utilize highly divergent units (Tons/Ha, Ha, Households, and Tails), raw data cannot be clustered directly without severe distance bias toward larger-scale variables. To standardize the metrics, a global Z-Score Normalization is executed:

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{\sigma_j} \tag{1}$$

Where X_{ij} represents the raw value of subdistrict i for variable j, $\bar{X}_j$ is the sample mean, and σ_j is the standard deviation of variable j.

Clustering Analysis

Clustering Analysis is a method in data science that aims to group objects or data into clusters that have certain similarities among the observed data. According to (Han et al., 2022), objects that fall within the similarity threshold will be grouped into one cluster, while objects that do not fall within the similarity threshold will be excluded from the cluster. The resulting clusters are expected to have high homogeneity between objects and high heterogeneity between clusters (Tan et al., 2021).

To mathematically justify the selection of k clusters, the Elbow Method utilizing the Within-Cluster Sum of Squares $WCSS$ is applied:

$$WCSS = \sum_{k=1}^{K} \sum_{i \epsilon C_k} \|Z_i - \mu_k\|^2 \tag{2}$$

The optimal number of clusters is chosen at the 'inflection point' where adding another cluster yields diminishing returns in reducing *WCSS*.

K-Means

K-Means is one of the non-hierarchical methods in cluster analysis. This method is carried out by placing objects with similar characteristics into one cluster. If objects have different characteristics, they will be removed and placed into another cluster (James et al., 2021). The following are the steps in performing cluster analysis using the K-Means method.

Determine the number of clusters k to be formed First, the number of clusters, k, to be formed is determined. Then, the initial cluster center points (k Centroid) are generated randomly using the following formula.

$$v = \frac{\sum_{1=1}^{n} x_i}{n} \quad ; i = 1,2,3, \dots, n \tag{3}$$

Where,

v = centroid in the cluster

x_i = object$-i$

n = number of objects

Next, the distance of each object to each centroid of each cluster is calculated using the formula.

$$d(x_i, y_i) = \|x_i - y_i\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad ; i = 1,2,3, \ldots, n \tag{4}$$

Where,

y_i = the i (th) y object

x_i = the i (th) x object

n = the number of objects

Each object is then assigned to the nearest centroid, and iteration is performed to determine the new centroid. This process is repeated until the centroid position no longer changes (Witten et al., 2023).

Clustering in Regional Development Analysis

Clustering has become an effective analytical tool in regional development planning. According to (Porter, 2021), clustering enables the identification of spatial patterns and unique characteristics of each region. Research by (Zhao et al., 2023) shows that the K-Means Clustering method is capable of grouping regions based on multiple indicators with an accuracy of up to 85%. The application of clustering in the agricultural sector has been developed by (Kumar & Singh, 2022) in India, who successfully identified four regional clusters based on crop productivity. Their research results show that a cluster-based approach can increase the effectiveness of policy interventions by up to 30% compared to a uniform approach.

Previous Research

Several previous studies have examined aspects of agricultural productivity in Indonesia. A study by (Saputra et al., 2021) in East Java successfully identified three rice productivity clusters using the K-Means method. Meanwhile, research by (Wahyuni & Fitriani, 2022) in South Sulawesi developed cluster-based mapping for *plantation* commodities. However, based on a literature review, there is still a gap in research on integrated cluster analysis that combines the three main sectors (*food*, *plantation*, and *livestock*) in a single analytical framework. In addition, specific studies in Mamasa Regency are still limited, even though this region has unique geographical and social characteristics.

RESULTS AND DISCUSSION

The following are the results of the analysis, which consist of descriptive analysis, analysis of leading commodities, and clustering analysis.

Descriptive Analysis

Food Sector

The results of the analysis of food sector productivity can be seen in Table 2 below.

Table 2. Characteristics of Food Sector Productivity Data

Variable	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Rice/Corn	5.2882	0.0492	0.2027	5	5.15	5.3	5.4	5.6

Based on Table 2, it can be seen that the average productivity value of the food sector for rice/corn commodities is 5.2882 with a standard deviation of 0.2027. This relatively small standard deviation value indicates that the variation in data between regions or observation periods is not too large, so that the level of rice/corn productivity tends to be uniform. The minimum value recorded is 5.00 and the maximum is 5.60, indicating that the difference between regions or times with the lowest and highest productivity is not too far apart.

Meanwhile, the median value of 5.30 shows that half of the data is below that figure. Overall, food sector data on rice/corn commodities show a relatively narrow distribution, indicating that productivity between regions or periods tends to be stable and does not show extreme deviations.

Plantation Sector

The results of the analysis of productivity in the *plantation* sector can be seen in Table 3 below.

Table 3. Characteristics of Food Sector Productivity Data

Variable	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Cocoa	422.2	56.7	233.8	0	230.8	440	634.6	707.9
Arabica coffee	552.3	85.4	352.3	0	300	648.6	760.3	1300
Robusta coffee	565.2	43.3	178.5	0	497.8	618.6	665.9	850
Patchouli	7.54	3.63	14.98	0	0	0	12	50

Based on Table 3, productivity in the *plantation* sector shows considerable variation between commodities. Cocoa has an average productivity of 422.2 with a standard deviation of 233.8, indicating a wide spread of data from 0 to 707.9. Arabica coffee has the greatest variation, with an average of 552.3 and a standard deviation of 352.3, as well as a productivity range of 0–1300, reflecting very high production disparities between regions. Robusta coffee has an average of 565.2 and a standard deviation of 178.5, indicating moderate but still significant variation, with values ranging from 0 to 850. Meanwhile, patchouli shows very low productivity with an average of 7.54, a standard deviation of 14.98, and initial quartile and median values of 0, indicating that most regions do not produce patchouli, although some areas reach up to 50. Overall, these data show that the *plantation* sector has a high level of productivity variation, especially in Arabica coffee and cocoa.

Farm Sector

The results of the analysis of livestock sector productivity can be seen in Table 4 below.

Table 4. Characteristics of Food Sector Productivity Data

Variable	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Cattle	4.86	0.59	2.434	1.023	2.49	5.326	6.875	8.725
Buffalo	4.94	1.14	4.7	1.7	2.66	3.41	5.54	22
Pig	4.797	0.567	2.336	0	3.29	5.144	7.008	7.832

Based on Table 4, productivity in the livestock sector shows moderate to high variation between commodities. Cattle have an average productivity of 4.86 with a standard deviation of 2.434, indicating a fairly wide distribution of data with a minimum value of 1.023 and a maximum of 8.725. Buffalo commodities have the greatest variation with an average of 4.94 and a standard deviation of 4.7, as well as a productivity range of 1.7–22, illustrating significant differences in production between regions. Meanwhile, the pig commodity has an average of 4.797 with a standard deviation of 2.336, with a minimum value of 0 and a maximum of 7.832, indicating that there are regions that do not produce at all to regions that show fairly high productivity. Overall, the livestock sector shows high productivity variability, especially in the buffalo commodity.

Analysis of Leading Commodities

Food Sector

The results of the diagram for the food sector can be seen in Figure 1 below. Based on Figure 1, rice/corn productivity between subdistricts is relatively stable, ranging from 5.0 to 5.6. The subdistricts of Mehalaan, East Rantebulahan, and Tabulahan have the highest productivity (5.6), while Balla, Tabang, and Tandukkalua have the lowest (5.0). Most subdistricts are in the range of 5.2–5.4, indicating a fairly even distribution of productivity with insignificant differences between regions.

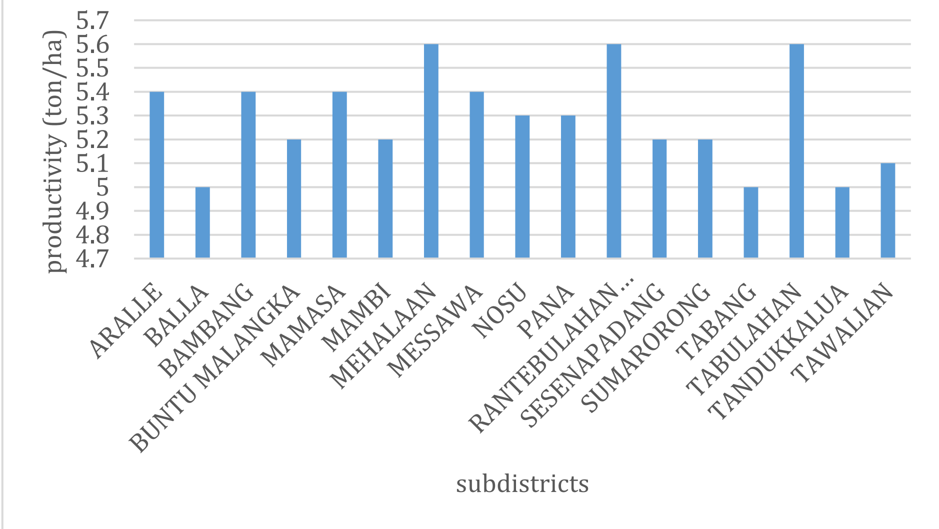


Figure 1. Food Sector Productivity

Plantation Sector

The results of the diagram for the food sector can be seen in Figure 1 below.

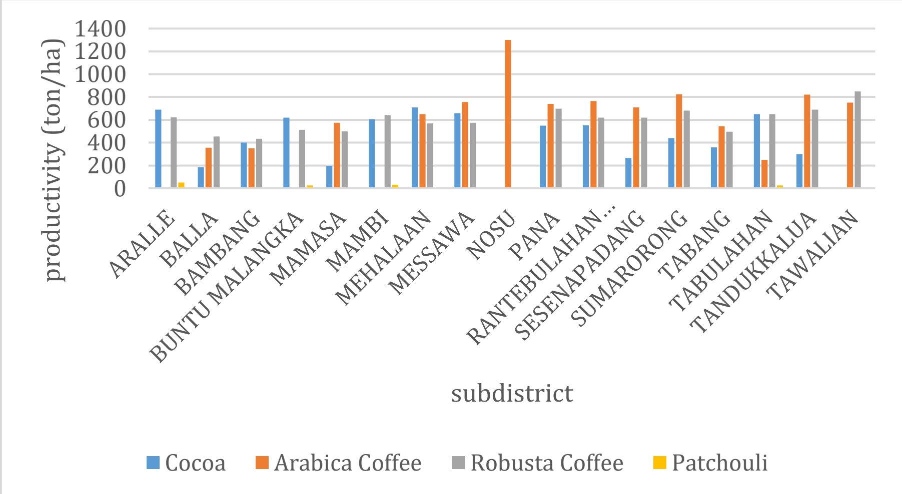


Figure 2. Productivity of the Plantation Sector

Based on Figure 2, the productivity of the *plantation* sector in various subdistricts shows that the productivity of Arabica and Robusta coffee tends to be higher than that of cocoa and patchouli in most areas. Nosu subdistrict has the highest Arabica coffee productivity (1,300) even though it does not produce cocoa or Robusta coffee, while Tawalian shows the highest Robusta coffee production (850) but does not produce cocoa or patchouli. The highest cocoa productivity is found in Mehalaan subdistrict (707.89), while patchouli productivity is generally low, with a maximum value of 50 in Aralle. Several subdistricts, such as Messawa, Pana, East Rantebulahan, and Sumarorong, show high productivity for Arabica and Robusta coffee, indicating that these areas have strong coffee *plantation* potential. Overall, *plantation* sector productivity is uneven across subdistricts, with coffee as the dominant commodity, while cloves and cocoa are only produced in certain areas.

Farm Sector

The results of the diagram for the livestock sector can be seen in Figure 3 below.

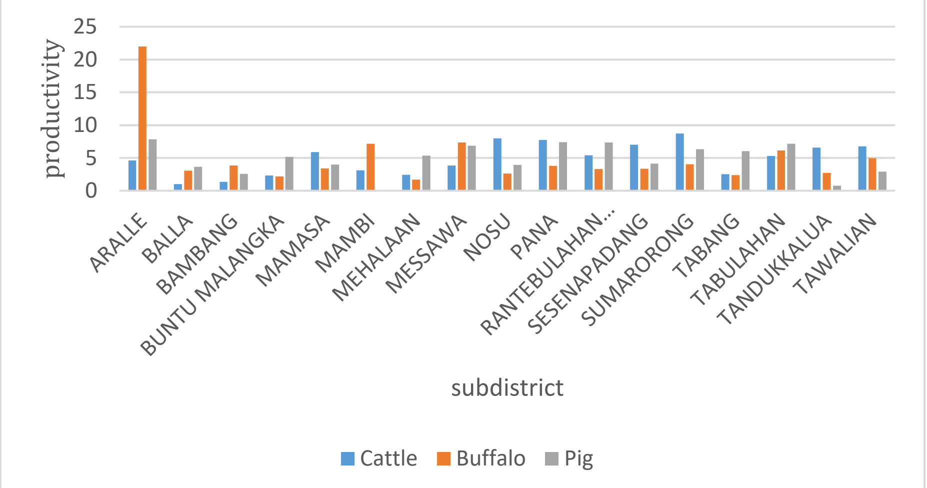


Figure 3. Farm Sector Productivity

Based on Figure 3, data on livestock sector productivity in various subdistricts shows considerable variation between commodities and regions. The highest cattle productivity was found in Sumarorong (8,725), followed by Nosu (8.0), Pana (7.75), Tawalian (6.75), and Tandukkalua (6,585), while the lowest was found in Balla (1,023) and Bambang (1,350). For buffalo, the highest value was recorded in Aralle (22), followed by Messawa (7,338), Mambi (7,167), and Tabulahan (6,133), while several subdistricts such as Mehalaan (1.7) and Buntu Malangka (2,182) had the lowest productivity. The highest pig productivity was found in Aralle (7,832), Pana (7,415), Tabulahan (7,151), and East Rantebulahan (7,362), while subdistricts such as Mambi (0) and Tandukkalua (0.793) had the lowest productivity. Overall, livestock productivity between subdistricts is uneven, with cattle, buffalo, and pigs having different distributions according to the focus of livestock farming in each region.

Clustering Analysis

Food Sector

The results of the clustering analysis for the food sector with rice/corn commodities obtained the cluster division given in Table 6 as follows.

The cluster division results in Table 6 show that subdistricts with relatively similar characteristics in terms of production, productivity, number of farmers, and food crop management patterns are grouped into the same cluster. In viewing the cluster levels based on the analysis results, the centroid values for each cluster are obtained through Table 7 below.

Table 6. List of Subdistricts for Each Food Sector Cluster

Commodities	Cluster	Subdistrict
Rice/Corn	Cluster 1	Aralle, Buntu Malangka, Mehalaan, Nosu, Rantebulahan Timur, Sesenapadang, dan Tabang
	Cluster 2	Balla, Mamasa, Mambi, Messawa, Pana, Sumarorong, Tabulahan, Tandukkalua, dan Tawalian
	Cluster 3	Bambang

The results of the analysis in Table 7 show that for the rice/corn productivity variable, cluster 3 has the highest centroid value, placing it in the group with the highest productivity, followed by cluster 1, while cluster 2 is below average. For the rice/corn production variable, cluster 2 actually has the highest value, indicating the fastest growth, while cluster 1 experiences slightly lower growth than the average and cluster 3 is below average.

Table 7. Centroid Values for Each Food Sector Cluster

Commodities	Variable	Cluster 1	Cluster 2	Cluster 3	Best Cluster
Rice/Corn	Productivity	0.199	-0.216	0.5513	3
	Production	-0.7767	0.7037	-0.8966	2
	Crop Area	-0.8693	0.7519	-0.6825	2
	Number of Farmers	-0.5915	0.2831	1.5921	3

For the rice/corn planting area variable, cluster 2 has the highest centroid value, making it the group with the largest planting area, while clusters 1 and 3 have planting areas below average. Furthermore, for the variable of the number of rice/corn farmers, cluster 3 shows the highest value, meaning that the number of farmers in that subdistrict is relatively high, followed by cluster 2, while cluster 1 has a small number of farmers. Based on productivity values for the food sector, the cluster division can be shown through the graph in Figure 5 below.

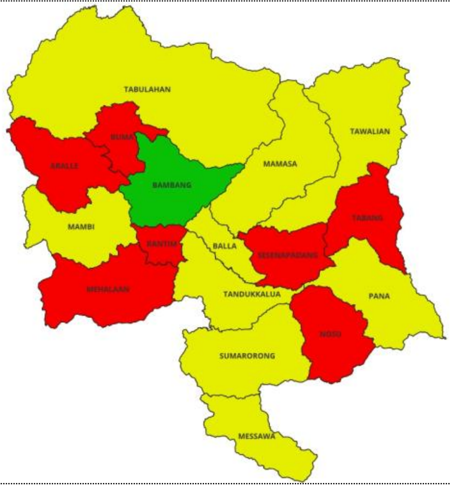


Figure 5. Clustering Results for Food Sector Productivity

Based on Figure 3, data on livestock sector productivity in various subdistricts shows considerable variation between commodities and regions. The highest cattle productivity was found in Sumarorong (8,725), followed by Nosu (8.0), Pana (7.75), Tawalian (6.75), and Tandukkalua (6,585), while the lowest was found in Balla (1,023) and Bambang (1,350). For buffalo, th the plantation sector has a high level of productivity variation, especially in Arabica coffee and cocoa.

Farm Sector

The results of the clustering analysis for the plantation sector with cocoa, Arabica coffee, Robusta coffee, and patchouli commodities obtained the cluster division given in Table 8 as follows.

The cluster division results in Table 8 show that subdistricts with relatively similar characteristics in terms of production, productivity, number of farmers, and plantation land management patterns are grouped into the same cluster

Table 8. Characteristics of Food Sector Productivity Data

Commodities	Cluster	Subdistrict
Cocoa	Cluster 1	Aralle
	Cluster 2	Balla, Buntu Malangka, Mamasa, Mambi, Mehalaan, Messawa, Nosu, Pana, Rantebulahan Timur, Sesenapadang, Sumarorong, Tabang, Tandukkalua, Tawalian
	Cluster 3	Bambang, Tabulahan
Arabica coffee	Cluster 1	Aralle, Buntu Malangka, Mambi, Tabulahan
	Cluster 2	Balla, Mamasa, Mehalaan, Rantebulahan Timur, Sesenapadang, Tawalian
	Cluster 3	Bambang, Messawa, Nosu, Pana, Sumarorong, Tabang, Tandukkalua
Robusta coffee	Cluster 1	Aralle, Buntu Malangka, Mambi, Mehalaan, Rantebulahan Timur, Sesenapadang, Tabulahan, Tandukkalua, Tawalian
	Cluster 2	Balla, Mamasa, Nosu
	Cluster 3	Bambang, Messawa, Pana, Sumarorong, Tabang
Patchouli	Cluster 1	Aralle, Mambi, Tabulahan
	Cluster 2	Balla, Bambang, Mamasa, Mehalaan, Messawa, Nosu, Pana, Rantebulahan Timur, Sesenapadang, Sumarorong, Tabang, Tandukkalua, Tawalian
	Cluster 3	Buntu Malangka

In viewing the cluster levels based on the analysis results, the centroid values for each cluster are obtained through Table 9 below.

Table 9. Centroid Values for Each Cluster in the Plantation Sector

Commodities	Variable	Cluster 1	Cluster 2	Cluster 3	Best Cluster
Cocoa	Productivity	1.1434	-0.1445	0.4396	1
Arabica coffee	Production	2.5775	-0.4265	1.6969	1
	Crop Area	1.7351	-0.4262	2.1157	3
	Number of Farmers	0.276	-0.3407	2.2471	3
Kopi Arabika	Productivity	-0.8879	-0.3706	0.8251	3
Robusta coffee	Production	-1.3904	0.2323	0.5953	3
	Crop Area	-1.3214	-0.0299	0.7807	3
	Number of Farmers	-0.8662	-0.6709	1.0701	3
Kopi Robusta	Productivity	0.4262	-1.3848	0.0636	1
Patchouli	Production	0.5407	-1.6878	0.0393	1
	Crop Area	0.0178	-1.3713	0.7906	3
	Number of Farmers	-0.601	-0.2569	1.2359	3
Cocoa	Productivity	1.8132	-0.503	1.0988	1
	Production	1.8183	-0.3888	-0.3999	1
	Crop Area	1.2657	-0.2788	-0.1725	1
	Number of Farmers	1.7619	-0.3871	-0.2534	1

The results of the analysis in Table 9 show that the characteristics of clusters for the plantation sector vary depending on the commodity. For cocoa, cluster 1 has the highest centroid value for productivity (1.1434) and production (2.5775), although the highest planted area and number of farmers are found in cluster 3 (2.1157 and 2.2471), indicating that cocoa production efficiency is most optimal in cluster 1, while cluster 3 has large land potential and a large number of farmers. For Arabica coffee, cluster 3 excels in almost all variables, with the highest productivity (0.8251), production (0.5953), planted area (0.7807), and number of farmers (1.0701), making this cluster the best performing group for Arabica coffee. For robusta coffee, cluster 1 has the highest productivity (0.4262) and production (0.5407), but the highest planted area and number of farmers are in cluster 3 (0.7906 and 1.2359), indicating that cluster 3 has greater development potential. Meanwhile, for patchouli, cluster 1 shows dominance in all variables, including productivity (1.8132), production (1.8183), planted area (1.2657), and number of farmers (1.7619), making cluster 1 the most optimal region for patchouli production. Based on these centroid values, the cluster division in the plantation sector can be visualized through the graph in Figure 6 to show areas with low, medium, and high production performance.

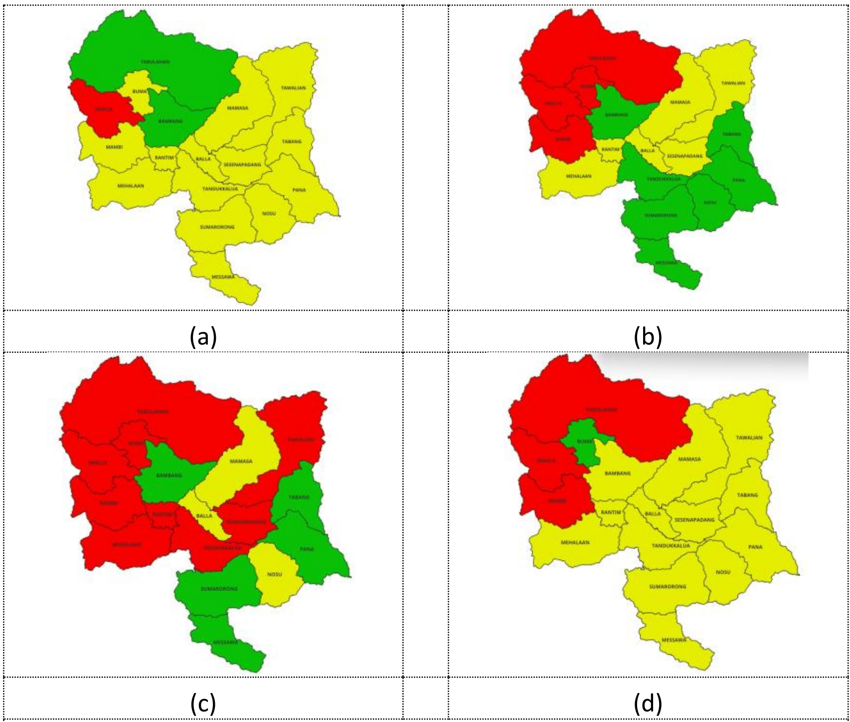


Figure 6. Clustering Results for Plantation Sector Productivity cocoa (a), Arabica coffee (b), Robusta coffee (c), and patchouli (d).

Based on Figure 6, in general, cluster 1 is dominated by subdistricts with high productivity and production for cocoa and patchouli, but has a more limited number of farmers and planted area for several other commodities. Cluster 2 is below average for almost all commodities and variables, indicating potential that still needs to be improved. Meanwhile, cluster 3 stands out in terms of planted area and number of farmers, particularly for cocoa and robusta coffee, and is the best cluster for arabica coffee, indicating high efficiency in land and labor utilization even though productivity per unit area is not always the highest.

Farm Sector

The results of the clustering analysis for the livestock sector with commodities of cattle, buffalo, and pigs obtained the cluster division given in Table 10 as follows.

The cluster division results in Table 10 show that subdistricts with relatively similar characteristics in terms of productivity, quantity, and number of farmers are grouped into the same cluster.

Table 10. List of Subdistricts for Each Livestock Sector Cluster

Commodities	Cluster	Subdistrict
Kattle	Cluster 1	Aralle, Mambi
	Cluster 2	Mamasa, Nosu, Pana, Rantebulahan Timur, Sesenapadang, Sumarorong, Tandukkalua, Tawalian
	Cluster 3	Balla, Bambang, Buntu Malangka, Mehalaan, Messawa, Tabang, Tabulahan
Buffalo	Cluster 1	Aralle
	Cluster 2	Balla, Bambang, Buntu Malangka Mamasa, Nosu, Pana, Sumarorong
	Cluster 3	Mambi, Mehalaan, Messawa, Rantebulahan Timur, Sesenapadang, Tabang, Tabulahan, Tandukkalua, Tawalian
Pig	Cluster 1	Aralle, Mehalaan, Pana, Rantebulahan Timur
	Cluster 2	Balla, Buntu Malangka, Mambi, Messawa, Nosu, Sesenapadang, Sumarorong, Tabang, Tabulahan, Tandukkalua, Tawalian
	Cluster 3	Bambang, Mamasa

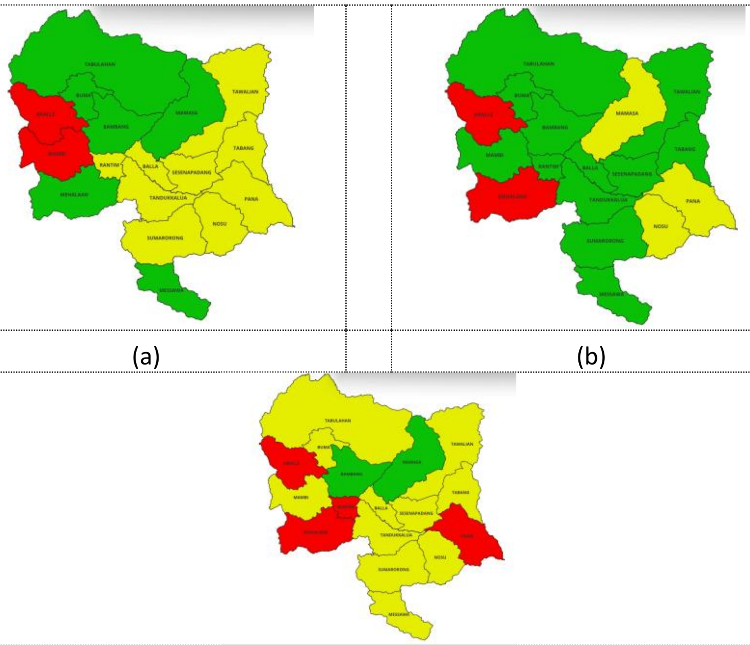
In viewing the cluster levels based on the analysis results, the centroid values for each cluster are obtained through Table 11 below.

Table 11. Centroid Values for Each Cluster in the Livestock Sector

Commodities	Variable	Cluster 1	Cluster 2	Cluster 3	Best Cluster
Kattle	Productivity	-0.411	0.8835	-0.8923	2
	Quantity	2.3479	-0.397	-0.2171	1
	Number of Farmers	2.3264	-0.6333	0.0591	1
Buffalo	Productivity	3.6338	-0.3546	-0.1279	1
	Quantity	-0.5054	0.8676	-0.6186	2
	Number of Farmers	-1.037	0.8936	-0.5798	2
Pig	Productivity	0.9395	-0.225	-0.6412	1
	Quantity	-0.6411	0.0496	1.0096	3
	Number of Farmers	-0.9899	0.0233	1.8516	3

The results of the analysis in Table 11 show the characteristics of clusters for different livestock sectors depending on the commodity. For cattle, cluster 2 excels in productivity (0.8835), while quantity (2.3479) and number of farmers (2.3264) are highest in cluster 1, indicating that cluster 2 is more effective in producing cattle per unit, while cluster 1 has the largest number of livestock and farmers. For buffalo, the highest productivity is recorded in cluster 1 (3.6338), while quantity (0.8676) and number of farmers (0.8936) are in cluster 2, indicating that cluster 1 produces buffalo with the best performance, while cluster 2 has greater potential for maintenance capacity. For pigs, cluster 1 has the highest productivity (0.9395), but the highest quantity (1.0096) and number of farmers (1.8516) are found in cluster 3, so cluster 3 shows greater potential for pig development even though productivity per unit is lower. Based on these centroid values, the cluster division in the livestock sector can be visualized through the graph in Figure 7 to show areas with low, medium, and high production performance.

Based on Figure 7, in general, cluster 1 stands out in terms of the number of farmers and the productivity of buffalo and pigs, cluster 2 excels in cattle productivity and buffalo maintenance capacity, while cluster 3 has the largest number of farmers and pigs. Based on these centroid values, the division of livestock sector clusters can be visualized through the graph in Figure 6 to show areas with low, medium, and high livestock performance.



(a) (b)

(c)

Figure 7. Clustering Results for Livestock Sector Productivity cattle (a), buffalo (b), and pigs (c).

Interpretation of Results

Based on the results of the analysis and discussion, the interpretation is as follows. In the food sector, rice and corn productivity between subdistricts is relatively stable but still shows variation. Mehalaan, East Rantebulahan, and Tabulahan subdistricts stand out as leading areas with the highest productivity (5.6 tons/ha), while Balla, Tabang, and Tandukkalua have the lowest productivity (5.0 tons/ha). Cluster analysis divides the region into three groups based on productivity, production, planted area, and number of farmers, showing that the determinants of productivity are uneven and that some subdistricts have higher cultivation efficiency than others.

In the plantation sector, the productivity of plantation commodities varies greatly between subdistricts. Nosu Subdistrict excels in Arabica coffee with the highest productivity (1,300 tons/ha), while Tawalian has the highest productivity of Robusta coffee (850 tons/ha). Tabulahan subdistrict stands out in patchouli production (50 tons), and Bambang stands out in plantation area (5,027 ha). Cluster analysis shows that Arabica and Robusta coffee have three distinct clusters based on productivity, production, planted area, and number of farmers, while cocoa and patchouli are relatively homogeneous and mostly fall into the high productivity cluster. This indicates the need for a specific development approach for coffee, whereas cocoa and patchouli can be developed using a general strategy.

In the livestock sector, productivity and quantity vary between subdistricts depending on the commodity. Cattle are superior in the subdistricts of Sumarorong (8,725), Pana (7,75), and Nosu (8.0), while buffalo have the highest productivity in Aralle (22). The highest pig populations are found in Aralle (7,832), Pana (7,415), and Tabulahan (7,151). Cluster analysis shows that cattle are superior in cluster 2, buffalo in cluster 1, and pigs in cluster 1 in terms of productivity, but the highest number of livestock and farmers for cattle and buffalo are in cluster 1, while for pigs it is in cluster 3. This indicates that cattle and buffalo development can use a general strategy, while pig development requires a different approach between subdistricts according to productivity clusters and maintenance capacity.

Overall, the three sectors show uneven productivity distribution across subdistricts, with some areas becoming centers of excellence for each commodity. Development strategies need to be tailored to the characteristics of each region and cluster in order to optimize efficiency and production output.

CONCLUSION

In conclusion, the spatial distribution of productivity across the food, plantation, and livestock sectors in Mamasa Regency is characterized by significant regional disparities and distinct commodity-driven patterns. Rather than reflecting a uniform level of development, the K-Means clustering results unveil unique subdistrict typologies that require differentiated policy interventions. In the food sector, subdistricts like Mehalaan, East Rantebulahan, and Tabulahan form a cluster characterized by high-efficiency rice and corn cultivation, suggesting their role as regional food security anchors. Conversely, the plantation sector is highly fragmented by specific commodities; Arabica coffee production creates a highly localized cluster in Nosu, whereas Robusta coffee dominates the Tawalian cluster. Meanwhile, cocoa and patchouli display a more homogeneous distribution across regions, making them suitable for broader, generalized structural support. In the livestock sector, cattle and buffalo production are relatively widespread, allowing for standardized regional programs, whereas pig farming is highly localized in Aralle, Pana, and Tabulahan, demanding cluster-specific biosecurity and marketing strategies.

Scientifically, this study proves that evaluating agricultural sub-sectors simultaneously through an integrated cluster approach prevents the misallocation of developmental resources. The regional government of Mamasa should move away from top-down, uniform agrarian policies. Instead, they must implement a decentralized, cluster-based development roadmap prioritizing infrastructural and technological intervention for the "Food Anchor" clusters, while optimizing supply chains and processing facilities tailored to the specific "Coffee" and "Livestock" cluster typologies.

ACKNOWLEDGEMENT

The author would like to express his deepest gratitude to various parties who have contributed to the completion of this research. Appreciation and gratitude are extended to Institut Teknologi Sepuluh Nopember (ITS) for providing academic support and adequate research facilities through the Research and Community Service Program (PPM); to the Ministry of Transmigration of the Republic of Indonesia, particularly the Mamasa Regency Transmigration Office, for providing access to data, technical support, and field facilities during the data collection and validation process; to the Mamasa Regency Agriculture and Food Security Agency, which provided secondary data and supporting information that was invaluable to the completeness of this research; to the respondents and informants in the 16 subdistricts of the research area, who were willing to give their time, share information, and provide valuable insights; and to research colleagues from the ITS Regional Planning and Development Laboratory, who provided constructive input during the analysis and writing process. The contributions of all these parties have been invaluable in ensuring the accuracy of the data, the depth of the analysis, and the quality of this paper. All individuals and institutions mentioned have given their consent to be included in this acknowledgments section.

DECLARATION

Author contribution

All authors contribute in the research and/or writing the paper, and approved the final manuscript.

Muhammad Zulfadhli	Formal analysis, Writing – Original draft, and Editing
Kesumaning Dyah Larasati	Investigation, Resources, Validation, Review, Funding acquisition and Editing
Rizky Fitria Ramadhanni	Methodology, Visualization, Writing - review and editing
Raynaldi Anggiat Samuel Siahaan	Software, Data curation, and Writing - review
Mochamad Fatih Romadlon	Investigation, Resources, and Validation. All authors discussed the results and contributed to the final manuscript.

Funding

This research did not receive any funding.

Conflict of interest

All authors declare that they have no competing interests.

Ethics declaration

We as authors acknowledge that this work has been written based on ethical research that conforms with the regulations of our institutions and that we have obtained the permission from the relevant institutes when collecting data. We support the Bulletin of Applied Mathematics and Mathematics Education (BAMME) in maintaining the high standards of personal conduct, practicing honesty in all our professional practices and endeavors.

The use of artificial intelligence

We do not use any generative AI tools to write any part of this paper.

Additional information

Not available.

REFERENCES

Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2020). *Statistics for business and economics*. Cengage Learning.

Arifin, B., Rahman, A., Prasetyo, D., & Lestari, S. (2021). Agricultural productivity disparities in Indonesia. *Journal of Indonesian Agriculture*, *15*(2), 45–60.

Chen, X., & Li, Y. (2022). Evidence-based agricultural policy making in developing regions. *Journal of Rural Development*, *45*(2), 112–128.

Fuglie, K., Heisey, P., King, J., & Schimmelpfennig, D. (2020). Agricultural productivity and economic growth. *Agricultural Economics*, *51*(1), 23–38.

Gupta, R., & Kumar, S. (2023). Spatial clustering for agricultural zoning: A case study of Northern India. *Agricultural Systems*, *195*, 103–115.

Han, J., Kamber, M., & Pei, J. (2022). *Data mining: Concepts and techniques*. Morgan Kaufmann.

Haryanto, A. E. P., & Zulfadhli, M. (2024). Grouping provinces based on Minimum Competency Assessment results to improve education quality in Indonesia. *Jurnal Nalar Pendidikan*, *12*(2), 155–163.

Ismail, F. A., Albarr, A. C., Zulfadhli, M., & Wulandar, S. P. (2024). Pengelompokan kabupaten di Jawa Tengah berdasarkan faktor yang memengaruhi Indeks Pembangunan Manusia tahun 2022 menggunakan analisis cluster. *Blantika: Multidisciplinary Journal*, *3*(1), 67–82.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning*. Springer.

Kumar, A., & Singh, R. (2022). Cluster-based agricultural development in India. *Indian Journal of Agricultural Economics*, *77*(3), 89–104.

Maryudi, A., Santoso, H., Putra, R., & Anwar, M. (2022). Location-specific approaches for agricultural development. *Indonesian Journal of Development Studies*, *10*(1), 34–49.

Porter, M. E. (2021). *Competitive advantage of regions*. Harvard Business Review Press.

Prasetyo, A., Nugroho, W., Hidayat, R., & Lestari, T. (2022). Regional disparities in agricultural productivity: Evidence from Indonesia. *Asian Journal of Agriculture and Development*, *19*(1), 45–62.

Saputra, R., Widodo, S., Hartono, A., & Lestari, D. (2021). Rice productivity clustering in East Java. *Journal of Agricultural Research*, *8*(2), 112–125.

Sari, D., & Wijaya, K. (2021). Cluster-based mapping for agricultural development planning. *Indonesian Journal of Agricultural Research*, *14*(3), 78–92.

Tan, P. N., Steinbach, M., & Kumar, V. (2021). *Introduction to data mining*. Pearson Education.

Todaro, M. P., & Smith, S. C. (2020). *Economic development*. Pearson Education.

Triola, M. F. (2021). *Elementary statistics*. Pearson Education.

Wahyuni, S., & Fitriani, N. (2022). Plantation commodity mapping in South Sulawesi. *Agro Economic Journal*, *13*(1), 67–82.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2023). *Data mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Zhao, L., Chen, M., Huang, X., & Li, Y. (2023). Advanced clustering methods for regional analysis. *Computers and Electronics in Agriculture*, *194*, 106–120.

Zulfadhli, M., Iswadi, A., Nurdin, S. Z., Muflikhah, S., Izzaty, N. R., Fatika, I., & Choirunnisa, T. F. (2025). Estimasi proporsi swing voter di kalangan pemilih pemula pada Pemilu 2024: Studi kasus Kota Salatiga dengan metode stratified two-stage cluster sampling. *Codeverse: Journal of Emerging Digital Realities*, *1*(1), 1–7.