

## **IMPLEMENTATION OF DATA MINING USING K-MEANS CLUSTERING METHOD TO DETERMINE SALES STRATEGY IN S&R BABY STORE**

**Tri Wahyudi<sup>1</sup>, Titi Silfia<sup>2\*</sup>**

Informatics Engineering Department, STIKOM Cipta Karya Informatika, Jakarta, Indonesia<sup>12</sup>  
fiviasilviatiti1101@gmail.com

Received : 15 August 2022, Revised: 02 September 2022, Accepted : 02 September 2022

\*Corresponding Author

---

### **ABSTRACT**

*The S&R Baby Store store is a Small and Medium Enterprise (SME) that is engaged in baby equipment, but there is a lot of competition between small and medium enterprises (SMEs) who are engaged in the same field, so that many products sold are of course not all sold out, some are lacking, in demand. Therefore the S&R Baby Store store needs a good sales strategy in order to increase sales profit. This study discusses the application of data mining, using the K-Means Clustering algorithm with the CRISP-DM method. Implementation using RapidMiner 9.10 which is done by entering sales transaction data with a total of 4 attributes and forming 4 clusters consisting of very in demand, in demand, moderate in demand and less in demand. the second cluster with 944 products, the third cluster with 2 products, and the fourth cluster with 43 products. The results of the cluster above are the products sold are the best-selling product categories, then the results of the cluster are validated using the Davies-Bouldin Index with a DBI value generated from clustering of 0.560.*

**Keywords :** Data Mining, K-Means, Clustering, Sales Strategy

### **1. Introduction**

S&R Baby Store is a Small and Medium Enterprise (SME) which is engaged in baby equipment. The S&R Baby Store store has been established since 2020 which focuses on selling various basic needs for babies ranging from milk, milk bottles, powder, cream, lotion, baby snacks, baby perfume, soap, Pampers, Shampoo and others. Along with the development of the times in the modern era where information technology (IT) is growing rapidly. In this business case, transactions that have occurred can be stored in digital form on a computer. The transaction process has also progressed in various ways and methods used to increase the sales profit of a product (Cartwright, et al., 2021; Setyowati, et al., 2021). However, not all of the products in the S&R Baby Store are selling very well, selling quite well and not selling well. The problem that often occurs is that there are many stocks of goods that have not been sold but are not paid attention to. So it is necessary to do several sales strategies to increase sales and minimize existing problems. Based on these problems, it is necessary to have the right solution for the problems being faced in determining the right sales strategy planning. the solution used to make it happen is to use data mining (Annur, 2019).

Data mining is a process of analyzing data activities to find a pattern from a data set, and the method used is the clustering method (Handoko, et al., 2020). Data mining can also be used to get a solution in a sales decision making to increase sales profit. Sales data storage can be seen by the large number of sales transaction records, where each record gives the product purchased by the customer for each sales transaction. The available data should be used as a decision-making system for business solutions and infrastructure support in the field of technology which is the cause of the emergence of a technology, namely data mining. The clustering method is a process of grouping data objects that are similar to each other into the same cluster and different from objects in other clusters (Bahar, et al., 2016; Hertina, et al., 2021).

The algorithm used in this study to determine the sales strategy is the K-Means algorithm. The K-Means algorithm is an algorithm that has become one of the most important algorithms in the field of data mining, because it has advantages as an algorithm that is easy to implement, relatively fast in terms of computational time and has been widely used to solve various computational problems (Rezaee, et al., 2021). is arguably the most popular grouping method. and if using the k-means method, it must transmit data so that it can be processed, then the data of

nominal data type must be initialized in the form of numbers first. and if the data is already in the form of numbers, it will be processed using Rapidminer tools. By doing this research, it is hoped that it can help the S&R Baby Store store in determining the right sales strategy to increase sales profits and be able to compete with competitors(Bramasta & Halilinta, 2021).

**2. Literature Review**  
**Survey Methodology**

According to experts, namely Kerlinger, stating that the methodology survey research is research conducted on large or small populations, but the data studied are data from samples taken from the population, to find relative events, distributions, and relationships between variables. sociological and psychological. In formulating the RQ on effectiveness in research, it should focus on 5 elements known as PICOC which can be seen in table 1. As follows:

Table 1 - PICOC table

<b>Survey Theme Title:</b> <b>Application of Data Mining Using K-Means Clustering Method to Determine Sales Strategy at S&amp;R Baby Store Stores.</b>	
<b>Population</b>	Data Mining Using K-Means Clustering Algorithm
<b>Intervention</b>	<ul style="list-style-type: none"> <li>a. How to apply the Data Mining method with the K-Means algorithm to determine which products are highly in demand, in demand, moderately in demand and not in demand?</li> <li>b. How can shop owners not have difficulty classifying what products are needed by customers and storage of data that is less effective?</li> <li>c. How can you find out which product categories are based on 4 clusters, namely very in demand, in demand, moderately in demand and not in demand which can be promoted so as to increase sales profit?</li> </ul>
<b>Comparison</b>	N/A
<b>Outcomes</b>	<ul style="list-style-type: none"> <li>a. Applying the concept of data mining using the K-Means clustering algorithm to determine which items are selling well, selling well and not selling well to increase sales profit.</li> <li>b. To be able to make it easier for the S&amp;R Baby Store Store in classifying sales products into the categories of very selling, selling well, selling enough and not selling well.</li> </ul>
<b>Context</b>	Private.

**Data Mining**

Data mining is an analysis of reviewing data sets to find unexpected relationships and summarizing data in different ways in a different way than before, which is understandable and useful for data owners. Data mining is a multidisciplinary field that combines techniques from machine learning, pattern recognition, statistics, databases, and visualization to solve problems of retrieving information from large databases. Data mining is a term used to describe the discovery of knowledge in databases. Data mining is a process that uses statistical techniques, mathematics, artificial intelligence, and machine learning to extract and identify useful information and knowledge bound from various large databases(Muttaqin, et al., 2022; Triyandana, et al., 2022).

**K-Means Algorithm**

K-means clustering is a non-hierarchical data clustering method that groups data in the form of one or more clusters/groups(Indriyani & Irfiani, 2019). Data that has the same characteristics are grouped into one cluster/group and data that has different characteristics are grouped with other clusters/groups so that data in one cluster/group has a small degree of variation(Pambudi & Witanti, 2022; Turnip & Fahmi, 2021; Borlea, et al., 2021). The steps for performing clustering with the K-Means method are as follows.

- a. Determine the value of k as the number of clusters to be formed.
- b. Determine the starting point of each cluster.
- c. Calculate the distance of each input data to each centroid using the distance formula

Euclidean (Euclidean Distance) until the closest distance is found from each data with centroids.

Here is the Euclidian Distance equation:

$$D(x, y) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2} \dots \dots \dots (1)$$

Information;

D = Distance x = Data y = Centroid

- d. Classify data based on their proximity to the centroid.
- e. Recalculate the cluster center with the current cluster members. The cluster center is the average value of all data objects in a particular cluster.
- f. Count each object again using the new cluster center. If the cluster center does not change again then the clustering process is complete. Or, go back to step number 3 until the cluster center doesn't change anymore.

**Davies-Bouldin Index (DBI)**

The Davies-Bouldin Index was introduced by David L. Davies and Donald W. Bouldin in 1979. Sum-of square within a cluster (SSW) is a metric of cohesion within a cluster. Separation by Sum-of-square-between-cluster (SSWB) by measuring the distance between the centroids □idanj.i,j is a measure of the ratio of how well the comparison value between the i-cluster and j-th cluster(Wijaya, et al., 2021; Ashari, et al., 2022).

The SSW formula is:

$$SSW = \frac{1}{N} \sum_{i=1}^N \|x_i - c_{p_i}\|^2$$

Rumus SSB :

$$SSB = \frac{2}{M(M-1)} \sum_{i=1}^M \sum_{j=1, j \neq i}^M \|c_i - c_j\|^2$$

Rumus R dan DBI

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}}$$

$$DBI = \frac{1}{K} \sum_{i=1}^K \max_{i \neq j} (R_{i,j})$$

**Sales Strategy**

A sales strategy is a tool that is fundamentally planned and designed. This design process is carried out as an effort by the company to develop competitive advantage through a special program in order to serve the market on an ongoing basis.

**3. Research Methods**

**Application of the Methodology**

In this study, the method used is K-Means Clustering. The data used in this study is sales transaction data for the S&R Baby Store Store in January-March 2022. At this stage of research, initial data collection will be carried out, the next process will be a Data Description, then after that an Evaluation of Data Selection is carried out, then the next stage is carried out Attribute selection, the next stage is to apply the CRISP-DM method which in this research process refers to the six stages of CRISP-DM, namely business understanding, data understanding, data preparation, modeling, evaluation and dissemination(Schröer, et al., 2021; Salt, 2021). The following are the stages of the research methodology which can be seen in Figure 1.

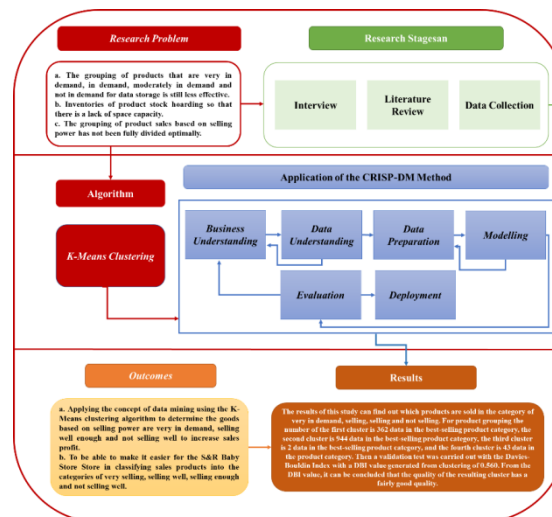


Fig 1. Stages of Methodology Application

**CRISP-DM method**

At this stage the approach used is the Cross Industry Standard for Data Mining (CRISP-DM) method. CRISP-DM is a method that uses a data development process model that is widely used by experts to solve problems. The research process refers to the six stages contained in the CRISP-DM, which are described as follows:

**Business Understanding**

Business understanding is done by studying the object of research, namely the S&R Baby Store. At this stage, the business objective of this research is to determine a sales strategy in order to increase sales profit. For this reason, product grouping will be carried out from the data obtained from observations in order to find out the products sold based on four clusters, namely very in demand, selling well, selling well and not selling well.

**Data Understanding**

Based on the sales transaction data that has been obtained, the next step in the CRISP-DM method is understanding the data needs related to achieving goals in determining effective and efficient sales strategies. The sales transaction data has a total of 1351 records with 10 attributes that will be carried out by selecting attributes into four attributes, namely the name of the item, the price of the item, the number of purchases, the total price. The data obtained in this study were obtained from January to March 2022. After the data was obtained, the next process was to understand the data.

**Data Preparation**

Data preparation includes all activities to build sales datasets that will be applied to modeling tools, from raw data in the form of sales transaction datasets and then data mining processes will be carried out. Its main function is specifically as a clustering modeling tool. Data preparation is a solid stage with data processing activities. The stage of creating a sales transaction dataset is the final data preparation stage for clustering modeling in data mining. In this process, the preparation of sales transaction data that has been carried out at the evaluation stage will be carried out and data improvements will be made based on the results of the evaluation. At this stage, business goals and data mining objectives have been set in determining product sales potential based on attributes,

**Modeling**

The modeling phase will use the clustering method with the K-means algorithm. In the application of the clustering method, it will be divided into 4 clusters which will classify sales products that are very in demand, in demand, moderately in demand and less in demand. The modeling tool used is Rapidminer.

### **Evaluation**

At this stage, an analysis or measurement of the accuracy of the modeling that has been carried out will be carried out. Evaluation is done by applying the Davies Bouldin Index (DBI) method, where this method is a quality test method based on the resulting cluster. The evaluation is intended to determine whether the modeling carried out is appropriate and appropriate to be applied to the case of this research and is in accordance with the initial plan of the study. Furthermore, from the results of the evaluation is to determine whether the next step can be continued or repeated from the beginning because it is not in accordance with the initial plan of the study.

### **Deployments**

The next stage is the stage of distributing the results of the research that has been carried out as a report or presentation of the knowledge that has been obtained based on modeling and evaluation of the data mining process.

## **4. Results and Discussions**

### **CRISP-DM**

This research was conducted using the CRISP-DM (Cross Industry Standard Process for Data Mining) method. In this method there are 6 stages used, namely:

#### **Business Understanding**

S&R Baby Store is a Small and Medium Enterprise (SME) that sells baby equipment. The business was initiated in 2020 and is located on Jl. Awareness, RT.1/RW.4, Pegangsaan Dua, Kec. Kelapa Gading, North Jakarta City, whose basic objective is to increase sales profit in determining the marketing strategy to be researched.

##### a. Define Business Goals

The business objective of this research is to determine which products are sold based on the best-selling, best-selling, moderate-selling and less-selling products. so that it can be used as a basis for increasing sales profits with decision makers to determine which products are in demand by customers consisting of products that are very in demand, in demand, moderately in demand and less in demand.

##### b. Assess the situation

S&R Baby Store is a small business whose main activity is selling various kinds of basic needs for babies, starting from baby milk, milk bottles, powder, cream, lotion, baby snacks, baby perfume, soap, pempers, shampoo and others. However, the current S&R Baby Store store has not utilized transaction data properly. Therefore, for this business, it is necessary to understand the sales strategy as in business goals and then translate it into data mining purposes.

##### c. Determine the Initial Data Mining Strategy

The initial strategy in this research is to review and request sales transaction data first at the S&R Baby Store Store.

#### **Data Understanding**

The dataset used in this study in the form of sales transaction data from January to March 2022 which is in the form of an excel document totaling 1351 records and 10 attributes. For data collection using sales transaction data consisting of its attributes, namely the date of sale, the name of the item, the purchase price of the item, the selling price of the item, the number of pieces, gross, net, profit and reports. Then the attribute selection is made into 4 attributes, namely, Item Name, Per item Price, Quantity, Total Price which will be processed into data mining. The following can be seen in table 2. the results of the selection of attributes.

Table 2 - Attribute Selection

Attributes	Description
Name of goods	The name of the item being sold.
Price per item	the price of goods sold is per item only.
Quantity	Quantity of goods sold
Amount	the sum of all products sold.

**Data Preparation**

Data preparation includes a sales transaction dataset that will be applied to the modeling tool, from the initial raw data in the form of a sales transaction dataset and then the data mining process will be carried out.

a. Data Selection

Where the raw data that will be selected in this study uses the attributes of the name of the item, the price of the item, the number of purchases, the total price at the S&R Baby Store.

b. Raw Data Processing (Preprocessed Data)

At this stage is the stage to ensure that the selected sales transaction data is feasible to be processed. After checking the data one by one and there is no data that has problems after cleaning, then the usage data is combined in one data cluster which will be tested in the K-means Algorithm process so that 4 clusters are used.

**Modeling**

Modeling is a stage that directly involves data mining techniques by selecting data mining techniques and determining the algorithm to be used. The data mining modeling in this study uses Excel for the calculation of the Euclidean Distance formula (Euclidean Distance) and for data processing software using Rapid Miner version 9.10. In this application the clustering algorithm is available in the form of the K-Means algorithm. The following is the implementation of the K-Means algorithm test which can be explained as follows.

a. Implementation Using Rapid Miner

Enter the 4 clusters specified in the clustering operator in the parameters. After that, click run to perform the test, it will result in testing in 4 clusters.

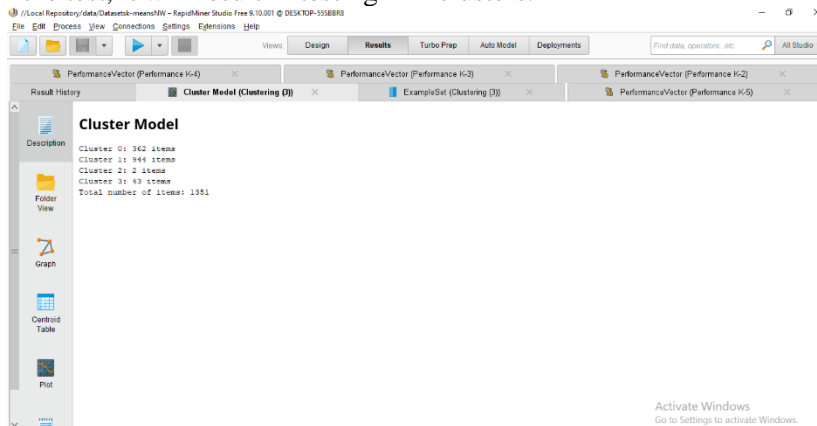


Fig 2. Display of Cluster Model Results

To see the results of the clustering, click Description on the display as shown above. Description to see the number of items contained in C0, C1, C2, and C3. Figure 2. is a display of the results of the cluster model from the processed data.

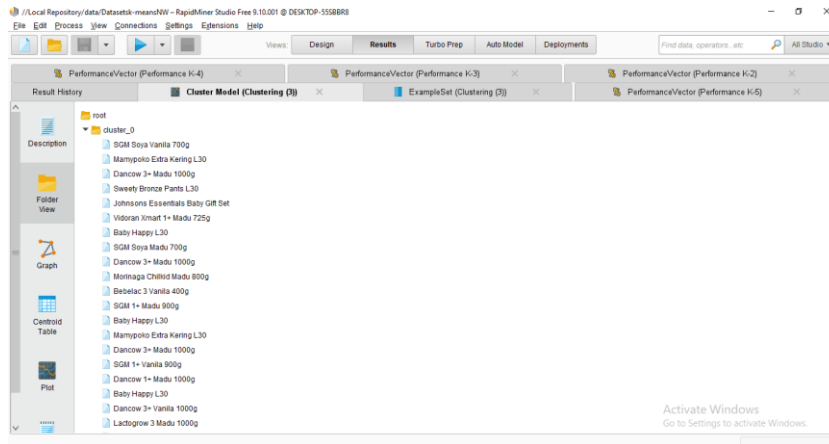


Fig 3. Display of Cluster 1 Results

The picture above is a display to see cluster 1 data that has been processed from Rapidminer which contains 362 data products.

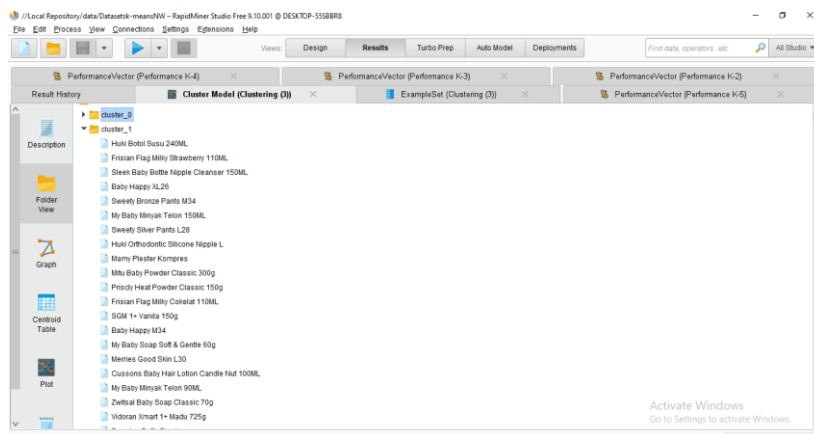


Fig 4. Display of Cluster 2 Results

Figure 4 is a display of results from cluster 2 that has been processed from Rapidminer which contains 944 data products.

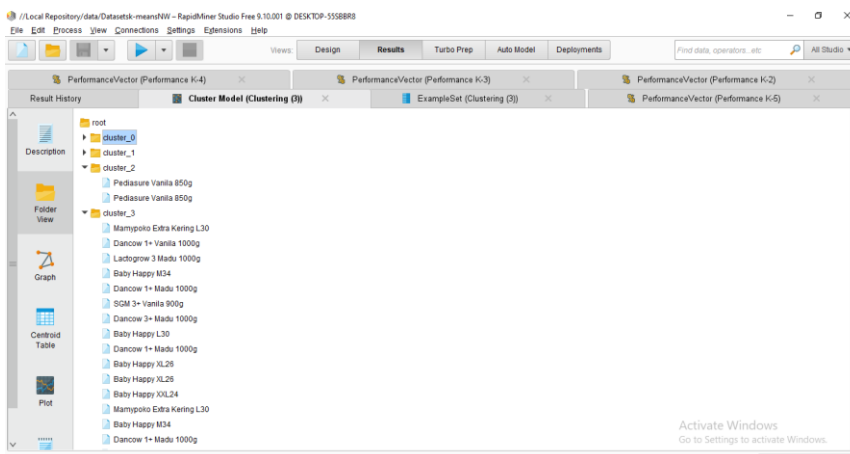


Fig 5. Display of the results of Cluster 3 and Cluster 4

Figure 5 is a display of results from cluster 3 that has been processed from Rapidminer which contains 2 data products. And the results of cluster 4 display which there are 43 data products.

Attribut	cluster_0	cluster_1	cluster_2	cluster_3
Harga Pentem	75.931	28.141	240	72.372
Jumlah	1.506	1.150	1	3.395
Harga Total	100.693	29.321	240	214.395

Fig 6. Centroid Table Display

Figure 6 is a display of results from the centroid table. In the centroid table there are 3 attributes and the results of the data processed from rapidminer are divided into 4 clusters.

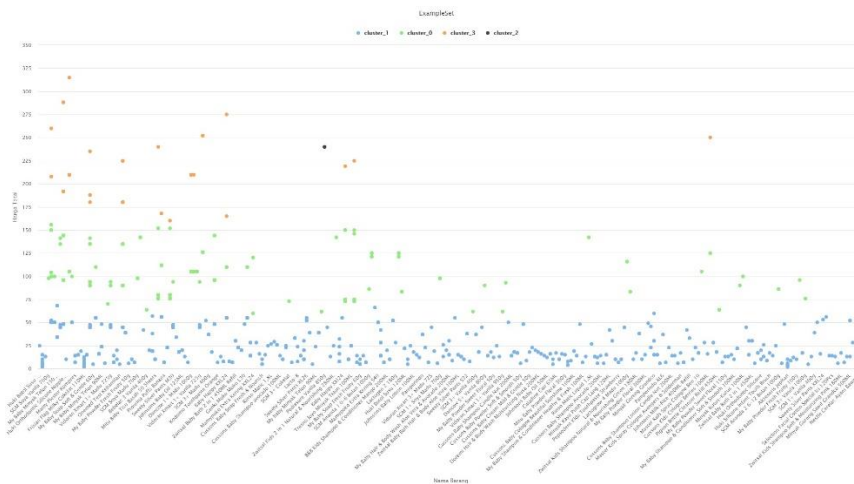


Fig 7. Results of Cluster Visualization in Rapidminer

Figure 7 is the result of Visualization Cluster on Rapidminer which is marked in blue as the result of cluster 1, green is the result of cluster 0 and orange is the result of cluster 3 and black is the result of cluster 4.

a. Calculation with K-Means Algorithm

Table 3 - Initial Centroid

Cluster to-	Selling Price	Quantity Of Purchase	Amount
C1	45000	4	180000
C2	80000	3	240000
C3	48000	3	144000
C4	45000	1	45000

Table 3 is Determining the Initial Cluster Center randomly taken from sales transaction data. The data selected for the initial cluster are the data of 292, 600, 955, 489. The distance of the centroid of the 1st data in cluster 1 (C1) is:



$$\sqrt{(SP_1 - C1_1)^2 + (QOP_2 - C2_2)^2 + (AM_3 - C3_4)^2} = \dots$$

Cluster 2

$$\sqrt{(25000^1 - 45000^1)^2 + (1^2 - 4^2)^2 + (25000^3 - 18000^3)^2} = 181.492,5$$

Cluster 3

$$\sqrt{(25000^1 - 80000^1)^2 + (1^2 - 3^2)^2 + (25000^3 - 24000^3)^2} = 295.013$$

Cluster 4

$$\sqrt{(25000^1 - 48000^1)^2 + (1^2 - 3^2)^2 + (25000^3 - 14400^3)^2} = 85.250$$

Cluster 4

$$\sqrt{(25000^1 - 45000^1)^2 + (1^2 - 1^2)^2 + (25000^3 - 45000^3)^2} = 25.000$$

**Evaluation**

Evaluation is an advanced stage of the purpose of data mining. Evaluation is carried out in depth with the aim that the results at the modeling stage are in line with the objectives to be achieved at the business understanding stage. Evaluation of the results of this stage assesses the extent to which the data mining modeling results meet the data mining objectives that have been determined at the business understanding stage. Where in testing the data that has been processed in Rapid Miner. Testing data on Rapid Miner used in this study is index.

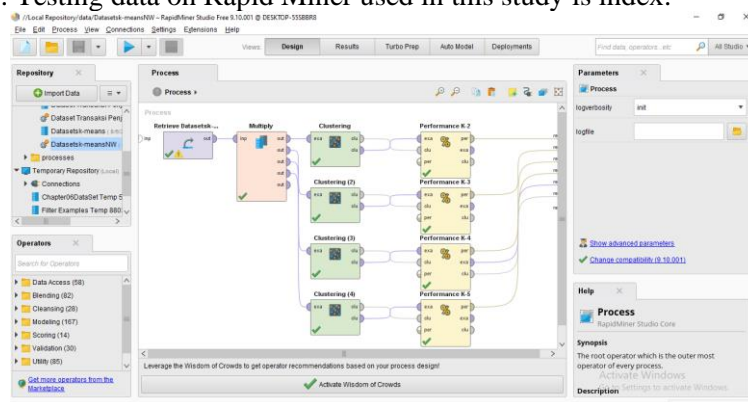


Fig 8. Display in the testing process

From Figure 8, the testing process is carried out with the multiply operator because the goal is to be connected to several operators, then the clustering operator is duplicated into four operators to form four clusters. In the first clustering operator it becomes 2 clusters, for the second clustering operator it is formed into 3 clusters then for the third clustering operator it is formed into 4 clusters and for the fourth clustering operator it is formed into 5 clusters, then connected to a distance performance cluster with crossed connectors. After all operators are connected to the distance performance cluster, change the cluster distance performance parameter on the right to change to Davies-Bouldin Index and then the test is successful, After that, click run, the results of the test data will appear which will show the Davies-Bouldin Index value for each performance in the four clusters. The following is a comparison of the Davies-Bouldin Index values based on 4 clusters which can be seen in table 3

Table 4 - Davies-Bouldin Index Value of Each Cluster Model

	2 Clusters	3 Clusters	4 Clusters	5 Clusters
Davies Bouldin Index	-0.700	-0.692	-0.560	-0.586

From table 4 it can be concluded that the result of the Davies-Bouldin Index value in the smallest performance is in cluster 4. For the DBI value of cluster 4 is -0.560. while for performance cluster 2 there is a DBI . value-0.700, the performance of cluster 3 has a DBI value of -0.692, and for the performance of cluster 5 there is a DBI value of -0.586, so from the results

of the cluster the DBI value has a fairly good quality is cluster 4 because the smaller the value of the cluster, the better and optimal. The following results from distance performance cluster 4 can be seen in Figure 9.

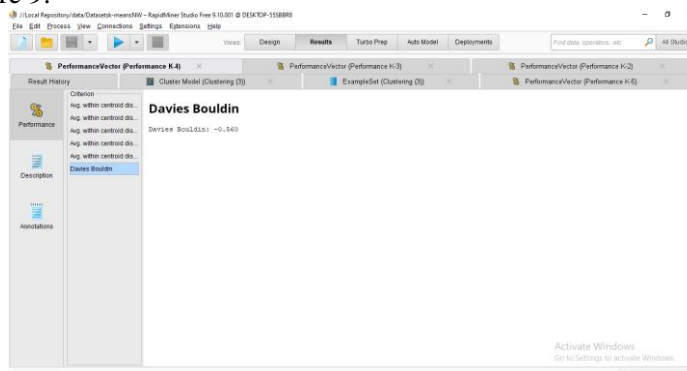


Fig 9. Display of the results of the Davies Bouldin Index (DBI)

## Deployment

Deployment is the final stage in making reports on the results of data mining activities. The final report containing the knowledge gained or pattern recognition in the data mining process. Here are some cluster analyzes in this study which can be explained as follows. There are 4 attributes used in the clustering process, namely the name of the item, the item price, the number of purchases, and the total price. There are also 4 clusters used, cluster 0 is taken from the best-selling products sold, cluster 1 is taken from products sold in the best-selling category, cluster 2 is taken from the category of products that sell well enough and cluster 3 is taken from the category of sold products that are not selling well. Here are the clustering results:

1. cluster 0, the products sold in the very best-selling category consist of 362 products.
2. Cluster 1, the best-selling products category consists of 944 products.
3. Cluster 2, the product sold in the category of selling well, consisting of 2 products.
4. Cluster 3, the product sold in the category of not selling well consists of 43 products.

Which is where the total of all products is 1351 products, and of the 4 clusters of products that are mostly sold in the best-selling category in cluster 1.

## 5. Conclusion

So from the results of the study, it can be concluded that the K-means method can be applied to the S&R Baby Store Store to determine which sales of baby equipment products are in demand, selling well, selling well and not selling well. Application of K-Means method on S&R Baby Store namely by grouping sales transaction data. Then choose 4 clusters randomly as the initial centroid. After the data in each cluster does not change, it can be seen that the final result is that there are 362 products that are selling very well, 944 products that are selling well, those that are selling well, consisting of 2 products and 43 products that are not selling well. Applying the K-means method to Rapidminer is done by entering sales transaction data, namely the name of the item, the price of the item, the number of purchases and the total price which will become a database on Ms. Excel, the data is then connected to the Rapidminer Tools, and will be processed and formed K-means. means. After that Rapidminer will produce which products are very in demand, in demand, moderate in demand and less in demand.

## References

- Annur, H. (2019). Penerapan Data Mining Menentukan Strategi Penjualan Variasi Mobil Menggunakan Metode K-Means Clustering. *Jurnal Informatika Upgris*, 5(1).
- Ashari, I. F., Banjarnahor, R., Farida, D. R., Aisyah, S. P., Dewi, A. P., & Humaya, N. (2022). Application of Data Mining with the K-Means Clustering Method and Davies Bouldin Index for Grouping IMDB Movies. *Journal of Applied Informatics and Computing*, 6(1), 07-15.
- Bahar, A., Pramono, B., & Sagala, L. H. S. (2016). Penentuan Strategi Penjualan Alat-Alat Tattoo

- di Studio Sonyxtattoo Menggunakan Metode K-Means Clustering. *Seman TIK*, 2(2).
- Borlea, I. D., Precup, R. E., Borlea, A. B., & Iercan, D. (2021). A unified form of fuzzy C-means and K-means algorithms and its partitional implementation. *Knowledge-Based Systems*, 214, 106731.
- Bramasta, F. A., & Halilintar, R. (2021, August). Penerapan Data Mining Untuk Menentukan Strategi Penjualan Toko Sepatu. In *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)* (Vol. 5, No. 2, pp. 236-241).
- Cartwright, S., Liu, H., & Raddats, C. (2021). Strategic use of social media within business-to-business (B2B) marketing: A systematic literature review. *Industrial Marketing Management*, 97, 35-58.
- Handoko, S., Fauziah, F., & Handayani, ETE (2020). Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering. *Jurnal Ilmiah Teknologi dan Rekayasa*, 25(1), 76–88. <https://doi.org/10.35760/tr.2020.v25i1.267>
- Hertina, H., Nurwahid, M., Haswir, H., Sayuti, H., Darwis, A., Rahman, M., ... & Hamzah, M. L. (2021). Data mining applied about polygamy using sentiment analysis on Twitters in Indonesian perception. *Bulletin of Electrical Engineering and Informatics*, 10(4), 2231-2236.
- Indriyani, F., & Irfiani, E. (2019). Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means. *JUITA: Jurnal Informatika*, 7(2), 109-113.
- Muttaqin, M. R., Hermanto, T. I., & Sunandar, M. A. (2022). Penerapan K-Means Clustering dan Cross-Industry Standard Process For Data Mining (CRISP-DM) untuk Mengelompokan Penjualan Kue. *KOMPUTASI*, 19(1), 38-53.
- Pambudi, W. T., & Witanti, A. (2022). Penerapan Algoritma K-Means Untuk Menganalisis Data Penjualan Pada Toko Ayu Collection Berbasis Web. *Jurnal Sistem Informasi dan Bisnis Cerdas*, 15(1), 1-9.
- Rezaee, M. J., Eshkevari, M., Saberi, M., & Hussain, O. (2021). GBK-means clustering algorithm: An improvement to the K-means algorithm based on the bargaining game. *Knowledge-Based Systems*, 213, 106672.
- Saltz, J. S. (2021, December). CRISP-DM for Data Science: Strengths, Weaknesses and Potential Next Steps. In *2021 IEEE International Conference on Big Data (Big Data)* (pp. 2337-2344). IEEE.
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, 181, 526-534.
- Setyowati, W., Widayanti, R., & Supriyanti, D. (2021). Implementation Of E-Business Information System In Indonesia: Prospects And Challenges. *International Journal of Cyber and IT Service Management*, 1(2), 180-188.
- Triyandana, G., Putri, L. A., & Umaidah, Y. (2022). Penerapan Data Mining Pengelompokan Menu Makanan dan Minuman Berdasarkan Tingkat Penjualan Menggunakan Metode K-Means. *Journal of Applied Informatics and Computing*, 6(1), 40-46.
- Turnip, H. N., & Fahmi, H. (2021). Penerapan Data Mining Pada Penjualan Kartu Paket Internet Yang Banyak Diminati Konsumen Dengan Metode K-Means. *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, 4(2), 36-41.
- Wijaya, Y. A., Kurniady, D. A., Setyanto, E., Tarihoran, W. S., Rusmana, D., & Rahim, R. (2021). Davies Bouldin Index Algorithm for Optimizing Clustering Case Studies Mapping School Facilities. *TEM J*, 10(3), 1099-1103.