Classification of Sachet Beverage Products Based on Sales Data with Using the K-Means Clustering Method

Eko Afrianto[1*], Iqbal Sabillirasyad[2], Hayatul Maspufah[3]

[1,2] Faculty of Science, Technology and Industry, Mandala Institute of Technology and Science, Indonesia
[3] Faculty of Economic and Business, Mandala Institute of Technology and Science, Indonesia
Corresponding Author: ekoafrianto@itsm.ac.id

Abstract

Sachet drinks are one of the popular products in Indonesia, especially among urban communities who are looking for convenience in consumption. With increasing amount products circulating in the market, classification product based on sales data become important for understand preference consumers and effective marketing strategies. Research This aim for classify product sachet drinks using K-Means Clustering method, which is technique data analysis that can grouping data based on similarities characteristics. Through analysis this, it is hoped can obtained more insight in about trend sales and preferences consumers, as well as help manufacturer in take more decisions appropriate.

1. Introduction

Sachet drinks have been become an integral part of life daily Indonesian society. Growth This No let go from change style life an increasingly diverse society busy, where are they? more choose convenience in consume product. Phenomenon This reflect shift in pattern consumption, where consumers more prioritize efficiency and practicality in choose product drinks. With Thus, it is important for do in -depth analysis against sales data for identify existing patterns and trends.

K-Means Clustering is one of the frequent methods used in data analysis for grouping data into in a number of categories based on similarities characteristics. This method operates with share data to in K groups, where each group has a centroid that represents the average of the data in group The use of K -Means Clustering in context industry sachet drinks provide opportunity for company for understand different market segments. For example, with grouping sales data based on characteristics demographics, such as age and taste preferences, company can adjust marketing strategy them to be more appropriate target [1].

In context product sachet drinks, analysis This can help in understand preference consumer to various type products, such as taste, packaging, and price. As example, product with fruit flavor Possible more in demand by consumers young people looking for freshness and innovation, while product with a classic taste like coffee or tea tend more preferred by more consumers old people who want a familiar and nostalgic taste. With information this, producer can optimize line product them and design campaign more marketing effective. For example, if the data shows that larger market segments young more like colored packaging bright and modern design, manufacturer can design packaging new and interesting attention segment mentioned [2].

In addition, the classification products can also help in management inventory. With know which product has sale high and which ones are not, companies can reduce risk excess stock and maximize benefits. As for example, if analysis show that product certain experience declines sales, company can take step proactive for reduce production or replace product the with other more variants in accordance with preference consumer research This aims for giving more insight in about classification product sachet drinks based on sales data use K-Means Clustering method, as well as give recommendation for manufacturer in develop a more effective marketing strategy effective [3].

With background behind said, research This will discuss methodology used in data analysis, results obtained from implementation K-Means method, as well as implications from results the for industry sachet drinks in Indonesia. It is hoped that study This can give significant contribution for marketing strategy development and management products in the sector this. Through greater

understanding Good about preference consumers and market trends, producers can formulate a more strategic strategy adaptive and responsive to continuous market needs changed [4].

In analysis deep this is important For investigate more carry on How variables like price, packaging, and taste are mutual interact and influence decision purchase consumers. Shows that competitive prices and innovative tastes can become factor determinant in decision purchase sachet drinks. With Thus, entrepreneurs need consider No only One aspects, but also combinations from various factor For create products that can accepted by the market [5].

As closing, in -depth analysis against sales data sachet drinks using K-Means Clustering method does not only give outlook about preference consumers, but also allows manufacturer For formulate a better marketing strategy effective and efficient. With understand characteristics and needs every market segment, company can optimize product them and improve Power competition in an increasingly competitive market strict. Research This expected can give significant contribution for development industry sachet drinks in Indonesia, as well as open opportunity For study more carry on in field This.

2. Methods

In research this, sales data product sachet drink taken from various brand and type beverages available in the Indonesian market. This data covers estimate sale monthly, average sales daily, and percentage market share of each product. For example, the product Nutrisari Orange with type drink powder fruit own estimate sale of 2,700 units per month, which is equivalent to with 90 units per day, and controlling 5.4% of the total market share. This data is very important For understand market dynamics and behavior consumer in choose product sachet drinks.

Product classification based on sales data can help producers and marketers in determine more marketing strategies effective. With use K-Means Clustering method, we can grouping products the based on similarities in sales data. For example, instant coffee products such as Good Day Original and Torabika The duo has estimate relative sales high, namely 2,300 and 2,000 units per month, respectively, with market share of 4.6% and 4.0%. This show that instant coffee products own Power strong attraction among consumers [6].

In addition, the product drink Milk powders like Pop Ice also show significant sales, with Chocolate Ice Pops reached 1,100 units per month. This shows that category drink milk powder has huge market potential. With utilizing this data, we can do analysis more carry on For determine factors that influence sales and preferences consumers.

Data collection was carried out with method survey and secondary data analysis from report available sales and market statistics. The data obtained Then cleaned and prepared For analysis more further. In research Here, in table 1, the data used covers 40 products sachet drinks from various categories, including coffee, tea, and beverages powder fruit, which provides comprehensive overview about the sachet beverage market in Indonesia [7].

No	Brand / Product	Types of Drinks	Sales /Month	Average/Day	% Market
1	Nutrisari Orange	Drink powder fruit	2,700	90	5.4%
2	Nutrisari Mango	Drink powder fruit	1,700	57	3.4%
3	Good Day Original	Instant coffee	2,300	77	4.6%
4	Good Day Mocacinno	Instant coffee	1,800	60	3.6%
5	Good Day Chococinno	Instant coffee	1,500	50	3.0%
6	Torabika Duo	Instant coffee	2,000	67	4.0%
7	Torabika Cappuccino	Instant coffee	1,700	57	3.4%
8	Torabika Creamy Latte	Instant coffee	1,300	43	2.6%
9	Indocafe Coffeemix	Instant coffee	1,500	50	3.0%

10	Indocafe 3 in 1	Instant coffee	1,200	40	2.4%
11	Maxtea Tarikk	Tea pull	2,100	70	4.2%
12	Tea Sachet Top	Instant sweet tea	1,300	43	2.6%
13	Sachet Glass Tea	Instant sweet tea	1,000	33	2.0%
14	Chocolate Pop Ice	Drink milk powder	1,100	37	2.2%
15	Strawberry Pop Ice	Drink milk powder	1,000	33	2.0%
16	Pop Ice Durian	Drink milk powder	900	30	1.8%
17	Energen Chocolate	Drink cereals	1,100	37	2.2%
18	Ginger Energen	Drink cereals	800	27	1.6%
19	Energen Vanilla	Drink cereals	800	27	1.6%
20	Chocolatos Drink	Drink chocolate	1,000	33	2.0%
21	Ultra Tea Instant	Tea powder	800	27	1.6%
22	Beng Beng Drink	Instant milk chocolate	1,000	33	2.0%
23	Top Coffee 3 in 1	Instant coffee	1,000	33	2.0%
24	Kapal Api Signature Sachet	Black coffee instant	1,000	33	2.0%
25	Kapal Api Special Mix	Black coffee instant	900	30	1.8%
26	Kopiko Blanca	White coffee instant	1,000	33	2.0%
27	Luwak White Coffee Sachet	White coffee instant	1,000	33	2.0%
28	Dancow Fortigro Sachet	Milk powder instant	800	27	1.6%
29	Milo Sachet	Instant milk chocolate	1,000	33	2.0%
30	ABC Fragrant Ginger	Drink ginger	700	23	1.4%
31	Sariwangi Tea Instant	Tea instant	700	23	1.4%
32	Nestlé Carnation Choco	Drink chocolate	700	23	1.4%
33	Frisian Flag Sweetened Condensed Milk	Condensed milk instant	700	23	1.4%
34	Red Ginger Plus Sido Muncul	Herbal drinks	600	20	1.2%
35	Sosro Bottled Tea Instant	Instant sweet tea	700	23	1.4%
36	ABC Coffee Milk Sachet	Coffee milk	800	27	1.6%

37	Nestlé Hot Milo Sachet	Chocolate hot instant	800	27	1.6%
38	Indomilk Chocolate Sachets	Chocolate milk	800	27	1.6%
39	Kuku Bima Ener-G Sachet	Drink energy	600	20	1.2%
40	Lasegar Sachet (powder)	Drink refresher body	600	20	1.2%
	Total		50,000	1,667	100%

Table 1. Sales Data Sachet Drinks
Source Data: Apindo 2025

The K-Means Clustering method is one of the technique clustering of large amounts of data used in quantitative data analysis, including in context marketing like sales data grouping product sachet drinks. Clustering itself is something technique purposeful data exploration For grouping objects to in a number of cluster based on similarity characteristics certain, so that object in one cluster has high similarity, while object between clusters have significant difference [8].

K-Means method works with method minimize amount square distance between data with its cluster center. In its implementation against sales data sachet drinks, every product can represented by attributes numeric like amount monthly sales and average sales daily or weekly. This method is very suitable used For know patterns group sale like product with sale high, medium, or low, which then can used For taking marketing strategy decisions, arrangements stock, or promotion [9].

In a way Mathematically, the K-Means process begins with choose number of clusters (k) manually, then done centroid initialization (point cluster center) random or based on method specific. Iterative process done For grouping data to the nearest centroid and updating the centroid based on the average position of the data in each cluster [10]. Formula main in The K-Means method is formula Euclidean distance is used For measure distance between data points to the centroid, namely:

$$d(x, c_j) = \sqrt{\sum_{i=1}^{n}(x_i - c_{ji})^2}$$

Information:

- $d(x, c_j)$ is the distance between the æ-th data and the j-th centroid,

- X_i is the value of the i-th attribute of the data,

- C_{ji} is the value of the i-th attribute of the j-th cluster centroid,

- n is the number of attributes.

After all data is classified to in the nearest cluster, the cluster center is updated with using the average of all over data points in the cluster. This process repeated until cluster center no changed Again or has reach iteration maximum. Formula centroid (cluster center) update for each jth cluster is :

$$c_j = \frac{1}{|C_j|}\sum_{x \in C_j} x$$

Information :

- Cj is the new centroid vector for the j-th cluster,
- Cj is the data set in the jth cluster,
- x is the data in the j-th cluster,
- Cj is the number of data in the j-th cluster.

3. Results and Discussion

1) Data Selection

Total sales every product during period time certain are 40 data points that will be entered to in category existing products in study this, includes amount transaction product, total sales, and average quantity transactions, each of which has three feature description. Period transaction for 1 month, or June 2025 which is used in study This is the processed data. can seen in Table 2.

No	Brand / Product	Sale	Average
1	Nutrisari Orange	2,700	90,000
2	Nutrisari Mango	1,700	56,667
3	Good Day Original	2,300	76,667
4	Good Day Mocacinno	1,800	60,000
5	Good Day Chococinno	1,500	50,000
6	Torabika Duo	2,000	66,667
7	Torabika Cappuccino	1,700	56,667
8	Torabika Creamy Latte	1,300	43,333
9	Indocafe Coffeemix	1,500	50,000
10	Indocafe 3 in 1	1,200	40,000
11	Maxtea Tarikk	2,100	70,000
12	Tea Sachet Top	1,300	43,333
13	Sachet Glass Tea	1,000	33,333
14	Chocolate Pop Ice	1,100	36,667
15	Strawberry Pop Ice	1,000	33,333
16	Pop Ice Durian	900	30,000
17	Energen Chocolate	1,100	36,667
18	Ginger Energen	800	26,667
19	Energen Vanilla	800	26,667
20	Chocolatos Drink	1,000	33,333
21	Ultra Tea Instant	800	26,667
22	Beng Beng Drink	1,000	33,333
23	Top Coffee 3 in 1	1,000	33,333
24	Kapal Api Signature Sachet	1,000	33,333
25	Kapal Api Special Mix	900	30,000
26	Kopiko Blanca	1,000	33,333
27	Luwak White Coffee Sachet	1,000	33,333
28	Dancow Fortigro Sachet	800	26,667
29	Milo Sachet	1,000	33,333
30	ABC Fragrant Ginger	700	23,333
31	Sariwangi Tea Instant	700	23,333
32	Nestlé Carnation Choco	700	23,333
33	Frisian Flag Sweetened Condensed Milk	700	23,333
34	Red Ginger Plus Sido Muncul	600	20,000
35	Sosro Bottled Tea Instant	700	23,333
36	ABC Coffee Milk Sachet	800	26,667
37	Nestlé Hot Milo Sachet	800	26,667
38	Indomilk Chocolate Sachets	800	26,667
39	Kuku Bima Ener-G Sachet	600	20,000

40	Lasegar Sachet (powder)	600	20,000

Table 2. Processed Data

Source Data: Data processed by researchers

2) Determine Number of Clusters and Initial Centroid Values

At the stage early, has set that two clusters is the most effective numbers For classify interest from customers. Number of clusters and initial Centroid value must determined next. The clusters that are formed namely high cluster (C1) and cluster (C2). C1 is obtained from mark The highest results in Table 2 and C2 were obtained from mark lowest listed in table 2. How to look for initial centroid value For iteration 1, namely:

C1 = Max (90,000; 56,667; 76,667; 60,000; 50,000; 66,667; 56,667; 43,333; 50,000; 40,000; 70,000; 43,333; 33,333; 36,667; 33,333; 33,333; 33,333; 23,333; 23,333; 23,333; 23,333; 20,000; 23,333; 26,667; 26,667; 26,667; 20,000; 20,000) = 90,000

C2 = Min (90,000; 56,667; 76,667; 60,000; 50,000; 66,667; 56,667; 43,333; 50,000; 40,000; 70,000; 43,333; 33,333; 36,667; 33,333; 33,333 ; 33,333; 23,333 ; 23,333; 23,333; 23,333; 20,000; 23,333; 26,667 ; 26,667; 26,667; 20,000; 20,000) = 20,000

Based on calculation determine number of clusters and initial centroid values above can known initial centroid value, where For look for mark maximum and minimum through the iteration process. Then it will the initial data centroid is obtained, as can be seen seen in table 3 below This.

Information	Amount
C1	90,000
C2	20,000

Table 3. Initial Data Centroid

After determine the initial centroid from each cluster, steps furthermore in the K-Means Clustering process is count distance each data against each centroid, namely C1 and C2. Calculation distance This aim For know how close is it data for each cluster center, so that can done grouping beginning. Generally, the distance used is Euclidean Distance (Euclidean distance). With count distance each data to C1 and C2, then can the data is determined more near to which centroid, and then the data will be classified to in the appropriate cluster. This process will repeated until the centroid experiences convergence or No Again changed in a way significant in iteration next.

3) Counting distance between each data C1 and C2

Distance used is Euclidean Distance, which is formulated as following:

$$d = \sqrt{(x-c)^2}$$

Information:

X = Data Value

C = Centroid Value

Then will perform the calculation process distance each data in C1, calculation This based on distance closest from every product sachet drink with centroid:

Product 1 = $\sqrt{(90,000 - 90,000)^2}$ = 0

Product 2 = $\sqrt{(56,667 - 90,000)^2}$ = 34,333

Product 3 = $\sqrt{(76,667 - 90,000)^2}$ = 13,333

Product 4 = $\sqrt{(60,000 - 90,000)^2}$ = 30,000

Product 5 = $\sqrt{(50,000 - 90,000)^2}$ = 40,000

Product 6 = $\sqrt{(66,667 - 90,000)^2}$ = 23,333

Product 7 = $\sqrt{(56,667 - 90,000)^2}$ = 34,333

Product 8 = $\sqrt{(43,333 - 90,000)^2}$ = 46,667

Product 9 = $\sqrt{(50,000 - 90,000)^2}$ = 40,000

…

Product 40 = $\sqrt{(20,000 - 90,000)^2}$ = 70,000

After perform the calculation process distance each data in C1 then will the calculation process is carried out on C2, the calculation This based on distance closest from every product sachet drink with centroid:

Product 1 = $\sqrt{(90,000 - 20,000)^2}$ = 70,000

Product 2 = $\sqrt{(56,667 - 20,000)^2}$ = 36,667

Product 3 = $\sqrt{(76,667 - 20,000)^2}$ = 56,667

Product 4 = $\sqrt{(60,000 - 20,000)^2}$ = 40,000

Product 5 = $\sqrt{(50,000 - 20,000)^2}$ = 30,000

Product 6 = $\sqrt{(66,667 - 20,000)^2}$ = 46,667

Product 7 = $\sqrt{(56,667 - 20,000)^2}$ = 36,667

Product 8 = $\sqrt{(43,333 - 20,000)^2}$ = 23,333

Product 9 = $\sqrt{(50,000 - 20,000)^2}$ = 30,000

…

Product 40 = $\sqrt{(20,000 - 20,000)^2}$ = 0

Based on calculation the distance between each data in C1 and C2, then can be seen results from calculation distance with point center on Iteration, in table 4.

No	Brand / Product	Shortest Distance	C1	C2
1	Nutrisari Orange	0	1	
2	Nutrisari Mango	34,333	1	
3	Good Day Original	13,333	1	
4	Good Day Mocacinno	30,000	1	
5	Good Day Chococinno	30,000		1
6	Torabika Duo	23,333	1	
7	Torabika Cappuccino	34,333	1	
8	Torabika Creamy Latte	23,333		1
9	Indocafe Coffeemix	30,000		1
10	Indocafe 3 in 1	20,000		1
11	Maxtea Tarikk	20,000	1	
12	Tea Sachet Top	23,333		1
13	Sachet Glass Tea	13,333		1
14	Chocolate Pop Ice	16,667		1
15	Strawberry Pop Ice	13,333		1
16	Pop Ice Durian	10,000		1
17	Energen Chocolate	16,666		1
18	Ginger Energen	6,667		1
19	Energen Vanilla	6,667		1
20	Chocolatos Drink	13,333		1
21	Ultra Tea Instant	6,667		1
22	Beng Beng Drink	13,333		1
23	Top Coffee 3 in 1	13,333		1
24	Kapal Api Signature Sachet	13,333		1
25	Kapal Api Special Mix	10,000		1
26	Kopiko Blanca	13,333		1
27	Luwak White Coffee Sachet	13,333		1
28	Dancow Fortigro Sachet	6,667		1
29	Milo Sachet	13,333		1
30	ABC Fragrant Ginger	3,333		1
31	Sariwangi Tea Instant	3,333		1
32	Nestlé Carnation Choco	3,333		1
33	Frisian Flag Sweetened Condensed Milk	3,333		1
34	Red Ginger Plus Sido Muncul	0		1
35	Sosro Bottled Tea Instant	3,333		1
36	ABC Coffee Milk Sachet	6,667		1
37	Nestlé Hot Milo Sachet	0		1
38	Indomilk Chocolate Sachets	6,667		1

| 39 | Kuku Bima Ener-G Sachet | 0 | 1 |
| 40 | Lasegar Sachet (powder) | 0 | 1 |

Table 4. Results of Data Distance with Clusters
Source Data: Data processed by researchers

Based on the data presented in Table 4, it can be seen that results calculation distance between every sales data with initial cluster center (centroid) that has been determined. The calculation process This aim For determine the proximity of each data to the two existing cluster centers, namely C1 and C2. From the results calculation said, can identified that the data with mark sale highest own further distance near with cluster center C1, so that the data grouped to in cluster 1 (C1). On the other hand, data with mark sale lowest show greater closeness big with cluster center C2 and therefore That entered to in cluster 2 (C2).

Grouping This reflect characteristics from each cluster, where cluster C1 tends to consists of on products with level high sales, while cluster C2 reflects group product with level relative sales more low. This step is part important in K-Means Clustering method because help in identify pattern data distribution based on intensity sales, which is next can used For marketing strategy analysis or taking decision business other.

4) Determine Group and Recalculate Cluster Values

In the initial process grouping use K-Means Clustering method, system moreover formerly do calculation distance between every data points in matter this is sales data product with point cluster center (centroid) which has been determined previously. The calculated distance use formula certain, such as Euclidean Distance, works For measure how much close to each data every cluster center. After all distance calculated, the data will be grouped based on proximity distance namely, with choose mark the smallest distance. The more small mark the distance, the more big the possibility of the data is at in the same group with the center of the cluster.

If results calculation show that distance shortest is in cluster 1 (C1), then the data in a way automatic will entered to in cluster 1. On the other hand, if distance smallest is in cluster 2 (C2), then the data will classified to in cluster 2. This process done For all data in systematic and repetitive. The results of stage grouping beginning or iteration First This can seen in a way clear in Table 5, which shows sales data sharing to in each cluster based on distance closest towards the centroid. Stages This is part crucial in the clustering process because become base For calculation of new centroids on iteration next until achieved stability in grouping.

Cluster	Mark
C1	7
C2	33

Table 5. Cluster Iteration Results

5) Iteration of K-Means Clustering Process

After the grouping process beginning finished, system will continue to stage next with do calculation repeat use point cluster center (centroid) obtained from results grouping previous. New centroid value counted based on the average of all combined data in each cluster in the iteration previously. This process will Keep going repetitive in form iteration, where each iteration will generate new centroids and clustering repeat data, until achieved condition convergent. Convergence happen when No Again there is change in data sharing to in cluster and fixed centroid value The same as in iteration previously.

In case this, after done calculation repeat on iteration second, found that No There is changes in the centroid value as well composition member each cluster. This means that the system has reach stability or point stopping point, so that the K-Means Clustering process is stopped. The results end from this process show that The cluster center is at the value of C1 = 7 and C2 = 33. The final result from grouping the served in Table 6, which shows distribution final sales data to in each cluster optimally based on approach iterative mentioned. Stages this is very important Because ensure that results obtained grouping truly reflect structure experience from the analyzed data.

No	Brand / Product	Sale	Average	C1	C2	Cluster
1	Nutrisari Orange	2,700	90,000	0	70,000	C1
2	Nutrisari Mango	1,700	56,667	34,333	36,667	C1

3	Good Day Original	2,300	76,667	13,333	56,667	C1
4	Good Day Mocacinno	1,800	60,000	30,000	40,000	C1
5	Good Day Chococinno	1,500	50,000	40,000	30,000	C2
6	Torabika Duo	2,000	66,667	23,333	46,667	C1
7	Torabika Cappuccino	1,700	56,667	34,333	36,667	C1
8	Torabika Creamy Latte	1,300	43,333	46,667	23,333	C2
9	Indocafe Coffeemix	1,500	50,000	40,000	30,000	C2
10	Indocafe 3 in 1	1,200	40,000	50,000	20,000	C2
11	Maxtea Tarikk	2,100	70,000	20,000	50,000	C1
12	Tea Sachet Top	1,300	43,333	46,667	23,333	C2
13	Sachet Glass Tea	1,000	33,333	56,667	13,333	C2
14	Chocolate Pop Ice	1,100	36,667	53,333	16,667	C2
15	Strawberry Pop Ice	1,000	33,333	56,667	13,333	C2
16	Pop Ice Durian	900	30,000	60,000	10,000	C2
17	Energen Chocolate	1,100	36,667	53,333	16,666	C2
18	Ginger Energen	800	26,667	63,333	6,667	C2
19	Energen Vanilla	800	26,667	63,333	6,667	C2
20	Chocolatos Drink	1,000	33,333	56,667	13,333	C2
21	Ultra Tea Instant	800	26,667	63,333	6,667	C2
22	Beng Beng Drink	1,000	33,333	56,667	13,333	C2
23	Top Coffee 3 in 1	1,000	33,333	56,667	13,333	C2
24	Kapal Api Signature Sachet	1,000	33,333	56,667	13,333	C2
25	Kapal Api Special Mix	900	30,000	60,000	10,000	C2
26	Kopiko Blanca	1,000	33,333	56,667	13,333	C2
27	Luwak White Coffee Sachet	1,000	33,333	56,667	13,333	C2
28	Dancow Fortigro Sachet	800	26,667	63,333	6,667	C2
29	Milo Sachet	1,000	33,333	56,667	13,333	C2
30	ABC Fragrant Ginger	700	23,333	66,667	3,333	C2
31	Sariwangi Tea Instant	700	23,333	66,667	3,333	C2
32	Nestlé Carnation Choco	700	23,333	66,667	3,333	C2
33	Frisian Flag Sweetened Condensed Milk	700	23,333	66,667	3,333	C2
34	Red Ginger Plus Sido Muncul	600	20,000	70,000	0	C2
35	Sosro Bottled Tea Instant	700	23,333	66,667	3,333	C2
36	ABC Coffee Milk Sachet	800	26,667	63,333	6,667	C2
37	Nestlé Hot Milo Sachet	800	26,667	63,333	0	C2
38	Indomilk Chocolate Sachets	800	26,667	63,333	6,667	C2
39	Kuku Bima Ener-G Sachet	600	20,000	70,000	0	C2
40	Lasegar Sachet (powder)	600	20,000	70,000	0	C2

Table 6. Final Iteration Classification
Source Data: Data processed by researchers

6. Product Classification Results

     After all over stages iteration in K-Means Clustering method is complete done and conditions convergent has reached where the centroid value is not Again experience changes, then obtained results end in the form of grouping data into in each cluster that has been determined previously.

Final result from this process show that every data has been grouped to in the most suitable cluster based on proximity the distance with centroid. Information about results grouping the can seen in a way more detailed in Table 7, which displays distribution object or data to in each group (Cluster 1 and Cluster 2), so that can known pattern or characteristics from each cluster formed.

Information	Number of Products	Member Product
Cluster 1	7	Nutrisari Orange, Nutrisari Mango, Good Day Original, Good Day Mocacinno, Torabika Duo, Torabika Cappuccino, Maxtea Pull
Cluster 2	33	Good Day Chococinno, Torabika Creamy Latte, Indocafe Coffeemix , Indocafe 3 in 1, Tea Pucuk Sachet, Teh Gelas Sachet, Pop Ice Chocolate , Pop Ice Strawberry , Pop Ice Durian, Energen Chocolate , Energen Ginger, Energen Vanilla, Chocolatos Drink, Ultra Tea Instant , Beng Beng Drink, Top Coffee 3 in 1, Kapal Api Signature Sachet, Kapal Api Special Mix, Kopiko Blanca, Luwak White Coffee Sachet, Dancow Fortigro Sachet, Milo Sachet, ABC Wangi Ginger, Sariwangi Tea Instant , Nestlé Carnation Choco, Frisian Flag Sweetened Condensed Milk, Red Ginger Plus Sido Muncul , Sosro Bottled Tea Instant , ABC Coffee Milk Sachet, Nestlé Hot Milo Sachet, Indomilk Chocolate Sachet, Kuku Bima Ener-G Sachet, Lasegar Sachet (powder)

Table 7. Product Classification Results
Source Data: Data processed by researchers

4. Conclusion

Based on results analysis use K-Means Clustering method on sales data product sachet drinks, can concluded that implementation algorithm This succeed identify pattern interest consumer for each product. Classification process This produce two groups main, namely Cluster 1 which consists of of 7 products with level interest high consumer demand, and Cluster 2 which includes 33 products with interest more consumers low. With Thus, of the total of 40 products analyzed, only part small that shows performance significant sales.

Solutions that can take based on results classification This is with carry out more marketing strategies focused. Products in Cluster 1 can made into a mainstay or priority in promotion Because proven in demand by consumers. Meanwhile that, for products in Cluster 2, the company can evaluate revisit marketing strategy, improve packaging, customize price, or even consider reformulation product for increase Power pull it. In addition, the results grouping This can made into base in taking decision managerial related management stock, distribution, and innovation products to be more in accordance with market preferences.

References

Arisandi, SJ (2016). Prediction of Fruit Needs with Market Segmentation Using K-Means. *Annu. Res. Semin*, *2* (1), 156-158.

Association Indonesian Entrepreneurs. (2025). Report annual industry Food and beverage: Performance and trends sale product sachet drinks in 2025. Jakarta: APINDO.

Li, Y., & Wu, H. (2012). A clustering method based on K-means algorithm. *Physics Procedia*, *25*, 1104-1109.

Awalina, EFL, & Rahayu, WI (2023). Optimizing Marketing Strategy with Segmentation Customer Use Application of K-Means Clustering in Online Retail Transactions. *Journal Technology and Information*, *13* (2), 122-137.

Darmi, YD, & Setiawan, A. (2016). Implementation k-means clustering method in grouping sale product. *Infotama Media Journal*, *12* (2).

Dubey, A., & Choubey, APDA (2017). A systematic review on k-means clustering techniques. *Int J Sci Res Eng Technol (IJSRET, ISSN 2278–0882)*, *6* (6).

Morissette, L., & Chartier, S. (2013). The k-means clustering technique: General considerations and implementation in Mathematica. *Tutorials in Quantitative Methods for Psychology*, *9* (1), 15-24.

Rohmah, AN, & Subari, S. (2021). Preferences consumer to product coffee drinks at Kopi Janji Jiwa Volume 324 Surabaya. *Agriscience*, *1* (3), 548-562.

Situmorang , A., Tukino , T., Novalia , E., & Ahmad, S. (2022). Classification of Sales Results Drink Light on Cooperatives Based on the Type of Goods Used K-Means Clustering Algorithm. *Tika Journal*, *7* (3), 250-255.

Triyandana, G., Putri, LA, & Umaidah , Y. (2022). Application of Data Mining for Food and Beverage Menu Grouping Based on Sales Level Using the K-Means Method. *Journal of Applied Informatics and Computing*, *6* (1), 40-46.

Wardani, SDK, Ariyanto, AS, Umroh, M., & Rolliawati, D. (2023). Comparison results K-Means, DB Scanner & Hierarchical clustering methods for analysis market segmentation. *JIKO (Journal Informatics and Computers)*, *7* (2), 191-201.