JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 321 – 328 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3774
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

MODELING CLOTHING ORDER SIZE GROUPING AT RIZKY CONVECTION USING THE K-MEANS METHOD

Putri Rahmadani[1*], Jeperson Hutahaean[1], Santoso [1]
[1]Information System, Universitas Royal
email: *rahmadaniputri450@gmail.com

Abstract: Rizky Convection Business is a sportswear production company based on Jalan Elang Lestari Kisaran. Every day they receive orders from schools and agencies that need sportswear. However, Rizky Convection often faces challenges, especially in managing raw material inventory, allocating production time and processing orders that come in large quantities. For this reason, order data needs to be grouped to make it easier for employees to work on it. The data that will be processed in this study is order data from 2022-2024. One way to do this is to apply data mining techniques, one of which is the K-Means Clustering method. The purpose of this study is to model the use of K-Means Clustering to improve production management and procurement of raw materials for fabrics at Rizky Convection. K-Means Clustering is the grouping of a number of data into clusters (groups) so that each cluster will contain data that is as similar as possible. The results of K-Means Clustering grouping with 3 clusters, namely cluster 1, the large order group has 25 order data, cluster 2, the small order group has 361 order data and cluster 3, the medium order group has 100 order data.

Keywords: data mining; order data; K-Means; sportswear manufacturing.

Abstrak: Usaha Konveksi Rizky merupakan perusahaan produksi pakaian olahraga yang berpusat di Jalan Elang Lestari Kisaran. Setiap hari mereka menerima pesanan dari sekolah dan instansi yang membutuhkan pakaian olahraga. Namun Konveksi Rizky sering menghadapi tantangan terutama dalam pengelolaan persediaan bahan baku kain, pengalokasian waktu produksi dan pemrosesan pesanan yang datang dalam jumlah banyak. Untuk itu, data pesanan perlu dikelompokkan untuk mempermudah karyawan dalam mengerjakannya. Data yang akan diproses pada penelitian ini adalah data pesanan dari tahun 2022-2024. Salah satu cara untuk hal tersebut adalah dengan menerapkan teknik data mining, salah satunya metode K-Means Clustering. Tujuan penelitian ini adalah untuk memodelkan penggunaan K-Means Clustering untuk meningkatkan manajemen produksi dan pengadaan bahan baku kain di Rizky Konveksi. K-Means Clustering adalah pengelompokan sejumlah data ke dalam cluster (group) sehingga setiap dalam cluster tersebut akan berisi data yang semirip mungkin. Hasil pengelompokkan K-Means Clustering dengan 3 cluster yaitu cluster 1 kelompok pesanan banyak memiliki 25 data pesanan, cluster 2 kelompok pesanan sedikit memiliki 361 data pesanan dan cluster 3 kelompok pesanan sedang memiliki 100 data pesanan.

Kata kunci: data mining; data pesanan; K-Means; konveksi.

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 321 – 328 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3774
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

INTRODUCTION

Data mining is a method that enables users to access large amounts of data in a relatively short amount of time [1]. Data mining is the activity of discovering valuable patterns and information hidden within large datasets by applying specific techniques and methods. In data mining, many techniques are used, one of which is K-Means Clustering [2]. Clustering is an algorithm that analyzes data with the aim of categorizing items with similar values into similar regions. K-Means Clustering is the process of grouping similar data from a larger dataset to find and cluster data with similar characteristics (similarity) between one data point and another [3].

One of the industries that can benefit from this technique is the industry convection, such as Usaha convection Rizky a sportswear production company based on Jalan Elang Lestari Kisaran [4]. Every day, they receive orders from schools and organizations that need sportswear and sports t-shirts. These orders are also available in various sizes, including S, M, L, XL, XXL, XXXL, and XXXXL, depending on the size desired by the customer. After receiving an order, the employees start making a sample to measure the basic material, cut it according to the sample, and sew it into clothing. The orders are processed according to the availability of fabric raw materials and customer orders.

If the fabric raw material runs out, the Konveksi must order it first, which takes a long time. Therefore, the Konveksi negotiates with customers by offering a different color or asking them to wait until the order is processed as scheduled. However, this is not always effective, especially when large orders arrive at the same time and require different sizes.

Konveksi Rizky often faces challenges, especially in managing fabric raw material inventory, allocating production time, and processing large orders. These challenges become more complex when orders come in various sizes, requiring careful planning to avoid material shortages and delivery delays. This problem occurs because there is no system to predict fabric raw material inventory.

Therefore, Rizky Convection needs a system to predict raw material requirements and optimize production to ensure timely order fulfillment [5]. This is useful for grouping order patterns based on the most frequently ordered sizes. It gives Konveksi Rizky better insight into customer preferences and more accurate raw material needs [6].

The application of data mining can help Konveksi Rizky plan production and raw material procurement more efficiently and reduce the risk of delays. One method that can be applied is Clustering with the K-Means algorithm. By using order data as the basis for analysis, the company can develop a better production strategy [7].

Based on research from the journal titled "Determination of Sales Grouping with a Combination of K-Means and Hamming Distance," this study was conducted to determine which products are selling well and which are not. The program created can generate a report listing the 10 best-selling products and the 10 least-selling products. The results of this study can be used as a consideration for determining product production and marketing strategies [8].

"Implementation of the K-Means Clustering Method for Grouping Online Store Product Stock Data" was conducted

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 321 – 328 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3774
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

to analyze product stock sales in the Perdagangan Kaos store using the K-Means method, which can generate groupings [9].

"Application of Data Mining in Dameyra Fashion Product Sales Using the K-Means Clustering Method" was conducted to determine which products are in high demand and which are not. The results are generated by the K-Means clustering output using the RapidMiner tool [10].

"Implementation of Data Mining on Clothing Sales Data Using the K-Means Algorithm with Optimize Parameter Grid" was conducted to determine the optimal k value based on parameters. The Optimize Parameter Grid operator is needed to speed up the grouping of sales data and identify characteristics in the dataset using the ParameterMixed Measure. This research uses the Data Mining Method with the K-Means Clustering Algorithm and Optimize Parameter Grid [11].

"Clustering Clothing Sales to Improve Sales Strategy Using K-Means" This research was conducted for to help improve effective sales strategies by utilizing available sales data. The sales data is grouped based on similarities so that data with the same characteristics are in one cluster. One of the attributes used is clothing type and sales [12].

The study titled "Clustering Sales Data Based on Region Using the K-Means Method at PT XYZ" was conducted to help the company manage its sales transaction data and determine the sales performance of each branch. The K-Means clustering algorithm and the Python programming language were used in this research. The data will be clustered into three groups: highest sales, medium sales, and lowest sales [13].

The study titled "Application of the K-Means Clustering Algorithm for Grouping Clothing Sales Data (Case Study: UMKM Lima Media Kuningan)" was conducted to classify underwear sales data, which can be used for further analysis, such as inventory planning for the next period. The method used is data mining with the K-Means Clustering algorithm to group data into several clusters, allowing the identification of the most popular underwear among customers using the RapidMiner application version 10.3 [14].

Based on this, before implementation the convection company often ran out of stock due to irregular order processing. After grouping the order data, the production process became more organized, and there was less waste of fabric raw materials. Additionally, the company can identify which customers frequently order the most clothing and which fabric materials they prefer.

The implementation of K-Means Clustering also improves production efficiency and customer satisfaction. This study aims to examine the application of this technique for production management and raw material procurement at Konveksi Rizky.

METHOD

Data mining is an integrated part of knowledge discovery in databases. Data mining is also referred to as knowledge discovery in databases (KDD) or pattern recognition. The term KDD, or knowledge discovery in data, refers to the main objective of data mining, which is to utilize data in a database by processing it to generate new, useful information. Meanwhile, the term pattern recognition

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 321 – 328 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3774
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

refers to the goal of discovering knowledge from the data set being analyzed [15].

The K-Means Clustering algorithm is a data analysis method or data mining technique that performs unsupervised modeling and is one of the methods used to cluster data through a partitioning system. K-Means is an iterative clustering algorithm. The K-Means algorithm assigns cluster values (k) randomly, where these values become the center of the cluster, known as the centroid, mean, or means[16].

The steps of performing clustering or grouping using the K-Means method can be explained as follows:

Determining Clusters

Determine the number of clusters, we can use the Elbow Method by testing data in the Anaconda Jupyter Notebook. This method helps us find the "elbow point," where the Within-Cluster Sum of Squares (WCSS) starts to decrease more slowly.

Select Centroid

Select and set the centroid or cluster center randomly. The initial centroid is determined based on assumptions.

Calculate The Distance Between The Data and The Cluster

Assign all data points to the nearest cluster. The proximity of two data points is determined based on the distance between the two data points and the calculation of the distance from the data to the cluster centroid. To calculate the distance of all data points to each cluster centroid, the Euclidean distance formula can be used, which is expressed as follows:

$$D(i,j) = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \ldots + (xk_i - xk_j)^2} \ldots (1)$$

Description:
D(i,j) : The distance of data point i to cluster centroid j.
X_{ki} : Data point i on attribute k,
X_{kj} : Cluster centroid j on attribute k.

Group The Data

Classify each data point based on its proximity to the centroid (the smallest distance).

Calculate The New Centroid Value

Update the centroid values. The new centroid value is obtained from the average of the corresponding cluster using the following formula:

$$Vij = \frac{1}{N_i} \sum_{k=0}^{N_i} Xk_j \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Description:
Vij : The centroid / the average of cluster i for variable j.
N_i : The number of data points that belong to cluster i.
i : k are the indices of the clusters,
j : The index of the variable,
X_{kj} : The value of data point k within the cluster for variable j.

Repeat the process until the members of each cluster do not change[17]. The initial process of this research assumes that there are three centroids or cluster points, following previous studies. The following table contains sample data used to perform manual calculation experiments.

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 321 – 328 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3774
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

Table 1. Sample Data Calculated

No.	Customer	S	M	L	XL	XXL	XXXL	XXXXL
1	SDN 010174 Sei Balai	2	22	18	2	5	0	0
2	SDN 015603 Tanah Rakyat	10	18	16	4	0	0	0
3	MIS Pendidikan Islam	0	28	27	10	0	0	0
4	MIS AW Desa Durian	4	30	0	0	0	0	0
5	SDN 014697 Banjar	4	38	54	24	4	1	0
6	SDN 132402 Tanjung Balai	0	24	10	7	10	0	0
7	MIS Jabat Hindi	2	15	20	0	0	0	0
8	MTS Jabat Hindi	0	5	76	25	6	0	0
	….....	…	….	…	….	…….	…….	....
483	MTS Taufik Kamil	0	0	10	14	0	0	0
484	MAS Cipta	0	0	17	111	22	0	0
485	SDN 132407 TJ.Balai	0	10	13	4	0	0	1
486	MTS Huswatun Hasana	0	0	10	3	2	0	0

At this stage, a manual calculation of Data Mining is performed using K-Means Clustering. There are several steps to calculate using K-Means Clustering, including determining clusters, which consist of 3 clusters based on the available data. In the research, the first cluster or cluster 0 is categorized as a high-order cluster, the second cluster or cluster 1 is categorized as a low-order cluster, and the third cluster or cluster 2 is categorized as a medium-order cluster.

For the initial cluster centers, it is assumed that the first cluster is data number 52, the second cluster is data number 189, and the third cluster is data number 452.

Table 2. Assumed Centroid

C	S	M	L	xl	xxl	xxxl	xxx xl
C0	10	30	66	177	44	1	2
C1	5	35	35	15	10	10	10
C2	33	22	48	27	33	17	0

After calculating using the Euclidean distance equation, clusters are formed in the first iteration. Then, calculate the next centroid for each cluster using the centroid average formula, which will be used for the second iteration. The first centroid is as follows:

Table 3. First Centroid

C	S	M	L	xl	xxl	xx xl	xxx xl
C0	1, 11	19, 93	95, 96	148, 25	35, 71	5	3,82
C1	1, 75	16, 79	19, 93	11,04	3,31	0, 60	0, 15
C2	2, 72	16, 32	62, 28	59,84	25, 4	5, 68	0, 40

Do this until the same members are found in an iteration. In this study, the iteration stops at iteration 7. The iteration stops because the members in iteration 6 and iteration 7 are the same.

RESULTS AND DISCUSSION

This research successfully grouped order data based on clothing sizes using the K-Means Clustering method. The analysis results showed that customer

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 321 – 328 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3774
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

orders could be divided into several clusters, where each cluster represents a different pattern of clothing size demand. With this clustering, Konveksi Rizky can understand which sizes are most frequently ordered and which are less popular.

In this study, the visualization of order data can be seen using Jupyter Notebook as follows:

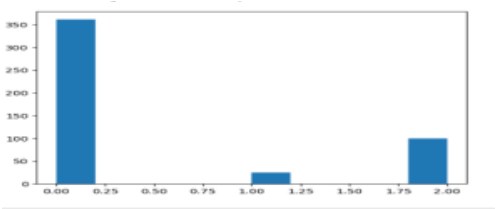

Image 1. Visualization

In the image, you can see three groups. The first diagram represents the "High Orders" group, the smallest diagram represents the "Low Orders" group, and the last one represents the "Medium Orders" group. In addition, we can also see the data distribution as follows:

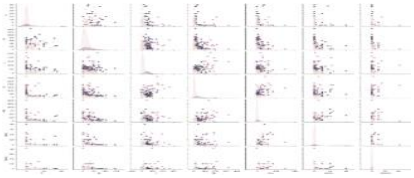

Image 2. Visualization of Order Data Distribution

From this data distribution, the K-Means Clustering calculation for each cluster was obtained. The following is the high-order cluster at Konveksi Rizky, which has 25 members, as shown below:

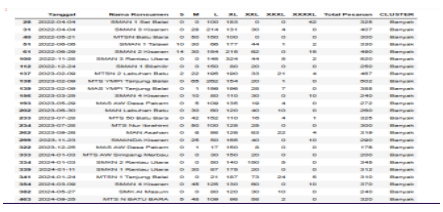

Image 3. Multiple Order Cluster

The medium-order cluster with 100 members from this research is as follows:

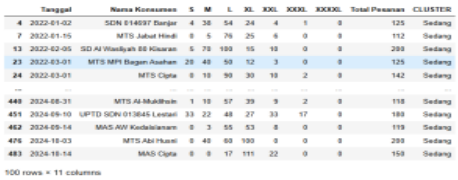

Image 4. Medium Order Cluster

Finally, the low-order cluster with 361 members is as follows:

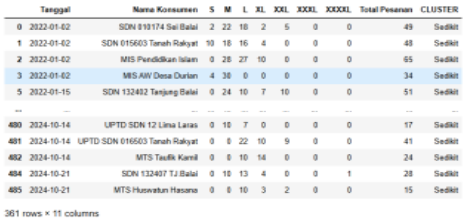

Image 5. Small Order Cluster

The conclusion from the K-Means Clustering calculation is as follows:

Tabel 4. Members of Clustering Results

Cluster	Number of Members
Cluster 0	25
Cluster 1	361
Cluster 2	100

To make it easier for the owner to group orders, this research has developed a website that simplifies the clustering process. The image 6, namely the clustering result display, showing the members and calculations from this research:

Image 6. Clustering Results

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 321 – 328 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3774
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

This website also generates reports that can be used for evaluation every year. The report display is as follows:

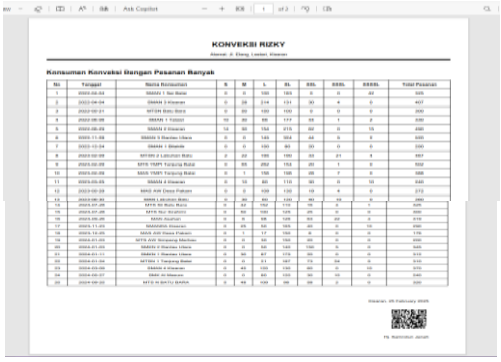

Image 7. Clustering Results Report

CONCLUSION

That this research successfully identified the demand patterns for sportswear orders at Rizky Convection using the K-Means Clustering method, which grouped the orders into three categories: high, medium, and low orders. The results of this clustering can support Rizky Convection in creating a more effective fabric procurement plan, helping to prevent shortages or excess inventory, and speeding up the production process. With the information from these demand patterns, the convection can more accurately project fabric material needs and reduce the likelihood of production delays.

BIBLIOGRAPHY

[1] A. Saputra,) ; Herlina, L. Sari, and D. Sartika, "implementasi metode association rule mining pada penjualan barang di toko bangunan ada mas menggunakan algoritma apriori," *Jurnal Multidisiplin Dehasen*, vol. 2, no. 4, pp. 709–718, Oct. 2023.

[2] P. Marpaung, I. Febrian, W. Putri, and E. P. Korespondensi, "Penerapan Data Mining Dalam Menentukan Tingkat Kedisiplinan Karyawan Perhotelan Menggunakan Algoritma K-Means Clustering," *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI*, vol. 7, no. 1, pp. 167–172, 2024.

[3] N. F. Adani, A. F. Boy, S. Kom, M. Kom, and R. Syahputra, "Implementasi Data Mining Untuk Pengelompokan Data Penjualan Berdasarkan Pola Pembelian Menggunakan Algoritma K-Means Clustering Pada Toko Syihan," *Jurnal CyberTech*, vol. x. No.x, Sep. 2021, [Online]. Available: https://ojs.trigunadharma.ac.id

[4] K. Butar-Butar, E. Husnita Hasibuan, R. Desti Dwi Meilasari, A. Murni, and N. Wahyuni, "Penyuluhan Tentang Pelaksanaan Quality Control (Qc) Pada Hasil Produksi Pakaian Jadi Di Konveksi Angkola Jaya Tapanuli Selatan," *Jurnal Pengabdian Kepada Masyarakat*, vol. 6, no. 1, 2022.

[5] E. Ummi Awaliyatul Ummah *et al.*, "Identifikasi Produksi Garment Terhadap Jumlah Reject Mati Dengan Metode Pareto," *Jurnal Komisi (Komputer dan Sistem Informasi)*, vol. 1, no. 3, pp. 83–89, 2024, [Online]. Available: https://komisijournal.indiepress.id/index.php/komisi/index

[6] R. Y. Firmansah, J. Dedy Irawan, and N. Vendyansyah, "Analisisrfm (Recency, Frequency And Monetary) Produk Menggunakan Metode K-Means," *Jurnal Mahasiswa Teknik Informatika)*, vol. 5, no. 1, 2021.

[7] S. Pujiono, R. Astuti, and F. Muhamad Basysyar, "Implementasi Data Mining Untuk Menentukan Pola Penjualan Produk Menggunakan Algoritma K-Means Clustering," 2024.

[8] R. T. Vulandari and D. Remawati, "Penentuan Pengelompokan Penjualan Dengan Kombinasi K-Means Dan Hamming Distance," *Postulat : Jurnal Inovasi Pendidikan Matematika*, vol.

1, no. 2, p. 122, Dec. 2020, doi: 10.30587/postulat.v1i2.2090.

[9] S. I. Wahyudi and A. Wibowo, "Implementasi Metode K-Means Clustering Untuk Pengelompokan Data Stok Produk Toko Online Perdagangan Kaos," *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI) Jakarta-Indonesia*, 2022, [Online]. Available: https://senafti.budiluhur.ac.id/index.php

[10] N. S. N. R. Iwan Pii, "Penerapan Data Mining Pada Penjualan Produk Pakaian Dameyra Fashion Menggunakan Metode K-Means Clustering," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, pp. 423–430, Feb. 2023.

[11] B. I. A. B. Dila Aryani, "Implementasi Data Mining Pada Data Penjualan Pakaian Menggunakan Algoritma K-Means Dengan Optimize Parameter Grid," *Jurnal Mahasiswa Teknik Informatika*, no. Implementasi Data Mining pada Data Penjualan Pakaian Menggunakan Algoritma K-Means dengan Optimize Parameter Grid, pp. 1673–1680, Apr. 2024.

[12] C. Purnama, W. Witanti, and P. N. Sabrina, "Klasterisasi Penjualan Pakaian Untuk Meningkatkan Strategi Penjualan Barang Menggunakan K-Means," *Jounal of Information Technology*, pp. 35–38, Mar. 2022.

[13] E. Mayoana Fitri, R. Randy Suryono, and A. Wantoro, "Klasterisasi Data Penjualan Berdasarkan Wilayah Menggunakan Metode K-Means Pada Pt Xyz," *Jurnal Komputasi*, vol. 11, no. 2, p. 2023, 2023.

[14] N. Wahyunisari, R. Kurniawan, J. Perjuangan No, and B. Karyamulya Kec Kesambi Kota Cirebon, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Data Penjualan Pakaian (Studi Kasus : UMKM Lima Media Kuningan)," 2024. [Online]. Available: https://djpb.kemenkeu.go.id

[15] D. N. M. W. D. S. Indah Puspitorini, "Penentuan Jenis Benang Untuk Pembuatan Produk Pada Konveksi Menggunakan Metode Algoritma C4.5 Dengan Rapid Miner," *Jurnal Minfo Polgan*, vol. 12, no. 1, Mar. 2023, doi: 10.33395/jmp.v12i1.12555.

[16] M. Arif Saifudin, H. Endah Wahanani, and A. Junaidi, "Implementasi Algoritma Asosiasi Fp-Growth Dan Klasifikasi K-Means Terhadap Pola Pembelian Konsumen Di Marketplace Shopee," 2024.

[17] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-Means Clustering Dalam Penentuan Siswa Kelas Unggulan," *Jurnal TEKNO KOMPAK*, vol. 15, no. 2, pp. 25–36, 2022.