Published online on the page: https://ejournal.iainbukittinggi.ac.id/index.php/ijokid/

Knowbase : International Journal of Knowledge in Database

| ISSN (Print) 2798-0758 | ISSN (Online) 2797-7501 |

K-Means Clustering in Determining the Category of Stock Items In Angkasa Mart

Roki Hardianto[1,], Hardiansyah Ramadhan[2], Eddissyah Putra Pane[3], Yogi Yunefri[4]*

[1234]Universitas Lancang Kuning, Pekanbaru, Indonesia

Article Information

Article History:
Accepted by the Editor: June 29, 2022
Final Revision: June 27, 2022
Published *Online*: June 30, 2022

Keyword

Clustering
K-Means
Web-Based

Correspondence

E-mail: roki@unilak.ac.id*

A B S T R A C T

A good and accurate stock management system is crucial in any institution or organization that conducts product buying and selling transactions. This is done to improve stock efficiency, reduce storage costs, be more effective, and meet customer expectations. This research produced three groups of the most desirable products for large stock quantities, moderate stock quantities for products of interest, and low stock quantities for products that were less or not in demand. Data was processed using the clustering method, specifically the K-Means method, which was based on historical sales data that includes the product code, number of transactions, and average sales. This study was carried out using web-based calculations. Testing with the rapid miner application. The research yielded three members of the product group for large stock, 238 large stock, 273 medium stock, and 25 low stock.

1. Introduction

Internet, computer, and technology (ICT) is rapidly expanding, particularly since the Covid-19 pandemic in Indonesia. This is characterized by the use of ICT in all fields in order to break the chain of Covid-19 spread. In the business world, especially for minimarkets, the use of ICT is crucial to carry out effective and efficient management and service processes for consumers. A minimarket can use storage media to store data about its inventory. Data for a minimarket can be very useful if it is further processed to produce information that can be used to make decisions.

Data is the main source from which information can be generated. Maximum data processing will result in high-quality information. Data mining functions convert warehouse data stored in a database into new knowledge that can be used to make decisions [1].

Angkasa Mart is a store that sells food, drinks, cosmetics, toiletries, baby gear, and other items. Similar to other minimarkets, Angkasa Mart has not implemented a computer-based system in its store management. With management still not fully ICT-based, there are certain aspects of decision-making that are not met. The fluctuating number of consumer requests results in an unstable supply that must be prepared. Furthermore, the variety of products makes stock management inaccurate, has high storage costs, is ineffective, and frequently disappoints consumers due to the vacancy of a specific product.

As a result, inventory or stock is an important activity because if there are no inventories, the customer may be dissatisfied if the goods required are not available. Determining the appropriate items in stock is also

important because if the category of stock items is not determined, there will be losses if items are not sold, expired, and production determination of the amount of product stock is less accurate because it must be based on knowledge of a large amount of sales transaction data. Thus, large transaction history data must be processed using data mining techniques and website-based system design. The Rapid Miner tool is used during the testing process to make the reporting process easier.

The clustering process using the website-based K-Means algorithm has the advantage of being multiplatform, as it can be used on computers, tablets, and smartphones. This certainly makes it easier for users to have flexible access and maintain work productivity. To determine product groups or sales data faster at the Angkasa Mart, a website-based system (desktop) is required.

2. Method

Data mining is one of the sciences used to determine specific information in a set of data to assist decision-making. Data mining is also a process to obtain useful information from a set of data. The discovery of knowledge in a database is referred to as data mining. Data mining is a process that uses statistical, mathematical, artificial intelligence, and machine learning techniques to interact with large databases in order to identify useful information and linked knowledge [2].

Clustering is a technique for locating and grouping data that have similar characteristics (similarity) to one another. Clustering is one of the undirected data mining techniques. Clustering is the process of categorizing data into classes or clusters based on their degree of similarity. Data that has similarities is included in the same cluster, while data that does not have similarities is included in different clusters [3].

The K-Means method is one of the methods used in the clustering or grouping function. Clustering is the process of grouping data, observations, or cases based on the object under study's similarity. A cluster is a group of data that is similar to or dissimilar to data in other groups [4].

K-means is a data mining method in which each data must be included in a specific cluster and it is possible for each data included in a particular cluster at one stage of the process to move to another cluster at the next stage [5]. K-means, which divides data into K sub-regions, is well known for its ability to classify and outline large amounts of data quickly. The k-means algorithm is thought to have a high enough level of accuracy to be used in this research to produce knowledge of the group characteristics of the area [6]. K-Means is a clustering algorithm that uses an iterative process. The number of clusters to be created is denoted by the letter K [7]. The value of K is then assigned at random. While the means is a temporary value that represents the cluster's center, also known as the centroid [8]. The distance from each existing data to the centroid is calculated using the Euclidean formula to produce the closest distance from each data to the centroid [9].

3. Results and Discussion

The research was conducted at Angkasa Mart in Pekanbaru City, Riau. This store is like Alfamart and Indomaret, selling essential items to the community. Why is Angkasa Mart a research location ?, because the management process in this store is still implementing a stock analysis system using manual methods and has not applied computerized methods such as the K-Means Clustering method.

In the pre-process of data clustering, raw data from Angkasa Mart is converted into relevant data. This process describes the transformation of raw data into data sets that can be used in the clustering process. The data used is complete data, and several attributes are chosen to obtain data that is relevant and in accordance with the needs to be processed and processed further.

Because beverages are among the most widely sold products, this research focused on product clustering in the type of beverage. The total number of beverage transactions in this study was 1184, with 536 products. The following data is used; however, only 100 of them are displayed in this table:

Table 1. Main Data

No	Item Code	Barcode	Item	July	Agustus	September	October	Total
1	Adm	8991002122017	ABC Eco Milk Coffe 230ml	2	3		4	9
2	ADC	8991002122000	Abc Exo chocomal 230ml		2	4		6
3	A0001	711844160057	ABC Jus Apel UHT 250	3			5	8
4	A0002	711844160071	ABC Jus Jambu UHT 250			5		5
5	A0004	711844160118	ABC Jus Mangga UHT 250	1	4		5	10
6	A0005	711844160019	ABC Jus Orange UHT 250	3	5			8
7	A0043	711844162419	ABC Sari Kacang Hijau UHT 250		3		6	9
8	ABCS	711844160095	ABC SOURSOP 250ML	6		3	2	11
9	ABCSPC	711844154506	ABC SP Cocopandan 600	7	3	5		15
10	ABCSPM	711844154568	ABC SP Melon 585	9	1	4	2	16
11	A0003	711844150003	Abc squash delight 525ml		3	5		8
12	AWC22	8991002101722	ABC WHAIE COFFE			1		1
13	AK0641	8998667300910	ACTIV WATER APPLE PEACH 380ML		9			9
14	AK0644	8998667300927	ACTIV WATER LEMON 380ML			6		6
15	AK0642	8998667300934	ACTIV WATER ORANGE 380ML	2			3	5
16	AK0641	8998667300910	ACTIV WATER APPLE PEACH 380ML		3		4	7
17	AK0644	8998667300927	ACTIV WATER LEMON 380ML	8		6		14
18	AK0642	8998667300934	ACTIV WATER ORANGE 380ML		1	3	4	8
19	A0056	8992761139018	ADES	1		5		6
20	AD6168	8992761166168	ADES 350ML	7	2			9
21	ADS	8992761136277	ADES Air Mineral 1.5L	9		6		15
22	IACLB	896867700326	AIR MINERAL CLUB 600ML	7		2		9
23	AK0252	8997233240049	AIR MINERAL INDODES 1500ML	3	6			9
24	AK0251	8997233240032	AIR MINERAL INDODES 600ML		1		9	10
25	AK0171	AIR GREEN	AIR MINUM GREEN	7	5			12
26	8992772125024	8992772125024	AIR SEJUK ADEM SARI 600ML			3	6	9
27	AS220	8995077800177	AMDK SMS Cup 220Ml	8	3			11
28	A0071	8993199160001	AMIA Botol 600 Ml		2	1		3
29	AMIA	AMIA CUP	AMIA CUP 240 ML			9		9
30	8992772663090	8992772663090	AMUNIZER DUS 4S		8	5		13
31	AK0312	8992772665049	AMUNIZER VIT C 1000	3		5		8
32	A330	8886008101336	AQUA 330 ML		4	8		12
33	A0110	8886008101091	AQUA Botol 1500 ml	1		9		10

34	A0111	8886008101053	AQUA Botol 600ml				6	6
35	A0112	8886008101046	AQUA Gelas 240 ml Dus		8	2		10
36	A0113	8886008101138	AQUA Isi Ulang 19 Liter		1	7		8
37	A8500	8992761122485	AQUARIUS 250ML	1	3			4
38	B2448	8998009020520	BUAVITA Mini Guava UHT 125ml		6	8		14
39	B9367	8999999055714	BUAVITA Guava 1L	5			2	7
40	B313	8998009020193	BUAVITA Mango Uht 250			9		9
41	AK0112-NL	8998009020179	BUAVITA ORANGE 250ML		4			4
42	B6350	8998009020186	BUAVITA UHT Guava 250ml			2		2
43	Chg	8992802080101	chil go pisang 140ml	5				5
44	8992775002155	8992775002155	CHOCOLATOS DRINK RTD TETRA 200ML		2		2	4
45	661312	8993200661312	CIMORY GRAPE 250ML		5			5
46	CIMOY	8993200665006	CIMORY LOW FAT 250 ML			6	9	15
47	8993200664993	8993200664993	CIMORY LOW FAT BANANA 250ML	8		4		12
48	61336	8993200661336	CIMORY MANGO 250ML		3		7	10
49	C1350	8993200661350	CIMORY MIXED FRUIT 250Ml	5		1		6
50	666959	8993200666959	CIMORY SQUEEZ BLUEBERRY 120ML	4		9		13
51	666935	8993200666935	CIMORY SQUEEZ ORIGINAL 120ML		3	8		11
52	666942	8993200666942	CIMORY SQUEEZ STRAWBEERY 120ML		2		4	6
53	668045	8993200668045	CIMORY SQUEEZE ALOE VERA 120ML		7			7
54	664986	8993200664986	CIMORY TROPICAL FRUITS 250ML			9		9
55	C1657	8993200661657	CIMORY Yd Blueberry 250Ml		2		6	8
56	C1299	8993200661299	CIMORY Yd Lyche 250Ml	1	9			10
57	C3057	8993200663057	CIMORY Yd Mix Berry 250Ml		6			6
58	C3064	8993200663064	CIMORY Yd Plain 250Ml	3				3
59	C1305	8993200661305	CIMORY Yd Straw 250Ml			3		3
60	C004	8992775717004	CLEVO Choco 125ML		1			1
61	CLO	8992775001585	CLEVO COKLAT ICE CREAM			9		9
62	A0617	8992761136017	COCA-COLA 1500Ml				6	6
63	145019	0119922311579106	COCA-COLA 250ML		4			4
64	A0618	8992761002015	COCA-COLA Btl 390 Ml			8		8
65	CC250	8992761145019	COCA-COLA Can 250 Ml		7			7
66	CCB2O	089686827729	COCO BIT SPLASH GUAVA	6				6

67	CCB1	089686827194	COCO BIT SPLASH PEACH			2		2
68	C5129	089686827781	COCOBIT 350ML		5			5
69	ak0587	8996001600665	COOL TIME COCONUT WATER 200ML	7				7
70	A0625	8992772485012	COOLANT Cooling Water 350Ml		3			3
71	CJBB	8992772485029	COOLANT Jus Belimbing			5		5
72	CRYS383	8991102026383	CRYSTALINE 1500ML		1		3	4
73	CR0263	8991102026352	Crystalline 600Ml	5		7		12
74	D219	8997214880219	DBEST CINCAU 320ML		9			9
75	DBG	DBG350	DBEST GREEN TEA 320ML			4		4
76	DK350	DBEST KUNDUR	DBEST KUNDUR 350ML	3				3
77	DL350	DBEST LECY	DBEST LECY 350ML			5		5
78	D127	8997214880127	DBEST SARANG BURUNG 320ML		7			7
79	D'BSSB	D'BSSB	D'BEST Sarang Burung krat	1	9			10
80	DS350	DBEST SOYA	DBEST SOYA 350ML			8		8
81	DS0028	8997214880028	D'best Soya Milk 320ml		6			6
82	DTB350	DBEST TEH BUNGA	DBEST TEH BUNGA350 ML			1		1
83	E63603	8992802636032	ENTRASOL Active Vanilla Latte 360gr		3			3
84	E63703	8992802637039	ENTRASOL GOLD PLAIN 370gr		2		4	6
85	8997214880578	8997214880578	EX-GOAL SOYA 310ML		5			5
86	FT15L	8992761136055	FANTA Botol 1,5L			7		7
87	FO390	8992761002046	FANTA ORANGE 350Ml	8				8
88	145033	8992761145033	fanta stbr 250ml			9		9
89	F039	8992761002039	FANTA Stroberri 350Ml	3	2		4	9
90	2763	8992753721214	FF CHOCO HAZELNUT PET 270 ML		1	5		6
91	2762	8992753721207	FF CHOCO PET 270 ML	7			4	11
92	F9513	8992753700103	FF coconut 225ml				9	9
93	721009	8992753721009	FF FULL CREAM GOLD 490g		7	2		9
94	AK076-NL	8998866202923	FLORIDINA COCO BITS	6	1			7
95	F0070	8998866500708	Floridina Orange 360 Ml			8		8
96	F3350	8992761166212	FRESTEA Apel 350 Ml		5		2	7
97	A0985	8992761122331	FRESTEA Apel 500 Ml		3		4	7
98	FA1.5	8992761135980	FRESTEA APPLE 1.5L		6	4	1	11
99	A0986	8992761122317	FRESTEA Green 500 Ml	7		9		16
100	A0987	8992761122430	FRESTEA Green Madu 500		3	5	6	14

The data is then cleaned to determine the attributes, the attributes used in this study are two: total transactions for four months and the average number of sales in four months. Where the outcome is shown in the Table 1.

Table 2. Attribute Data

No	Item	Total	Average
1	ABC Eco Milk Coffe 230ml	9	2.25
2	Abc Exo chocomal 230ml	6	1.5
3	ABC Jus Apel UHT 250	8	2
4	ABC Jus Jambu UHT 250	5	1.25
5	ABC Jus Mangga UHT 250	10	2.25
6	ABC Jus Orange UHT 250	8	2
7	ABC Sari Kacang Hijau UHT 250	9	2,25
8	ABC SOURSOP 250ML	11	2,75
9	ABC SP Cocopandan 600	15	3,75
10	ABC SP Melon 585	16	4
11	Abc squash delight 525ml	8	2
12	ABC WHAIE COFFE	1	0,25
13	ACTIV WATER APPLE PEACH 380ML	9	2,25
14	ACTIV WATER LEMON 380ML	6	1,5
15	ACTIV WATER ORANGE 380ML	5	1,25
16	ACTIV WATER APPLE PEACH 380ML	7	1,75
17	ACTIV WATER LEMON 380ML	14	3,5
18	ACTIV WATER ORANGE 380ML	8	2
19	ADES	6	1,5
20	ADES 350ML	9	2,25
21	ADES Air Mineral 1.5L	15	2,25
22	AIR MINERAL CLUB 600ML	9	2,25
23	AIR MINERAL INDODES 1500ML	9	2,25
24	AIR MINERAL INDODES 600ML	10	2,5
25	AIR MINUM GREEN	12	3
26	AIR SEJUK ADEM SARI 600ML	9	2,25
27	AMDK SMS Cup 220Ml	11	2,75
28	AMIA Botol 600 Ml	3	0,75
29	AMIA CUP 240 ML	9	2,25
30	AMUNIZER DUS 4S	13	3,25
31	AMUNIZER VIT C 1000	8	2
32	AQUA 330 ML	12	3
33	AQUA Botol 1500 ml	10	2,5
34	AQUA Botol 600ml	6	2,5
35	AQUA Gelas 240 ml Dus	10	2,5
36	AQUA Isi Ulang 19 Liter	8	2
37	AQUARIUS 250ML	4	1
38	BUAVITA Mini Guava UHT 125ml	14	3,5
39	BUAVITA Guava 1L	7	1,75
40	BUAVITA Mango Uht 250	9	2,25

41	BUAVITA ORANGE 250ML	4	1
42	BUAVITA UHT Guava 250ml	2	0,5
43	chil go pisang 140ml	5	1,25
44	CHOCOLATOS DRINK RTD TETRA 200ML	4	1,25
45	CIMORY GRAPE 250ML	5	1,25
46	CIMORY LOW FAT 250 ML	15	3,75
47	CIMORY LOW FAT BANANA 250ML	12	3
48	CIMORY MANGO 250ML	10	2,5
49	CIMORY MIXED FRUIT 250Ml	6	1,5
50	CIMORY SQUEEZ BLUEBERRY 120ML	13	3,25

Process K Means:

1. Determine the Number of Clusters

   To know the results of the beverage product grouping and to get the best cluster, the number of clusters must be determined, and in this research, the number of clusters determined by the researcher was three to match the groups that would be produced: large, medium, and low.

2. Determine the Centroid

   In the rapid miner application, the centroid is determined by equating the centroid value. The cluster centers are shown in the table below.

Table 3. Centroid

Cluster	Total	Average
C1	3,403	0,901
C2	7,429	1,841
C3	13,12	3,24

3. Determine the Centroid Distance [10]

   The centroid distance is calculated for all of the data to be tested. The Euclidean Distance formula is used to calculate the distance between each existing data set.

$$D(ij) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \cdots + (X_{ki} - X_{kj})^2} \quad \text{.....(1)}$$

Where D(I,j) = Data distance to i to cluster center j

Xki = Data to i attribute data to k

Xkj = Center point to j on attribute k

The first distance with Euclidean Distance for iteration 1 with c1 is calculated as follows. And this pattern was repeated throughout the data. $\text{ABC Eco Milk Coffe 230ml} = \sqrt{(9 - 3{,}403)^2 + (2{,}25 - 0{,}901)^2} = 5{,}76$

Then, for iteration 1, compute the first distance with Euclidean Distance using c2. And this pattern was repeated throughout the data. $\text{ABCEco Milk Coffe 230ml} = \sqrt{(9 - 7{,}429)^2 + (2{,}25 - 1{,}84)^2} = 1{,}62$. Then, in the first iteration, use c3 to calculate the first distance. And the data kept going. $\text{ABC Eco Milk Coffe 230ml} = \sqrt{(9 - 13{,}12)^2 + (2{,}25 - 3{,}24)^2} = 4{,}24$

You can also conduct manual calculations in MS Excel, as shown in the figure below.

Jarak ke Pusat Cluster				
Pusat Cluster				
	C1	C2	C3	Keanggotaan Cluster
1	5,76	1,62	4,24	C2
2	2,67	1,47	7,33	C2
3	4,73	0,59	5,27	C2
4	1,63	2,50	8,36	C1
5	6,73	2,60	3,27	C2
6	4,73	0,59	5,27	C2
7	5,76	1,62	4,24	C2
8	7,82	3,68	2,18	C3
9	11,94	7,81	1,95	C3
10	12,97	8,84	2,98	C3
11	4,73	0,59	5,27	C2
12	2,49	6,62	12,48	C1

Figure 1. Centroid Distance

Calculating the second cluster center. The formula below can be used to calculate the center of the second centroid [11]:

$$R_{k=\frac{1}{N_k}}(X_{1k}+X_{2k}+\cdots+X_{nk}) \quad \text{.....(2)}$$

Rk = New average.

Nk = Number of training patterns in the cluster (k).

Xnk = The pattern to (n) that is part of the cluster (k).

The steps for determining the centroid are as follows:

1. C1 for the total attribute = the total number of selected attribute values in c1/the number of c1 in the total attribute.
2. C1 for the average attribute = the number of selected average attribute values in c1/the number of c1 on the average attribute
3. C2 for the total attribute = the total number of selected attribute values in c2/the number of c2 in the total attribute.
4. C2 for the average attribute = the number of selected average attribute values in c2/the number of c2 on the average attribute
5. C3 for the total attribute the total number of selected attribute values in c3/the number of c3 in the total attribute.
6. C3 for the average attribute = the number of selected average attribute values in c3/the number of c3 on the average attribute

Then the center of the new centroid is obtained as shown in the following table:

Table 4. The Second Centroid Center

Centroid 1	7,49	1,84
Centroid 2	3,40	0,90
Centroid 3	13,12	3,24

After calculating the new centroid value using the same formula as for the second centroid, the results are as follows:

Jarak ke Pusat Cluster

Pusat Cluster	C1	C2	C3	Keanggotaan Cluster
1	1,56	5,76	4,24	C1
2	1,53	2,67	7,33	C1
3	0,53	4,73	5,27	C1
4	2,56	1,63	8,36	C2
5	2,54	6,73	3,27	C1
6	0,53	4,73	5,27	C1
7	1,56	5,76	4,24	C1
8	3,62	7,82	2,18	C3
9	7,75	7,81	1,95	C3
10	8,78	8,84	2,98	C3
11	0,53	0,59	5,27	C1
12	6,68	6,62	12,48	C2
13	1,56	1,62	4,24	C1
14	1,53	1,47	7,33	C2

Figure 2. Clustering iteration 2

Because the pattern or membership is still insufficient, the next iteration calculation is performed, and we obtain:

Table 5. New Centroid

Centroid 1	7,49	1,84
Centroid 2	3,4	0,90
Centroid 3	13,12	3,24

Then the calculation of the first distance with Euclidean Distance for iteration 3 is obtained as follows:

Jarak ke Pusat Cluster

Pusat Cluster	C1	C2	C3	Keanggotaan Cluster
1	1,56	5,76	4,24	C1
2	1,53	2,67	7,33	C1
3	0,53	4,73	5,27	C1
4	2,56	1,63	8,36	C2
5	2,54	6,73	3,27	C1
6	0,53	4,73	5,27	C1
7	1,56	5,76	4,24	C1
8	3,62	7,82	2,18	C3
9	7,75	7,81	1,95	C3
10	8,78	8,84	2,98	C3
11	0,53	0,59	5,27	C1
12	6,68	6,62	12,48	C2
13	1,56	1,62	4,24	C1
14	1,53	1,47	7,33	C2

Figure 3. Clustering in the third iteration

The result of clustering is when the results of the previous clustering have the same pattern. Clustering did not occur until the third iteration because the pattern in the third iteration was identical to the pattern in the second iteration.

Table 6. Final Result of Clustering Calculation

Item	Iteration 1	Iteration 2	Iteration 3
ABC Eco Milk Coffe 230ml	2	1	1
Abc Exo chocomal 230ml	2	1	1
ABC Jus Apel UHT 250	2	1	1
ABC Jus Jambu UHT 250	1	2	2
ABC Jus Mangga UHT 250	2	1	1
ABC Jus Orange UHT 250	2	1	1
ABC Sari Kacang Hijau UHT 250	2	1	1
ABC SOURSOP 250ML	3	3	3
ABC SP Cocopandan 600	3	3	3
ABC SP Melon 585	3	3	3
Abc squash delight 525ml	2	1	1
ABC WHAIE COFFE	1	2	2
ACTIV WATER APPLE PEACH 380ML	2	1	1
ACTIV WATER LEMON 380ML	2	1	1
ACTIV WATER ORANGE 380ML	1	2	2
ACTIV WATER APPLE PEACH 380ML	2	1	1

The clustering process is complete because iterations 2 and 3 already have the same pattern. Cluster 1 (a total of 238 products) is the most popular beverage product, so existing stock must be increased. Cluster 2 with a total of 273 products can be interpreted as a group of beverage products of interest, with a moderate amount of stock that must be provided. Cluster 3, with a total of 25 products, can be interpreted as a product group that Angkasa Mart finds less appealing.

Table 7. Large Stock Grouping of Products

No	Item
1	ABC Jus Jambu UHT 250
2	ABC WHAIE COFFE
3	ACTIV WATER ORANGE 380ML
4	AMIA Botol 600 Ml
5	AQUARIUS 250ML
6	BUAVITA ORANGE 250ML
7	BUAVITA UHT Guava 250ml
8	chil go pisang 140ml
9	CHOCOLATOS DRINK RTD TETRA 200ML
10	CIMORY GRAPE 250ML
11	CIMORY Yd Plain 250Ml
12	CIMORY Yd Straw 250Ml
13	CLEVO Choco 125ML
14	COCA-COLA 250ML
15	COCO BIT SPLASH PEACH
17	COOLANT Cooling Water 350Ml
18	COOLANT Jus Belimbing
19	CRYSTALINE 1500ML
20	DBEST GREEN TEA 320ML
21	DBEST KUNDUR 350ML
22	DBEST LECY 350ML
23	DBEST TEH BUNGA350 ML
24	ENTRASOL Active Vanilla Latte 360gr
25	EX-GOAL SOYA 310ML
26	FRESTEA MARKISA 350 Ml
27	FRISIAN FLAG 40G
28	FRUIT TEA STOBERI350ml

Table 8. Medium Product Stock Grouping

No	Item
275	Abc Exo chocomal 230ml
276	ABC Jus Apel UHT 250
277	ABC Jus Mangga UHT 250
278	ABC Jus Orange UHT 250
279	ABC Sari Kacang Hijau UHT 250
280	Abc squash delight 525ml
281	ACTIV WATER APPLE PEACH 380ML
282	ACTIV WATER LEMON 380ML
283	ACTIV WATER APPLE PEACH 380ML
284	ACTIV WATER ORANGE 380ML
285	ADES
286	ADES 350ML
287	AIR MINERAL CLUB 600ML

288	AIR MINERAL INDODES 1500ML
289	AIR MINERAL INDODES 600ML
290	AIR SEJUK ADEM SARI 600ML
291	AMIA CUP 240 ML
292	AMUNIZER VIT C 1000
293	AQUA Botol 1500 ml
294	AQUA Botol 600ml
295	AQUA Gelas 240 ml Dus
297	BUAVITA Guava 1L
298	BUAVITA Mango Uht 250
299	CIMORY MANGO 250ML
300	CIMORY MIXED FRUIT 250Ml
301	CIMORY SQUEEZ STRAWBEERY 120ML
302	CIMORY SQUEEZE ALOE VERA 120ML
303	CIMORY TROPICAL FRUITS 250ML
304	CIMORY Yd Blueberry 250Ml
305	CIMORY Yd Lyche 250Ml
306	CIMORY Yd Mix Berry 250Ml
307	CLEVO COKLAT ICE CREAM
308	COCA-COLA 1500Ml
309	COCA-COLA Btl 390 Ml
310	COCA-COLA Can 250 Ml
311	COCO BIT SPLASH GUAVA
312	COOL TIME COCONUT WATER 200ML
313	DBEST CINCAU 320ML

Table 9. Low Product Stock Grouping

No	Item
512	ABC SP Cocopandan 600
513	ABC SP Melon 585
514	ACTIV WATER LEMON 380ML
515	ADES Air Mineral 1.5L
516	AIR MINUM GREEN
517	AMDK SMS Cup 220Ml
518	AMUNIZER DUS 4S
519	AQUA 330 ML
520	BUAVITA Mini Guava UHT 125ml
521	CIMORY LOW FAT 250 ML
522	CIMORY LOW FAT BANANA 250ML
523	CIMORY SQUEEZ BLUEBERRY 120ML
524	CIMORY SQUEEZ ORIGINAL 120ML
525	Crystalline 600Ml
526	FF CHOCO PET 270 ML
527	FRESTEA APPLE 1.5L
528	FRESTEA Green 500 Ml

529	FRESTEA Green Madu 500
531	FRESTEA Jasmine 350Ml
532	FRESTEA LYCHEE 350ML
533	FRISIAN FLAG KM 200G
534	JAHE GULA AREN
535	POCKY COLORFULL 36G
536	TEH SURYA CELUP 25'S

The K-Means process was also used in the study, as illustrated below:

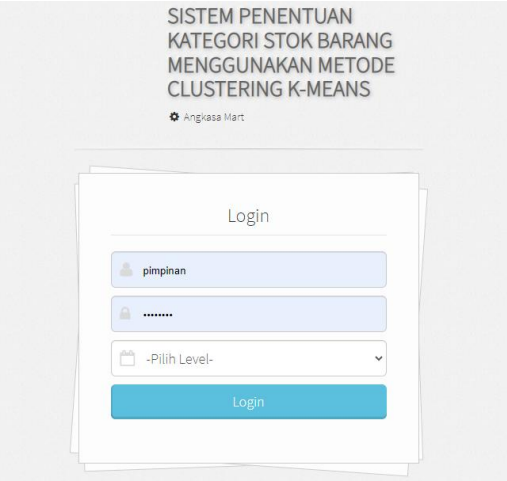

Figure 2. K-Means Login Page

In the K-Means process submenu, the K-Means Clustering process will be executed. This menu produces grouping calculations in order to find the same pattern.

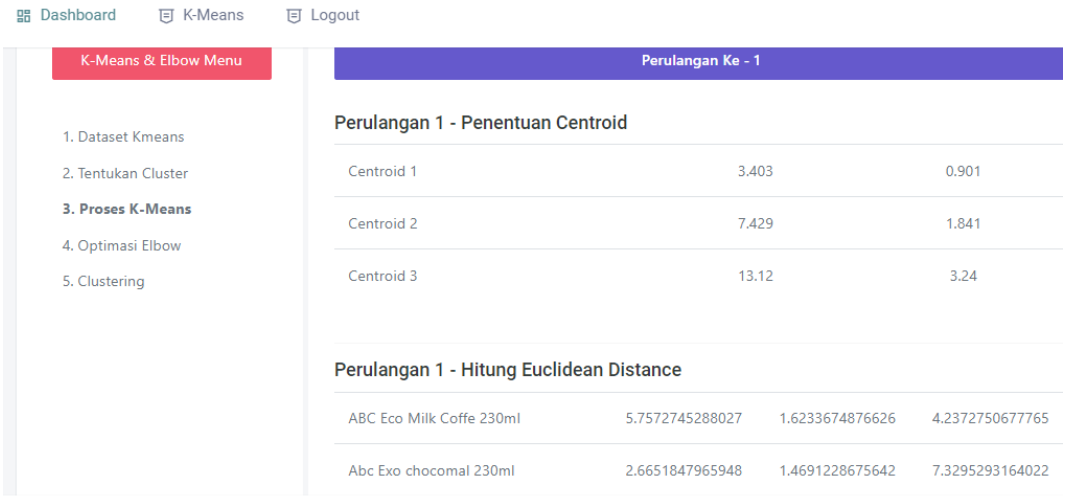

Figure 3. Clustering Process

The grouping produces a report on the leader's role. This menu displays the k-means process results, as shown in the figure below:

SISTEM INFORMASI CERDAS
Angkasa mart
PEKANBARU
Jalan Labersa Telp. 0341-210-0232
Email : admin@angkasa.or.id Kode Pos : 2828
Toko angkasa mart

LAPORAN DATA			
No	Nama produk	total	rata
537	TROPICANA Slim Diabetics 50'S	4	1
536	TROPICANA Slim Diabetics 25'S	3	0.75
535	TROPICANA Slim Diabetics 100'S	7	1.75
534	TROPICANA Slim Classic Ref 500	4	1
533	TROPICANA Slim Classic Ref 250	2	0.5
532	TROPICANA Slim Classic 50'S	7	1.75
531	TROPICANA Slim Classic 25'S	8	2
530	TROPICANA Slim Classic 100'S	5	1.25

Figure 4. Data Report

To perform clustering on Rapidmier, first, perform data input and the k-means algorithm. The trick is to type the data type and then select the file from the operation menu. It's the same with entering the K-Means algorithm.

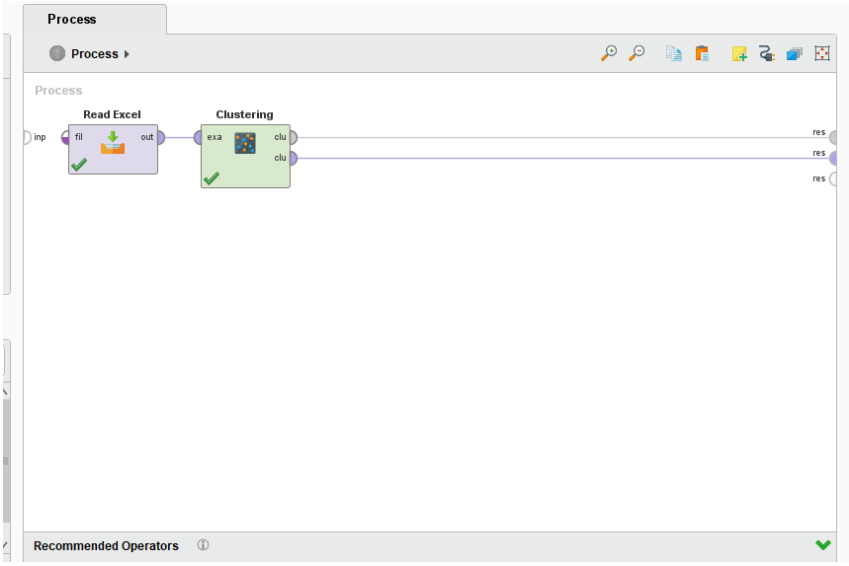

Figure 5. Input Data on Rapid Miner

After entering the data, press the run button to run it. The outcomes will be displayed in the results menu.

Result History | ExampleSet (Clustering) | Cluster Model (Clustering)

Data | Statistics | Visualizations | Annotations

Open in | Turbo Prep | Auto Model | Filter (536 / 536 exampl

Row No.	nama_barang	cluster	total	rata_rata
1	ABC Eco Milk ...	cluster_1	9	2.250
2	Abc Exo choc...	cluster_1	6	1.500
3	ABC Jus Apel...	cluster_1	8	2
4	ABC Jus Jam...	cluster_0	5	1.250
5	ABC Jus Man...	cluster_1	10	2.250
6	ABC Jus Ora...	cluster_1	8	2
7	ABC Sari Kac...	cluster_1	9	2.250
8	ABC SOURS...	cluster_2	11	2.750
9	ABC SP Coco...	cluster_2	15	3.750
10	ABC SP Melo...	cluster_2	16	4
11	Abc squash d...	cluster_1	8	2
12	ABC WHAIE ...	cluster_0	1	0.250
13	ACTIV WATE...	cluster_1	9	2.250
14	ACTIV WATE...	cluster_1	6	1.500
15	ACTIV WATE...	cluster_0	5	1.250

ExampleSet (536 examples, 2 special attributes, 2 regular attributes)

Figure 6. Clustering Result

Clustering results processed with Rapidminer software generate graphs for each stock item category. C1 are the blue ones, C2 are the green ones, and C3 are the red ones. As seen in the figure below.

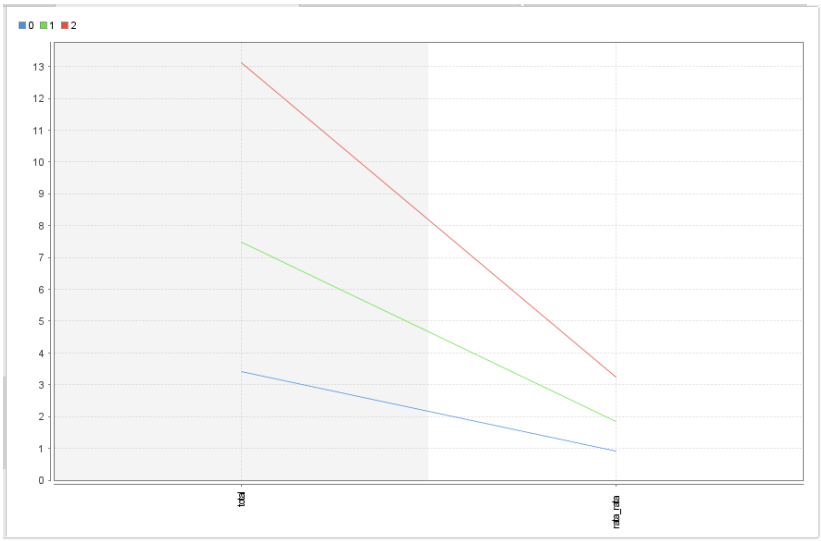

Figure 7. Clustering Chart

4. Conclusion

Angkasa Mart research was conducted using reference data of approximately 536 product data and 1184 transactions. Using the K-Means method by manual calculation, it was discovered that stock category data consisted of three groups, namely 238 product stock, 273 product stock, and 25 product stock. Calculations using the web-based k-means method to determine the category of stock items at the Angkasa Mart result in three groups: large product stock (238 items), medium product stock (273 items), and low product stock (25 items).

References

[1] Y. R. Sari, A. Sudewa, D. A. Lestari, and T. I. Jaya, "Penerapan Algoritma K-Means Untuk Clustering Data

Kemiskinan Provinsi Banten Menggunakan Rapidminer," *CESS (Journal Comput. Eng. Syst. Sci.*, vol. 5, no. 2, p. 192, 2020, doi: 10.24114/cess.v5i2.18519.

[2] M. L. Sibuea and A. Safta, "Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustring," *Jurteksi*, vol. 4, no. 1, pp. 85–92, 2017, doi: 10.33330/jurteksi.v4i1.28.

[3] D. N. Batubara, A. P. Windarto, D. Hartama, and H. Satria, "Analisis Metode K-MEANS Pada Pengelompokan Keberadaan Area Resapan Air Menurut Provinsi," *Semin. Nas. Sains Teknol. Inf.*, no. x, pp. 345–349, 2019, [Online]. Available: https://prosiding.seminar-id.com/index.php/sensasi/article/view/324 (05 Juni 2020).

[4] E. Muningsih and S. Kiswati, "Penerapan Metode K-Means Untuk Clustering Produk Online Shop Dalam Penentuan Stok Barang," *J. Bianglala Inform.*, vol. 3, no. 1, pp. 10–17, 2015.

[5] R. D. Ramadhani, "Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Universitas Dian Nuswantoro," *Ind. Mark. Manag.*, vol. 1, no. 1, pp. 1–9, 2014, [Online]. Available: http://dx.doi.org/10.1016/j.indmarman.2016.05.016.

[6] R. Ordila, R. Wahyuni, Y. Irawan, and M. Yulia Sari, "Penerapan Data Mining Untuk Pengelompokan Data Rekam Medis Pasien Berdasarkan Jenis Penyakit Dengan Algoritma Clustering (Studi Kasus : Poli Klinik PT.Inecda)," *J. Ilmu Komput.*, vol. 9, no. 2, pp. 148–153, 2020, doi: 10.33060/jik/2020/vol9.iss2.181.

[7] P. A. Ariawan, "Optimasi Pengelompokan Data Pada Metode K-means dengan Analisis Outlier," *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 2, pp. 88–95, 2019, doi: 10.25077/teknosi.v5i2.2019.88-95.

[8] R. Rosmini, A. Fadlil, and S. Sunardi, "Implementasi Metode K-Means Dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah," *It J. Res. Dev.*, vol. 3, no. 1, pp. 22–31, 2018, doi: 10.25299/itjrd.2018.vol3(1).1773.

[9] W. M. P. Dhuhita, "Clustering Menggunakan Metode K-Means Untuk," *J. Inform.*, vol. 15, no. 2, pp. 160--174, 2016.

[10] K. Fatmawati and A. P. Windarto, "Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi," *Comput. Eng. Sci. Syst. J.*, vol. 3, no. 2, p. 173, 2018, doi: 10.24114/cess.v3i2.9661.

[11] N. Rofiqo, A. P. Windarto, and D. Hartama, "Penerapan Clustering Pada Penduduk Yang Mempunyai Keluhan Kesehatan Dengan Datamining K-Means," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, pp. 216–223, 2018, doi: 10.30865/komik.v2i1.929.