ISSN 2502-5872 (Print)
ISSN 2622-3627 (Elektronik)

Volume 9 Nomor 1, February 2024, 1-12

ENHANCING ORCHID INVENTORY MANAGEMENT WITH K-MEANS CLUSTERING: A CASE STUDY IN SALES OPTIMIZATION

Dwi Ramadhani[1], Tabah Heri Setiawan[2*], Choirul Basir[3], Dewi Purnama Sari[4]
[1,2,3,4]Department of Mathematics, Universitas Pamulang, Province of Banten, Indonesia
*Correspondence: tabah.ibnubara@gmail.com

ABSTRACT

In the ornamental industry, particularly in orchid sales, manual inventory management has presented significant challenges. These challenges include excessive stock accumulation, financial pressure due to overstocks, and limited land resources to store overstocks. Utilizing information technology, data mining, specifically the K-Means Clustering algorithm, emerges as a solution. This research aims to optimize orchid inventory management by grouping products into three clusters: highly sellable, sellable, and less sellable. Historical sales data was used as the main dataset, considering parameters such as sales volume, demand patterns, and seasonality. The K-Means Clustering algorithm categorized the orchid products, with 97 products in the highly sellable cluster, 68 products in the sellable cluster, and 22 products in the less sellable cluster. This analysis offers insights. Orchids that are highly sellable require higher stock levels to meet demand. Sellable varieties require standard stock levels, while less sellable types should maintain lower stock levels to avoid overstocking. In conclusion, using k-means clustering for orchid inventory management can optimize sales strategies. Orchids in high demand receive adequate stock, ensuring customer satisfaction. Effective stock allocation improves financial stability by overcoming the problem of overstock in the ornamental industry.

Keywords: Inventory Management, K-Means Clustering, Orchid

How to Cite: Ramadhani, D., Setiawan, T. H., Basir, C., & Sari, D. P. (2024). Enhancing Orchid Inventory Management With K-Means Clustering: A Case Study In Sales Optimization. *Mathline: Jurnal Matematika dan Pendidikan Matematika*, *9*(1), 1-12. http://doi.org/10.31943/mathline.v9i1.529

PRELIMINARY

Indonesia is an archipelago blessed with abundant biodiversity, both in terms of flora and fauna (Von Rintelen et al., 2017). Its tropical climate, characterized by distinct wet and dry seasons, provides ideal conditions with sufficient rainfall, temperature, and humidity levels to support a wide variety of plant and animal species. Indonesia's landscapes are adorned with lush greenery, and one of the standout representatives of this tropical abundance is the orchid (Irvani & Susandarini, 2022).

Orchids, known for their aesthetic appeal, hold a special place in the hearts of horticulture enthusiasts, both locally and internationally. Their unique shapes and vibrant colors make them highly sought-after ornamental plant (Prayoga et al., 2022). Orchids are highly popular in various forms, including cut flowers and potted plants (Hinsley et al., 2015). They are valued for their beauty and unique characteristics, making them sought after in the horticultural and floral industries (Auvira et al., 2021).

South Tangerang is one of the cities in Banten that cultivates orchids and has made orchids as city icon (Ardiansah, 2023). The multitude of orchid varieties available, coupled with increasing customer demands, pose challenges in maintaining adequate stock levels. Stockouts and excess are recurring issues that hinder stock management at orchid cultivation entrepreneur. As a result, businesses face stock shortages, financial inefficiencies, and storage space constraints.

Inventory management is considered an essential function of an inventory management system (Munyaka & Yadavalli, 2022). Inventory management, also known as materials management, is the coordination and organization of the creative flow in a mechanized project by organizing, acquiring, storing and distributing the right materials, in the right quality, in the right quantity, at the right time and in the right place (Ajibade et al., 2022). The goal of inventory management is to add value in the form of customer satisfaction, which is a useful indicator of organizational performance, while maintaining a certain level of stock at a low cost.

The advent of advanced information technology solutions offers numerous data processing techniques to address these challenges through algorithm approach (Ridwan & Rakhmawati, 2023). Data mining, specifically the k-means clustering algorithm, stands out as a promising solution (Patel, 2019; Wu, 2012). In many researches k-means clustering is utilized to group sales data, help store owners identify buying patterns and manage inventory more effectively (Adani et al., 2022). Moreover, k-means clustering is also widely used in recognizing popular (Aggarwal & Aggarwal, 2012; Thogarcheti et al., 2020) and less popular items to avoid accumulation (Metisen & Sari, 2015).

The utilization of the k-means clustering method has become one of the most commonly applied approaches in the field of inventory management. Usually, this method is used to cluster inanimate objects in order to understand demand patterns, optimize storage, and plan stock more efficiently (Goncalves et al., 2021; Liu et al., 2019; Witthaut & Kalbe, 2022). However, this research extends the concept by applying K-means clustering to different objects, namely plants. This creates new opportunities to apply the method in different contexts, such as agribusiness or agricultural research, to cluster crops based on certain characteristics, facilitating more efficient and effective crop management.

METHODS

The type of data used is secondary data. Then this study uses quantitative data expressed by numbers to show the value of the amount of the variable it represents. The population used is the result of an annual recapitulation in 2022 using a sample of 187, which is the name of the vanda-type orchid plant. Can be seen in Table 1.

Table 1. Sales Data of Orchid Plants in 2022

Code	Orchid Plant Items	Bimonthly Sales					
		B1	B2	B3	B4	B5	B6
Y_1	Vanda Denisoniana x Vanda Denisoniana	36	34	43	33	45	37
Y_2	Aerides Lawrenceae x Vanda blue	23	36	35	22	34	24
Y_3	Ascda Bigness x Kobfah	32	37	31	37	39	33
Y_4	Ascda. Bangkuntien Gold	17	22	18	21	15	25
Y_5	Ascda. Crown Fox Red Gems)	43	24	20	36	31	36
Y_6	Ascda. Crownfox Magic	30	33	22	40	25	41
Y_7	Ascda. Fuch Sunglow	37	43	29	25	32	33
Y_8	Ascda. Kulvadee Fragrant Black spot	58	49	43	47	49	54
Y_9	Ascda. Kulvadee Fragrant Red spot	25	27	31	21	34	24
Y_{10}	Ascda. Princess Mikasa Blue	23	16	21	25	19	20
Y_{11}	Ascda. Princess Mikasa White	44	28	22	31	34	24
…	…	…	…	…	…	…	…
Y_{184}	V.Yellow Butterfly #497	21	34	30	29	28	28
Y_{185}	V.Yellow Butterfly #597	27	35	20	32	28	22
Y_{186}	Vasco. Koishi Ikai Pink	31	30	35	65	48	37
Y_{187}	Vasco. Ploenplit Prize	24	31	35	31	27	35

In this study, the k-means data mining algorithm is used, which aims to optimize orchid clustering using k-means clustering in order to distinguish datasets with similar characteristics from other datasets. The algorithm was chosen because the k-means algorithm is an algorithm designed for clustering that can be done quickly and efficiently on data that is not overly large. The k-means algorithm uses a process of computational iteration to obtain the resulting pattern of the algorithm.

The k-means method aims to divide data into many groups that have certain characteristics in some groups and different in others (Wanto, 2020). The following will present a flowchart of the k-means clustering algorithm:

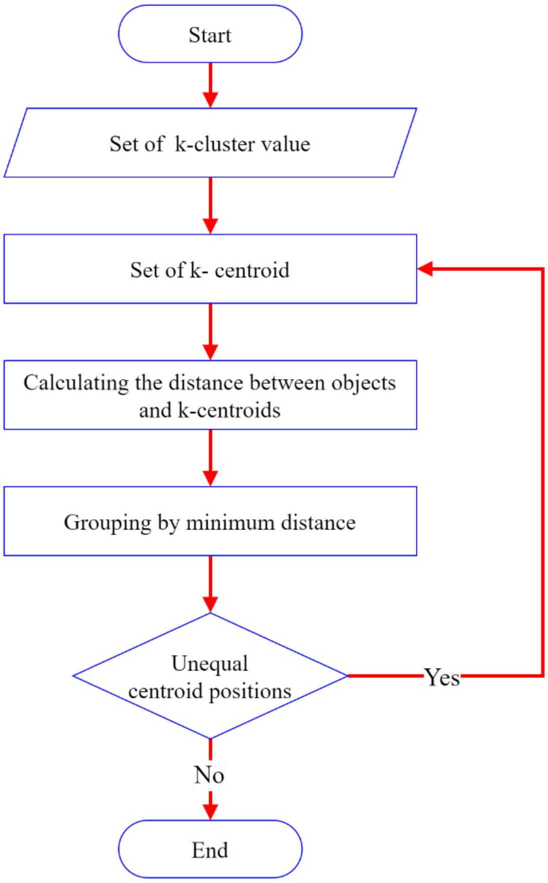

Figure 1. Flowchart of K-Means Clustering Algorithm

The steps of k-means clustering algorithm are explained as follows (Wanto, 2020):

1. Set of k-cluster value.
2. The initial step's centroid is chosen at random. The formula used in the iterative phase is Formula (1) below:

$$v_{ij} = \left(\frac{1}{N_i}\right)\sum_{k=0}^{N_i} x_{kj} \tag{1}$$

Description:

v_{ij} = The j-th variable's average centroid for the i-th cluster

N_i = The number of individuals in the i-th cluster

i, k = index for clusters

j = index of the variable

x_{kj} = The cluster's j-th variable's k-th data value.

3. Calculate the distance between objects and the cluster center.

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \tag{2}$$

Description:

De = *Euclidean Distance*

i = Number of objects

(x, y) = Object Coordinates

(s, t) = Centroid Coordinates

4. After the calculation is complete, assigns all calculated data and objects to the closest cluster. The closest distance determines which cluster the object or data belongs to.
5. Recalculate all data using the latest cluster centers. The clustering step is complete when the cluster centers no longer change; otherwise, return to step 3 until the cluster centers do not change.

Calculations carried out using the help of Google Collaboratory. Google Collaboratory is a cloud-based platform that allows users to write, store, and share executable documents (Iskandar et al., 2022). It provides a collaborative environment where multiple users can simultaneously edit a document (Musen, 2015). Google Colab supports various programming languages, including Python, and provides access to powerful computing resources such as GPUs for model training (Sampathila et al., 2022). It can be accessed through a web browser and does not require any installation or setup.

The following are the steps to get clustering results using the k-means clustering algorithm method.

1. Inserting Libraries

   To run the k-means clustering algorithm in Google Collaboratory, libraries such as pandas, numpy, and matplotlib are needed.
2. Inserting Datasets

   The dataset used in this study is a *.csv file which is then imported using the pandas library and make sure that the output results do not experience errors.
3. View Variable Information on Data

   Ensure that the variables in the data, both the column name and the data type, are not wrong.
4. Determining the Variables to be Clustered

   The variables to be used for clustering are the values in columns B1, B2, B3, B4, B5, and B6. While the KODE column in this step will be deleted.
5. Changing Variables

In order for the values in the dataset to be used to calculate the K value, the variables that were previously in the form of frame data need to be changed.

6. Entering the K Value

   Based on the value of K, it has been determined that 3 and 3 clusters are requests from the owner of the Setu Orchid Garden, therefore the number of clusters to be used is 3 clusters.

7. Finding the centroid value

   After the data is randomly clustered, it needs to be allocated to the respective clusters using the centroid of each cluster.

8. Display Cluster Results

   Display the clustering result as an array.

9. Visualization of Clustering Results

   To visualize the clustering results using a 2-dimensional plot where each group is represented with a different color.

RESULT AND DISCUSSION

In this section, we present the results and insights derived from our application of k-means clustering in the context of orchid inventory management, with a specific focus on sales optimization.

1. Variables to be Categorized

The variables used for clustering are the values in columns B1, B2, B3, B4, B5, and B6. The CODE column has been removed.

Table 2. Variables to be Categorized

	B1	B2	B3	B4	B5	B6
0	36	34	43	33	45	37
1	23	36	35	22	34	24
2	32	37	31	37	39	33
3	17	22	18	21	15	25
4	43	24	20	36	31	36
...	...	...	...	...	...	...
184	27	35	20	32	28	22
185	31	30	35	65	48	37
186	24	31	35	31	27	35

2. Centroid Value

Determine the new centroid position v_{ij} by calculating the average at the centroid using Equations (1).

Determine the Centroid in Cluster 1 for point B1:

$$v_{B11} = (23 + 32 + 43 + 30 + 37 + 25 + 44 + 29 + 25 + 26 + 28 + 44 + 31 + 19 + 43 + 55 + 30 + 38 + 29 + 24 + 23 + 27 + 32 + 27 + 25 + 29 + 39 + 32 + 33 + 25 + 32 + 21 + 27 + 23 + 33 + 32 + 39 + 34 + 23 + 25 + 27 + 38 + 32 + 23 + 22 + 46 + 27 + 25 + 35 + 21 + 25 + 27 + 22 + 26 + 17 + 22 + 17 + 22 + 16 + 23 + 27 + 30 + 23 + 28 + 25 + 20 + 28 + 22 + 36 + 34 + 30 + 43 + 26 + 38 + 26 + 37 + 20 + 20 + 25 + 26 + 36 + 23 + 33 + 37 + 27 + 35 + 34 + 27 + 21 + 20 + 22 + 23 + 21 + 28 + 21 + 27 + 24)/97$$

$$v_{B11} = \frac{2772}{97}$$

$$v_{B11} = 28.58$$

To determine Cluster 1 in B2 to B6, it is done as how to calculate Cluster 1.

Determine the Centroid in Cluster 2 for point B1:

$$v_{B12} = (17 + 23 + 19 + 16 + 13 + 16 + 16 + 15 + 20 + 19 + 20 + 13 + 28 + 15 + 25 + 15 + 19 + 18 + 19 + 22 + 19 + 27 + 17 + 29 + 13 + 21 + 18 + 22 + 37 + 29 + 19 + 25 + 24 + 18 + 18 + 18 + 19 + 23 + 25 + 16 + 19 + 15 + 24 + 17 + 15 + 27 + 20 + 15 + 26 + 13 + 24 + 16 + 13 + 18 + 22 + 17 + 24 + 18 + 22 + 26 + 25 + 15 + 19 + 21 + 24 + 16 + 16 + 21)/68$$

$$v_{B12} = \frac{1353}{68}$$

$$v_{B12} = 19.90$$

To determine Cluster 2 in B2 to B6, it is done as how to calculate Cluster 2.

Determine the Centroid in Cluster 3 for point B1:

$$v_{B13} = (36 + 58 + 33 + 65 + 46 + 37 + 53 + 85 + 49 + 56 + 40 + 55 + 33 + 40 + 43 + 47 + 38 + 50 + 46 + 37 + 31 + 31)/22$$

$$v_{B13} = \frac{1009}{22}$$

$$v_{B13} = 45.86$$

To determine Cluster 3 in B2 to B6, it is done as how to calculate Cluster 3.

Determination of the centroid has been done using Equation (1), then the new centroid is obtained as shown in Table 3.

Table 3. Centroids

Cluster	B1	B2	B3	B4	B5	B6
1	28.58	29.37	29.16	30.13	28.47	30.25
2	19.90	21.53	19.19	20.51	19.94	20.04
3	45.86	37.91	37.50	44.64	41.91	40.55

The data below is the result of the centroid value using Google Collaboratory.

```
[[28.57731959 29.37113402 29.16494845 30.13402062 28.4742268  30.24742268]
 [19.89705882 21.52941176 19.19117647 20.51470588 19.94117647 20.04411765]
 [45.86363636 37.90909091 37.5        44.63636364 41.90909091 40.54545455]]
```

3. Calculate The Distance Between Objects and The Cluster Center

The calculation of the distance from the first data to the centroids using Equation (2) is as follows below:

Calculation of the first data to Cluster 1:

$$d = \sqrt{\begin{array}{l}(36-28.58)^2 + (34-29.37)^2 + (43-29.16)^2 + \\ (33-30.13)^2 + (45-28.47)^2 + (37-30.25)^2\end{array}}$$

$$d = 24.39$$

Calculation of the first data to Cluster 2:

$$d = \sqrt{\begin{array}{l}(36-19.90)^2 + (34-21.53)^2 + (43-19.19)^2 + \\ (33-20.51)^2 + (45-19.94)^2 + (37-20.04)^2\end{array}}$$

$$d = 45.31$$

Calculation of the first data to Cluster 3:

$$d = \sqrt{\begin{array}{l}(36-45.86)^2 + (34-37.91)^2 + (43-37.50)^2 + \\ (33-44.64)^2 + (45-41.91)^2 + (37-40.55)^2\end{array}}$$

$$d = 17.33$$

The counting of the first data has been completed from Cluster 1 to Cluster 3, so do the same counting until the 187th data. Aim to determine whether each data is in Cluster 1, Cluster 2 or Cluster 3.

4. Clustering Based on Minimum Distance

The results of the calculation using Equation (2), obtained group members of Cluster 1, Cluster 2, and Cluster 3:

***Cluster* 1**: { Y_{2}, Y_{3}, Y_{5}, Y_{6}, Y_{7}, Y_{9}, Y_{11}, Y_{12}, Y_{13}, Y_{14}, Y_{15}, Y_{18}, Y_{21}, Y_{22}, Y_{23}, Y_{24}, Y_{25}, Y_{27}, Y_{28}, Y_{29}, Y_{30}, Y_{32}, Y_{34}, Y_{36}, Y_{37}, Y_{40}, Y_{44}, Y_{46}, Y_{47}, Y_{49}, Y_{52}, Y_{53}, Y_{55}, Y_{56}, Y_{57}, Y_{59}, Y_{63}, Y_{67}, Y_{68}, Y_{72}, Y_{76}, Y_{79}, Y_{83}, Y_{85}, Y_{86}, Y_{87}, Y_{89}, Y_{91}, Y_{93}, Y_{106}, Y_{107}, Y_{108}, Y_{109}, Y_{111}, Y_{117}, Y_{118}, Y_{120}, Y_{121}, Y_{124}, Y_{126}, Y_{127}, Y_{129}, Y_{130}, Y_{131}, Y_{132}, Y_{133}, Y_{134}, Y_{135}, Y_{136}, Y_{140}, Y_{141}, Y_{142}, Y_{143}, Y_{150}, Y_{151}, Y_{154}, Y_{156}, Y_{158}, Y_{159}, Y_{160}, Y_{161}, Y_{162}, Y_{164}, Y_{167}, Y_{168}, Y_{170}, Y_{171}, Y_{173}, Y_{178}, Y_{179}, Y_{180}, Y_{181}, Y_{182}, Y_{183}, Y_{184}, Y_{185}, Y_{187}}

***Cluster* 2**: { Y_{4}, Y_{10}, Y_{16}, Y_{17}, Y_{38}, Y_{39}, Y_{41}, Y_{42}, Y_{45}, Y_{50}, Y_{51}, Y_{54}, Y_{58}, Y_{62}, Y_{65}, Y_{70}, Y_{71}, Y_{74}, Y_{75}, Y_{78}, Y_{81}, Y_{82}, Y_{84}, Y_{88}, Y_{94}, Y_{95}, Y_{96}, Y_{97}, Y_{98}, Y_{99}, Y_{100}, Y_{101}, Y_{102}, Y_{103}, Y_{104}, Y_{105}, Y_{110}, Y_{112}, Y_{113}, Y_{114}, Y_{115}, Y_{116}, Y_{119}, Y_{122}, Y_{123}, Y_{125}, Y_{128}, Y_{137}, Y_{138}, Y_{139}, Y_{144}, Y_{145}, Y_{146}, Y_{147}, Y_{148}, Y_{149}, Y_{152}, Y_{153}, Y_{155}, Y_{157}, Y_{163}, Y_{165}, Y_{166}, Y_{169}, Y_{172}, Y_{175}, Y_{176}, Y_{177}}

***Cluster* 3**: {Y_{1}, Y_{8}, Y_{19}, Y_{20}, Y_{26}, Y_{31}, Y_{33}, Y_{35}, Y_{43}, Y_{48}, Y_{60}, Y_{61}, Y_{64}, Y_{66}, Y_{69}, Y_{73}, Y_{77}, Y_{80}, Y_{90}, Y_{92}, Y_{174}, Y_{186}}

Below is the result of clustering as an array using Google Colab:

array([2, 0, 0, 1, 0, 0, 0, 2, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 2, 2, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 2,
0, 2, 0, 2, 0, 0, 1, 1, 0, 1, 1, 2, 0, 1, 0, 0, 2, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 0, 2, 2,
1, 0, 2, 1, 2, 0, 0, 2, 1, 1, 0, 2, 1, 1, 0, 2, 1, 0, 2, 1, 1, 0, 1, 0, 0, 0, 1, 0, 2, 0,
2, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 0, 1, 0, 0,
1, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0,
1, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 2, 1, 1, 1, 0, 0, 0, 0,
0, 0, 0, 0, 2, 0], dtype=int32)

5. Results of Clustering Visualized

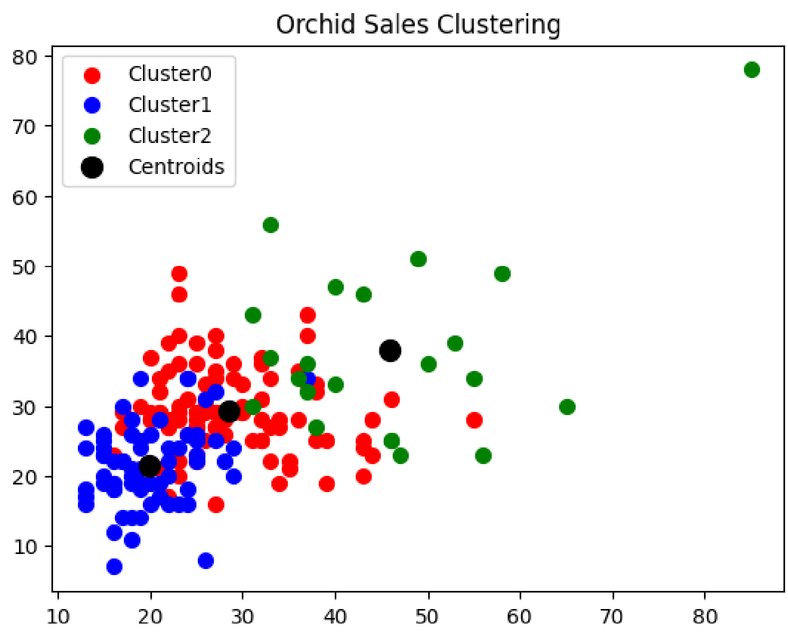

Figure 2. Results of Clustering Visualized

The k-means clustering divides the dataset into three clusters, namely clusters with highly sellable orchid, clusters with sellable orchid, and clusters with less sellable orchid, and calculates the distance between each cluster using the Euclidean distance.

In the highly sellable orchid sales cluster, 97 orchid plants are obtained consisting of Aerides Lawrenceae x Vanda Blue, Ascda Bigness x Kobfah, Ascda. Crown Fox Red Gems, and others. In the sellable orchid sales cluster, 68 orchid plants are obtained which consists of Ascda. Bangkuntien Gold, Ascda. Princess Mikasa Blue, Renanthera Renaworld x Nopporn Gold, and others. In the less sellable orchid sales cluster, 22 orchid plants are obtained consisting of Vanda Denisoniana x Vanda Denisoniana, Ascda. Kulvadee Fragrant Black spot, V. (Arunsri x Annet Jone), and others.

This research can find out which orchid plants are highly sellable, sellable, and less sellable. So that the goods in the garden do not experience accumulation and optimize to provide stock items for sales.

CONCLUSION

In this research, we have identified a significant number of advantages to applying K-means clustering to orchid inventory management. One of the key findings is that this method can change our view of how to manage live plant inventory. With intelligent clustering, we are able to better understand orchid characteristics and customer demand patterns. This is not just about stock efficiency but also about improving the knowledge underlying business decisions. By understanding the advantages of k-means clustering in this context, we open up new opportunities to apply it in various sectors involving inventory, from agriculture to e-commerce.

However, this study also highlights interesting future research opportunities. K-means clustering, as one of the powerful data mining techniques, can be used in various data cases, not only in orkid inventory management. The next challenge is to dig deeper into the use of this method in different contexts, such as agriculture, plantations, and other industries. Future research can also combine k-means clustering with other methods or develop more sophisticated clustering approaches. In this way, we will better understand the potential of this method for optimizing inventory management and decision-making in various sectors, creating abundant opportunities for further research and innovation.

REFERENCES

Adani, N. F., Boy, A. F., & Syahputra, R. (2022). Implementasi Data Mining Untuk Pengelompokan Data Penjualan Berdasarkan Pola Pembelian Menggunakan Algoritma K-Means Clustering Pada Toko Syihan. *Jurnal Cyber Tech*, *2*(5). http://ojs.trigunadharma.ac.id/index.php/jct/article/view/4648

Aggarwal, N., & Aggarwal, K. (2012). *An Improved K-means Clustering Algorithm For Data Mining* (1. Aufl). Lambert Academic Publishing.

Ajibade, T. . A., Odunayo, J., Osunusi, A. K., & Aguguom, T. A. (2022). Markov Chain and Inventory Management of Manufacturing Companies in Nigeria. *European Journal of Accounting, Auditing and Finance Research*, *10*(11), 21–35. https://doi.org/10.37745/ejaafr.2013/vol10n112135

Ardiansah, R. (2023, March 3). *Anggrek Vanda Douglas, Bunga Cantik Nan Mempesona Ikon Kota Tangerang Selatan*. https://www.goodnewsfromindonesia.id/2023/03/03/anggrek-vanda-douglas

Auvira, F. D. P., Mercuriani, I. S., & Aloysius, S. (2021). *Genetic Variability Analysis of Terrestrial Spathoglottis plicata Orchid Variants Based on RAPD Marker*. *528*, 70–75. https://doi.org/10.2991/assehr.k.210305.011

Goncalves, J. N. C., Cortez, P., & Carvalho, M. S. (2021). K-means clustering combined with principal component analysis for material profiling in automotive supply chains. *European J. of Industrial Engineering*, *15*(2), 273–294. https://doi.org/10.1504/EJIE.2021.114009

Hinsley, A., Verissimo, D., & Roberts, D. L. (2015). Heterogeneity in consumer preferences for orchids in international trade and the potential for the use of market research methods to study demand for wildlife. *Biological Conservation*, *190*, 80–86. https://doi.org/10.1016/j.biocon.2015.05.010

Irvani, M. D., & Susandarini, R. (2022). Keanekaragaman spesies anggrek di jalur pendakian Cemara Kandang, Gunung Lawu, Jawa Tengah. *Jurnal Biologi Udayana*, *26*(2), 175–185. https://doi.org/10.24843/JBIOUNUD.2022.v26.i02.p03

Iskandar, D. M., Yel, M. B., & Sitohang, A. (2022). Detection of The Deaf Signal Language Using The Single Shot Detection (SSD) Method. *Journal of Applied Engineering and Technological Science (JAETS)*, *4*(1), 215–222. https://doi.org/10.37385/jaets.v4i1.966

Liu, Q., Miao, K., & Lin, K. (2019). Inventory Management of Automobile After-sales Parts Based on Data Mining. *Proceedings of the 2019 3rd High Performance Computing and Cluster Technologies Conference*, 195–199. https://doi.org/10.1145/3341069.3342975

Metisen, B. M., & Sari, H. L. (2015). Analisis clustering menggunakan metode K-Means dalam pengelompokkan penjualan produk pada Swalayan Fadhila. *Jurnal Media Infotama*, *11*(2), 110–118. https://doi.org/10.37676/jmi.v11i2.258

Munyaka, J.-C. B., & Yadavalli, S. V. (2022). Inventory Management Concepts and Implementations: A Systematic Review. *South African Journal of Industrial Engineering*, *32*(2), 15–36. https://doi.org/10.7166/33-2-2527

Musen, M. A. (2015). The protégé project: A look back and a look forward. *AI Matters*, *1*(4), 4–12. https://doi.org/10.1145/2757001.2757003

Patel, S. (2019). *K-means Clustering Algorithm: Implementation and Critical Analysis* (1. Auflage). Scholars' Press.

Prayoga, G. I., Henri, H., Mustikarini, E. D., & Anggyansyah, A. (2022). Diversity and morphological relationship of orchid species (Orchidaceae) in Bangka Island, Indonesia. *Biodiversitas Journal of Biological Diversity*, *23*(10), 5323–5332. https://doi.org/10.13057/biodiv/d231042

Ridwan, M., & Rakhmawati, F. (2023). Analysis of The Selection of The Best Arabica Coffee Beans Using Apriori Algorithms. *Mathline : Jurnal Matematika Dan Pendidikan Matematika*, *8*(2), 739–752. https://doi.org/10.31943/mathline.v8i2.437

Sampathila, N., Chadaga, K., Goswami, N., Chadaga, R. P., Pandya, M., Prabhu, S., Bairy, M. G., Katta, S. S., Bhat, D., & Upadya, S. P. (2022). Customized Deep Learning Classifier for Detection of Acute Lymphoblastic Leukemia Using Blood Smear Images. *Healthcare*, *10*(10), 10101812. https://doi.org/10.3390/healthcare10101812

Thogarcheti, H. S., Pulabaigari, V., & K, M. (2020). *Improvements over k-means clustering methods for large datasets Prototype based hybrid techniques* (1. Auflage). Lambert Academic Publishing.

Von Rintelen, K., Arida, E., & Häuser, C. (2017). A review of biodiversity-related issues and challenges in megadiverse Indonesia and other Southeast Asian countries. *Research Ideas and Outcomes*, *3*, e20860. https://doi.org/10.3897/rio.3.e20860

Wanto, A. (2020). *Data Mining: Algoritma dan Implementasi*. Yayasan Kita Menulis.

Witthaut, M., & Kalbe, N. (2022). Sheet metal assortment optimization with k-Means. *The 21st International Conference on Modelling and Applied Simulation*, 003. https://doi.org/10.46354/i3m.2022.mas.003

Wu, J. (Ed.). (2012). *Advances in K-means clustering: A data mining thinking*. Springer.