Cluster Analysis of Determining the Location of Oil Palm Replanting in Mukomuko District

Emlan Fauzi, Linda Harta, Rahmat Oktavia, Sudarmansyah, Herlena Bidi Astuti, Jhon Firison, Andi Ishak, Shannora Yuliasari, Yudi Sastro

Assessment Institute for Agricultural Technology of Bengkulu
Email addres: emlan.81@gmail.com

ABSTRACT : Based on oil palm yield data in Bengkulu Province, it shows several regions with varying production numbers. For this reason, it is necessary to group the potential oil palm producing areas to find out which areas produce palm oil in large or small quantities. Production shared is usually done based on the name of the oil palm producing district. Therefore, a method is needed to facilitate the grouping of oil palm producing areas. With the clustering approach, the division of regional groups can be carried out based on planted area and production. Analysis K-Means makes it easier to group an area with the largest, medium and low yields of oil palm. From the analysis results, it can be seen that the priority areas for oil palm rejuvenation in Bengkulu Province are Mukomuko District with locations in Air Rami, Pondok Suguh, Teramang Jaya and Penarik Districts.

Keywords : Palm Oil, Cluster, K-Means Algorithm

INTRODUCTION

Oil palm is the main plantation crop in Bengkulu Province. This commodity has become the backbone of the economy for most farmer households in Bengkulu Province (Sukiyono *et al*, 2017). Most of the oil palm plantations in Bengkulu are smallholder plantations. The area of oil palm planted in Bengkulu Province reaches 208,627.11 ha (BPS Bengkulu Province, 2020). This area accounts for 50.36% of the plantation area in Bengkulu Province. Oil palm planted areas are not evenly distributed in all districts. The largest district is Mukomuko Regency.

Oil palm plantations in Mukomuko District began to develop since 2000, this shows that the age of oil palm trees in this district is relatively old and production tends to decline. Oil palm production over the age of 16 tends to decline (Wibowo and Junaedi, 2017; Pahan, 2008). For this reason,

it is necessary to make oil palm rejuvenation efforts in this district. The location of replanting must be in accordance with the proportion of the existing land area. For this reason, it is necessary to analyze the planting area and production based on district clusters even down to the sub-district level.

Determination of priority locations can be done using cluster analysis. Cluster analysis is an analysis that is used to group observations or variables into several groups of observations or variables whose numbers are more little based on characteristics (Wardhani, 2016; Setiawan, 2016; Yulianto and Hidayatullah, 2014). The benefit of cluster analysis is that it helps in developing a generalization method inductively, namely making general conclusions based on specific facts and describing the characteristics or characteristics of each group (Windarto, 2017; Hajar *et al*, 2020; Suhaeni, 2018 ; Alfina *et al*, 2012).

In previous studies, the k-means method has been has been widely appliedto classify an object / observation related to agriculture and non-agriculture, such as the Indonesian provincial cluster based on food production (Tendean and Purba, 2020), grouping areas based on the potential of agricultural products in The city of Cilegon (Sugiyani, 2016), increases the production of maize plants based on K-Means (Siregar, 2018), the quality grouping of palm kernels (Sirait, 2017) and the grouping of productive palm (Saragih *et al*, 2019). Therefore, the aim of this study is to analyze the location of replanting oil palm based on the planted area and production at the district and sub-district levels using the k-means cluster method.

MATERIALS AND METHODS

The data used in this study is secondary data, which was obtained from the Central Bureau of Statistics (BPS) Bengkulu Province in the form of production data and oil palm planted areas in all districts in Bengkulu Province and data in Mukomuko Regency in 2019.

Method of Analysis

The analytical method used in this research is the K-Means cluster analysis. Data processing was carried out with the help of SPSS 16.0 software, with the following stages:
1. Determine the number of clusters k
2. Allocating data into clusters randomly
3. Calculating the distance of each existing data to each cluster center in each cluster using the Euclidian Distance formula (Nugroho, 2008):

$$D(x_i y_j) = \sqrt{(x_{i1} y_{j1})^2 + (x_{i2} - y_{j2})^2 + \ldots + (x_{ip} - y_{jp})}$$

$$D(x_i y_j) = \sqrt{\sum_{i=1}^{p} (x_{i1} - y_{j1})^2}$$

where:
$D(x_i y_j)$ = The distance between the i-th data and the j-th cluster center

x_{i1} = The ith data on the 1st variable

y_{j1} = The center of the j cluster in the 1st variable

4. Group each data to the closest distance to its center.
5. Determine the position of the new Ckj cluster center by calculating the average value of the existing data in the same cluster with the formula (Rivani, 2010):

$$C_{kj} = \left(\frac{1}{n_k}\right) \sum_{j=1}^{p} d_{jt}$$

where:
C_{kj} = The center of the new ke-j cluster in the-1 variable

n_k = The number of object members in the cluster ke-k

d_{j1} = The data in the ke-k cluster is the ke-1 variable

6. If the cluster center does not change again, the cluster process is complete, or return to step 3 if there is still data moving clusters.

RESULTS AND DISCUSSION

Palm Oil Planted Area and Production in Bengkulu Province

The production and planted area of oil palm in Bengkulu Province are not evenly distributed as seen in Figure 1.
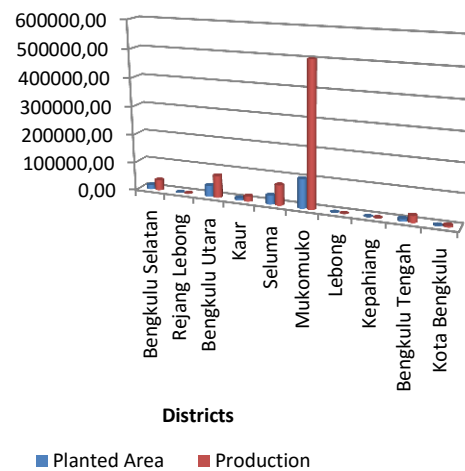

Figure 1. Planted Area and Production of Oil Palm in Bengkulu Province

Visually, Mukomuko Regency has a dominant planted area and oil palm production compared to other districts in Bengkulu Province. The planted area of oil palm in Mukomuko District is 102,822 hectares or contributes 49.29 percent of the planted area in Bengkulu Province. Meanwhile, the total oil palm production in this district was 502,345 tonnes or contributed 68.03 percent (Table 1).

Based on the planted area and production of oil palm in Bengkulu Province shown above, it can be divided into 3 clusters. Cluster 1 is Mukomuko District, cluster 2 is Bengkulu Utara and Seluma Districts, while cluster 3 is South Bengkulu, Rejang Lebong, Kaur, Lebong, Kepahiang, Central Bengkulu and Bengkulu City Regencies (Figure 2).

Table 1. Oil palm planted area and production in bengkulu province

| Districts	Planted Area	Percentage	Production	Percentage
Bengkulu Selatan	14930.37	7.16	36825.00	4.99
Rejang Lebong	731.00	0.35	889.00	0.12
Bengkulu Utara	38922.00	18.66	78413.00	10.62
Kaur	8574.00	4.11	18243.00	2.47
Seluma	31456.00	15.08	71687.00	9.71
Mukomuko	102822.00	49.29	502345.00	68.03
Lebong	259.65	0.12	72.00	0.01
Kepahiang	113.00	0.05	230.00	0.03
Bengkulu Tengah	9001.00	4.31	23500.00	3.18
Kota Bengkulu	1818.09	0.87	6173.00	0.84
Jumlah	208627.11	100.00	738377.00	100.00

Source: BPS Bengkulu Province, 2020.
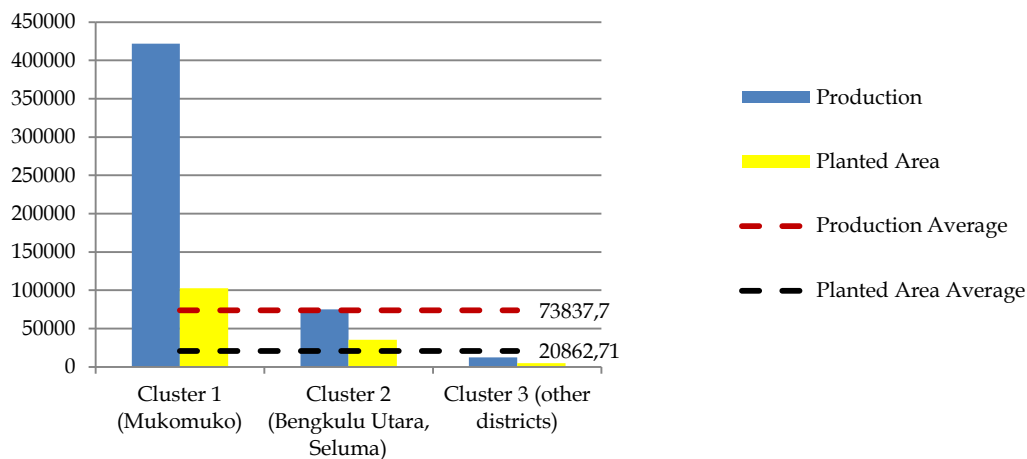
Figure 2. Oil Palm Clusters Based on Planted Area and Production in Bengkulu Province

Cluster 1 and cluster 2 are districts that have planted areas and oil palm production above the provincial average. While cluster 3 is the districts that have planted areas and oil palm production below the provincial average based on the normal distribution table (Table z) as presented in Table 2.

In cluster 2, the planting and production area of North Bengkulu Regency is higher than that of Seluma Regency. Meanwhile, in cluster 3, the area of planting and production in South Bengkulu Regency is higher when compared to other districts. The significant difference between cluster 1, cluster 2 and cluster 3 can be seen from the ANOVA test results (Table 3).

Based on the ANOVA test results table, it can be seen that the significant figures for all variables of planting area and oil palm

production are less than 0.05. This shows that there are significant differences between cluster 1, cluster 2 and cluster 3, which are associated with all of these variables. The greater the F value, the greater the difference between clusters for the variable planting area and production. In other words, Mukomuko District as cluster 1 has a higher planted area and production compared to cluster 2 (North Bengkulu and Seluma) and cluster 3 (other districts). Furthermore, the number of planted and oil palm production in cluster 2 is also significantly higher than in cluster 3. This suggests that the cluster division is correct.

Table 2. Value Z Score of Planted Area and Palm Oil Production Based on Clusters.

Clusters	Districts	Planted Area		Production	
		Z Score	Average	Z Score	Averge
1	Mukomuko	2.570	2.576	2,796	2.795
2	Bengkulu Utara	0.568	0.450	0.030	0.008
	Seluma	0.333		-0.014	
3	Bengkulu Selatan	-0.19	-0.498	-0.242	-0.402
	Rejang Lebong	-0.63		-0.476	
	Kaur	-0.39		-0.363	
	Lebong	-0.65		-0.481	
	Kepahiang	-0.65		-0.480	
	Bengkulu Tengah	-0.37		-0.328	
	Kota Bengkulu	-0.60		-0.442	

1) a negative number (-) means the data is average below the total.
2) a positive number (+) means above the total average.

Table 3.The results of the significance test for cluster differentiating variables.

Uraian	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Zscore (Luas_Tanam)	4.386	2	.032	7	135.099	.000
Zscore (Produksi)	4.473	2	.008	7	589.454	.000

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

Oil Palm Planted Area and Production in Mukomuko District

Mukomuko Regency with its dominant production and planting area in Bengkulu Province can form its own cluster. Oil palm production and planted area in Mukomuko District are relatively evenly distributed in each sub-district, as shown in Figure 3.
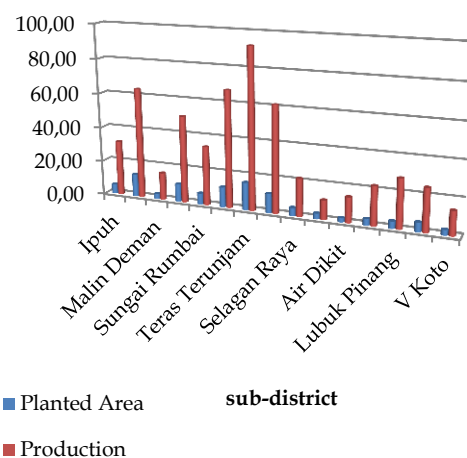

Figure 3. Planted Area and Production of Oil Palm in Mukomuko District

Teras Terunjam sub-district has relatively high planted area and palm oil production compared to other sub-districts in Mukomuko District. The planted area of oil palm in Teras Terunjam District is 15.89 hectares or contributes 15.48 percent of the planted area in Mukomuko District. Meanwhile, the total oil palm production in this sub-district was 92.41 tonnes or contributed 16.68 percent (Table 4).

Based on the planted area and production of oil palm in Mukomuko District shown above, it can be divided into 3 clusters. Cluster 1 is Teras Terunjam District, cluster 2 is Air Rami, Pondok Suguh, Teramang Jaya, and Penarik Districts. Meanwhile, cluster 3 is Ipuh, Malin Deman, Sungai Rumbai, Selagan Raya, Mukomuko, Air Dikit, XIV Koto, Lubuk Pinang, Air Munjunto and V Koto (Figure 4)

Table 4. Oil Palm Planted Area and Production in Mukomuko District

Sub-District	Planted Area	Percentage	Production	Percentage
Ipuh	5.51	5.37	31.73	5.73
Air Rami	12.95	12.62	63.86	11.53
Malin Deman	2.73	2.66	15.76	2.85
Pondok Suguh	10.17	9.91	50.28	9.08
Sungai Rumbai	6.36	6.20	34.03	6.14
Teramang Jaya	11.56	11.26	67.39	12.17
Teras Terunjam	15.89	15.48	92.41	16.68
Penarik	10.99	10.71	61.55	11.11
Selagan Raya	4.64	4.52	21.73	3.92
Kota Mukomuko	3.19	3.11	11.18	2.02
Air Dikit	2.35	2.29	14.57	2.63
XIV Koto	3.63	3.54	22.48	4.06
Lubuk Pinang	4.16	4.05	28.25	5.10
Air Munjunto	5.33	5.19	24.71	4.46
V Koto	3.19	3.11	13.98	2.52
Total	102.65	100.00	553.91	100.00

Source: Mukomuko in Numbers, 2020

Cluster 1 and cluster 2 are sub-districts that have planted areas and oil palm production above the district average. Whereas cluster 3 is a sub-district that has planted area and oil palm production below the district average based on the normal distribution table (Table z) as presented in Table 5.
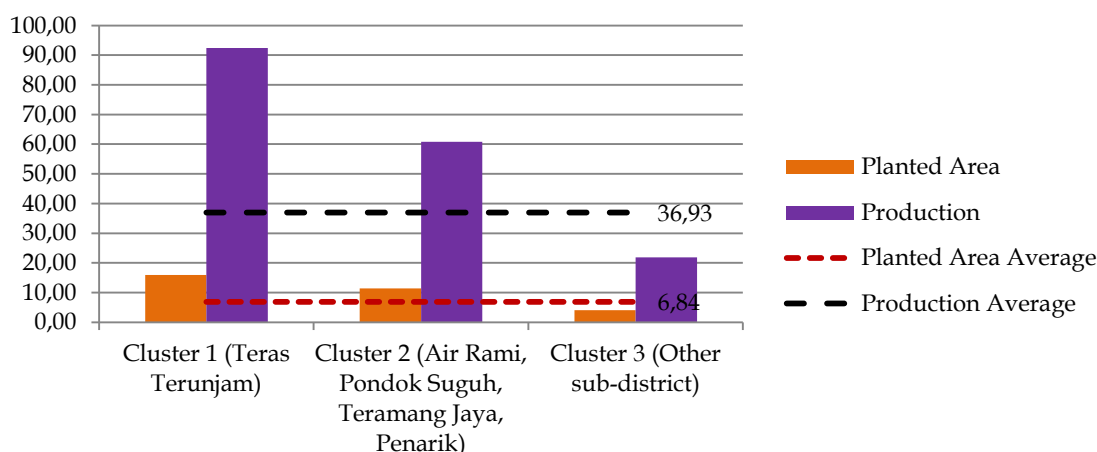

Figure 4. Oil Palm Clusters Based on Planted Area and Production in Mukomuko District

Table 5. Value Z Score of Planted Area and Palm Oil Production Based on Clusters.

Cluster	Sub-Districts	planted areas		Production	
		Z Score	Average	Z Score	Average
1	Teras Terunjam	2,098	2,098	2.271	2.271
2	Air Rami	1,417	1,061	1.103	0.976
	Teramang Jaya	1,094		1.247	
	Penarik	0,962		1.008	
	Pondok Suguh	0,772		0.546	
3	Ipuh	-0,309	-0.634	-0.212	-0.617
	Malin Deman	-0,954		-0.866	
	Sungai Rumbai	-0,112		-0.118	
	Selagan Raya	-0,511		-0.622	
	Kota Muko-muko	-0,847		-1.054	
	Air Dikit	-1,042		-0.915	
	XIV Koto	-0,745		-0.591	
	Lubuk Pinang	-0,622		-0.355	
	Air Munjunto	-0,351		-0.500	
	V Koto	-0,847		-0.939	

1) a negative number (-) means the data is average below the total.
2) a positive number (+) means above the total average.

Based on Table 5 above, it shows that Cluster 2 (Air Rami, Pondok Suguh, Teramang Jaya and Penarik Districts) is the main priority area for replanting oil palm because this area has a relatively high planting area but relatively low production. It is suspected that the age of oil palm trees in this area is relatively old. The planted area of oil palm in Air Rami Subdistrict is wider when compared to Teramang Jaya, Penarik and Pondok Suguh Districts, but in terms of production, Teramang Jaya Districts are relatively higher than those of Air Rami, Penarik and Pondok Suguh Districts.

Table 6. The results of the significance test for cluster differentiating variables.

Uraian	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Zscore (Luas_Tanam)	6.466	2	.089	12	72.718	.000
Zscore (Produksi)	6.392	2	.101	12	63.092	.000

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

In cluster 3, the planted area and production in Sungai Rumbai District are higher when compared to other sub-districts. The significant difference between cluster 1, cluster 2 and cluster 3 can be seen from the ANOVA test results (Table 6).

Anova test results showed a significant value for the variable planting area and oil palm production is less than 0.05. This shows that there are significant differences between cluster 1, cluster 2 and cluster 3, which are associated with all of these variables. The greater the F value, the greater the difference between clusters for the variable planting area and production. In other words, Teras Terunjam District as cluster 1 has a higher planting and production area than cluster 2 (Air Rami, Pondok Suguh, Teramang Jaya and Penarik) and cluster 3 (other sub-districts).

Furthermore, the planted area and oil palm production in cluster 2 are also significantly higher than in cluster 3. This suggests that the cluster division is appropriate.

CONCLUSION

The priority areas for oil palm rejuvenation in Bengkulu Province are Mukomuko District with locations in Air Rami, Pondok Suguh, Teramang Jaya and Penarik Districts.

SUGGESTION

Result of this research is a preliminary study that can be verified in the field in determining the location of CPCL replanting oil palm in Mukomuko District, Bengkulu Province.

REFERENCES

Alfina T., B Santosa, A R Barakbah. 2012, Analisa perbandingan metode hierarchical clustering, k means dan gabungan keduanya dalam cluster data (studi kasus : problem kerja praktek jurusan teknik industri its). Jurnal Teknik ITS. 1(1). 521-525. DOI: 10.12962/j23373539.v1i1.1794.

BPS. 2020. Provinsi Bengkulu Dalam Angka 2020. Bengkulu. Retrieved from https://bengkulu. bps.go.id/index.php/publikasi.

BPS. 2020. Kabupaten Mukomuko Dalam Angka 2020. Mukomuko. Retrieved from https://mukomukokab.bps.go.id. index.php/publikasi.

Hajar S., A. A. Novany, A. P. Windarto, A. Wanto, E. Irawan. 2020. Penerapan K-Means Clustering Pada Ekspor Minyak Kelapa Sawit Menurut Negara Tujuan. Seminar Nasional Teknologi Komputer & Sains (SAINTEKS). Hal.314-318. https://prosiding.seminar-id.com/index.php/sainteks.

Nugroho, S. 2008. Statistika Multivariat terapan, edisi pertama. UNIB Press, Universitas Bengkulu.

Pahan, I. 2008. Panduan Lengkap Kelapa Sawit: Manajemen Agribisnis dari Hulu hingga Hilir. Jakarta (ID): Penebar Swadaya.

Rivani E. 2010. *Aplikasi K-means Cluster untuk Pengelompokkan Provinsi Berdasarkan Produksi Padi, Jagung, Kedelai, dan Kacang Hijau Tahun 2009*. Jurnal Mat Stat. 10 (02). *122-134*. https://research.binus.ac.id/public ation/BF93AEE7-7518-4FF4-9ACD-33C708FD7E0C/aplikasi-k-means-cluster-untuk-pengelompokkan-provinsi-berdasarkan-produksi-padi-jagung-kedelai-dan-kacang-hijau-tahun-2009/

Saragih, R. A. D., M. Safii, H. S. Tambunan. 2019. Penerapan Metode K-Means Clustering Untuk Mengelompokkan Kelapa Sawit Produktif. Prosiding Seminar Nasional Matematika dan Terapan. Volume 1, pp: 362-370. https://bulletin.indomsacehsumut. org/index.php/simantap/article/v iew/95.

Setiawan R. 2016. Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru (Studi Kasus : Politeknik Lp3i Jakarta). Jurnal Lentera ICT 3(1). 76 –92.

Sirait N. 2017. Implementasi K-Means Clustering pada Pengelompokan Mutu Biji Sawit (Studi Kasus: PT. Multimas Nabati Asahan). Jurnal Pelita Informatika. 6(2). Hal: 170-174.

Siregar AM. (2018). Implementasi Algoritma Clustering dengan Singular Vector Decomposition untuk Menunjang Keputusan dalam Meningkatkan Produksi Tanaman Jagung. Institut Pertanian Bogor. pp. 71-78. http://semilkom.apps.cs.ipb.ac.id/

Sugiyani Y. 2016. Pengelompokan Wilayah Berdasarkan Potensi Hasil Pertanian Menggunakan Algoritma

K-means Di Kota Cilegon. Jurnal ProTekInfo. 3(2). 60-67.

Suhaeni C., A. Kurnia, Ristiyanti. 2018. Perbandingan Hasil Pengelompokan menggunakan Analisis Cluster Berhirarki, K-Means Cluster, dan Cluster Ensemble (Studi Kasus Data Indikator Pelayanan Kesehatan Ibu Hamil). Jurnal Media Infotama 14(1). 31-38.

Sukiyono, K., Cahyadinata, I., Purwoko, A., Widiono, S., Sumartono, E., Asriani, N. N., & Mulyasari, G. (2017). Assessing Smallholder Household Vulnerability to Price Volatility of Palm Fresh Fruit Bunch in Bengkulu Province. International Journal of Applied Business and Economic Research, 15(3): 1 – 15.

Tendean T dan W Purba. 2020. Analisis Cluster Provinsi Indonesia Berdasarkan Produksi Bahan Pangan Menggunakan Algoritma K-Means. Jurnal Sains dan Teknologi. 1(2). 5-11.

Wardhani A K. 2016. Implementasi Algoritma K-Means Untuk Pengelompokkan Penyakit Pasien Pada Puskesmas Kajen Pekalongan. Jurnal Transformatika.14 (1). 30–37.

Wibowo, W. H dan A. Junaedi. 2017. Replanting of Palm Oil (Elaeis guineensis Jacq.) in Seruyan Estate, Minamas Plantation Group, Seruyan, Central Borneo. Buletin Agrohorti 5 (1) : 107 – 116.

Windarto A P. 2017. Penerapan Datamining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering Method. Jurnal Teknologi Informasi. 16(4). pp. 348–357. doi: https://doi.org/10.33633/tc.v16i4.1447.

Yulianto S., K. H. Hidayatullah. 2014. Analisis klaster untuk pengelompokan kabupaten/kota di Provinsi Jawa Tengah Berdasarkan indikator kesejahteraan rakyat. Jurnal Statistika. 2 (1). 56-63.