

DATA MINING UNTUK PENEMPATAN BARANG BERDASARKAN FREKUENSI PERMINTAAN DI PT LAUTAN LESTARI SHIPYARD

Nober Six Salvanius Mendrofa¹,
Sunarsan Sitohang²

¹Mahasiswa Program Studi Teknik Informatika, Universitas Putera Batam

²Dosen Program Studi Teknik Informatika, Universitas Putera Batam
email: pb180210121@upbatam.ac.id

ABSTRACT

Good arrangement of goods in the warehouse is one of the important things in supporting the activities of a company. At PT Lautan Lestari Shipyard, the placement of goods in the warehouse is still relatively irregular and looks messy, this is because the placement is not based on the frequency of the number of goods coming out. This study aims to apply data mining in the preparation of the layout of goods by utilizing the archive of goods expenditure data available at the company by using the algorithm method k-means clustering. From this research, 5 data clusters were found to be the most optimal with a Davies Bouldin Index (DBI) value of 0.288. This research is expected to help make it easier for companies to arrange the goods in the warehouse.

Keywords: Clustering; Data Mining; K-Means; Layout.

PENDAHULUAN

Tersedianya data yang melimpah, keperluan pada informasi ataupun pengetahuan untuk dijadikan sebagai bahan penyokong penentuan strategi dalam pemecahan permasalahan bisnis, serta sokongan infrastruktur di bidang teknologi data ialah awal mula dari munculnya data *mining*. Data *mining* dirancang sebagai pengganti pemecahan masalah yang nyata untuk para pimpinan dalam mengambil keputusan. Data *mining* bertujuan untuk menggali nilai tambah pada kumpulan informasi yang dapat berupa pengetahuan yang belum diketahui.

Menurut (Vulandari, 2017) *clustering* merupakan salah satu metode data *mining* atau sering disebut dengan

segmentation. Metode *clustering* dimanfaatkan untuk menentukan kelompok alami dari suatu kasus yang didasarkan pada sebuah kelompok atribut, kemudian mengelompokkan data yang memiliki persamaan atribut. *Clustering* juga merupakan metode yang *unsupervised*, karena seluruh atribut input diperlakukan sama.

Algoritma metode *clustering* yang masuk dalam kelompok *unsupervised learning* adalah algoritma *K-means*, algoritma ini dimanfaatkan untuk membagi data ke beberapa kelompok dengan sistem partisi. Algoritma ini menerima input berbentuk data tanpa label kelas. (Wanto et al., 2020)

Dalam melakukan operasionalnya PT Lautan Lestari Shipyard membutuhkan

peralatan maupun material penunjang dalam menggerakkan roda perusahaan. Penempatan barang yang baru masuk disusun di rak berdasarkan rak yang sudah kosong duluan, sistem penempatan seperti ini tentunya tidak berdasarkan dengan jumlah frekuensi pengambilan terbanyak dan terkesan tidak rapi, sehingga menyebabkan kesulitan dalam pengambilan barang ketika diminta oleh pekerja di lapangan.

Pengambilan alat dan bahan tersebut diatas akan dicatat dalam *form* dan direkap dengan menggunakan perangkat lunak *Microsoft Excel*. Jika data tersebut tidak dilakukan pengolahan yang tepat, data yang telah terkumpul akan sia-sia dan tidak bernilai guna. Arsip data yang telah terkumpul dapat memberikan pola pengambilan barang dan dari pola tersebut dapat diketahui barang dengan kegunaan yang banyak, sedang atau sedikit. Pola ini dapat berguna untuk mencari kaitan antara barang dan bisa digunakan untuk mengatur letak yang tepat dan penempatan barang yang harus berdekatan.

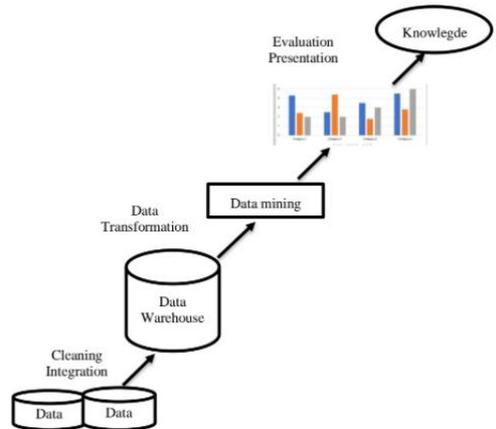
Hasil akhir penelitian ini adalah untuk menganalisis penempatan barang serta menerapkan metode *clustering* dalam penempatan barang berdasarkan frekuensi permintaan.

KAJIAN TEORI

2.1 *Knowledge Discovery in Database (KDD)*

Knowledge Discovery in Database (KDD) merupakan suatu proses untuk menemukan pengetahuan yang ada dalam database. Definisi lengkap dari KDD adalah keseluruhan proses memilah atau mengidentifikasi pola, pengetahuan dan informasi potensial dari kumpulan data besar. Output Pengetahuan dan informasi dari KDD

bersifat aktual, baru, mudah dipahami serta bernilai guna (Bulolo, 2020). Berikut ini adalah tahapan proses KDD.



Gambar 1. Tahapan Proses KDD

Sumber: (Bulolo, 2020)

2.2 *Data Mining*

Data mining dikenal sebagai *Knowledge discovery, Knowledge extraction, data/pattern analysis, information harvesting*, yang merupakan suatu proses menggali suatu hal penting yang dapat berupa informasi atau pola yang bernilai guna dari data yang besar. *Data mining* dimulai dari proses mengumpulkan data, mengekstraksi data, menganalisis data dan statistik data. (Arhami & Nasir, 2020).

Untuk mendapatkan skema ataupun informasi dari data, *data mining* memanfaatkan teknik atau metode tertentu. Dalam *data mining* mengekstrak dan mengidentifikasi suatu informasi yang dibutuhkan dari database yang besar diperlukan teknik seperti statistik, matematika, kecerdasan buatan, maupun *machine learning*. (Raja Gukguk & Sitohang, 2021)

2.3 Metode Data Mining

Pada tahun 1976, J.B. Mac Oueen adalah orang pertama yang memperkenalkan algoritma *K-Means*. Ketika memilah data berdasarkan ciri yang sama, algoritma inilah salah satu algoritma pengelompokan yang paling sering dipakai. Produk akhir yang dikenal sebagai *cluster* terbentuk sebagai hasil dari proses ini (Arhami & Nasir, 2020).

Clustering adalah pendekatan data *mining* yang berusaha untuk menemukan kelompok item dengan karakteristik serupa yang dapat diurutkan dari set lain dari objek, dengan objek ke kelompok yang sama yang lebih homogen daripada objek dalam kelompok yang terpisah (Jollyta et al., 2021).

Dengan metode *clustering*, dapat kita identifikasikan area yang padat, menemukan hubungan yang menarik antar atribut data dan menemukan pola distribusi secara keseluruhan. Teknik *clustering* digunakan untuk mengelompokkan data dengan kriteria serupa ke wilayah atau grup yang sama, sambil memisahkan data dengan kriteria yang berbeda. Data akan dikelompokkan dalam satu kelompok atau *cluster* mempunyai tingkat kemiripan yang maksimum dan setiap kelompok mempunyai tingkat kemiripan yang minimum (Ginantra et al., 2021)

Berikut merupakan beberapa langkah dalam pengelompokkan data (Witanto et al., 2019)

1. Tentukan jumlah/banyaknya kluster yang hendak dibentuk.
2. Inisialisasi awal dan pusat kluster dilakukan dengan cara random
3. Setiap titik data diposisikan di tengah *cluster* terdekat. Pada titik ini, metode *Euclidean* digunakan untuk menghitung jarak dengan menentukan kesamaan atau perbedaan data. Berikut rumusnya:

$$d(x, y) = \sqrt{\sum_i^n (x_i - y_i)^2}$$

4. Rata-rata objek dalam *cluster* digunakan untuk menghitung pusat *cluster*. Selain itu, median juga dapat digunakan untuk membuat perhitungan.
5. Untuk menyelesaikan proses pengklasteran, harus menghitung ulang jarak antara setiap objek dan pusat *cluster* sampai *cluster* tetap tidak berubah.

2.4 Algoritma Data Mining

Algoritma *K-Means* merupakan salah satu metode dalam data *mining* untuk mengklusterkan (*Clustering*) data kedalam beberapa kelompok berdasarkan jarak, kriteria, kondisi atau karakteristik. Data dalam satu kelompok harus memiliki jarak terpendek, kriteria, kondisi atau karakteristik yang sama atau mirip antara satu dengan lainnya. Algoritma *K-Means* dapat mengelompokkan objek yang memiliki kemiripan (Lestari, 2019).

2.5 Davies Bouldin Index (DBI)

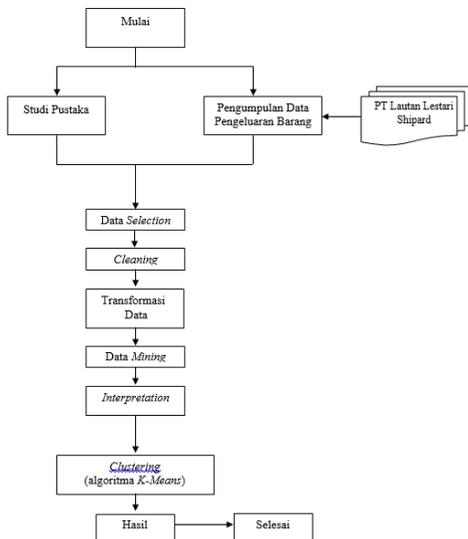
Pada tahun 1979 David L. Davies dan Donald W. Bouldin mengenalkan *Davies Bouldin Index* (DBI), sebuah teknik untuk mengevaluasi efisiensi pengelompokan dan mendefinisikan kohesi sebagai ukuran kedekatan data dengan titik pusat *cluster*. Ada skema evaluasi berdasarkan jumlah dan kedekatan data hasil *clustering* saat menggunakan *Davies Bouldin Index* untuk *clustering* internal (Muningsih et al., 2021: 97)

Davies Bouldin Index menghitung jarak (*dissimilarity*) antar *cluster*. Algoritma pengelompokan dengan DBI rendah menghasilkan kluster dengan jarak intra-kluster rendah (kesamaan intra-kluster tinggi) dan jarak antar-kluster yang tinggi (kesamaan antar-kluster

rendah). Semakin rendah nilai *Davies Bouldin Index* suatu algoritma *machine learning* yang digunakan untuk klusterisasi, maka semakin tinggi kinerjanya. (Santoso & Azis, 2020: 53). Tujuan evaluasi *cluster* ini untuk menentukan seberapa baik hasil pengelompokan cluster data. DBI adalah metode yang dapat dimanfaatkan untuk mengevaluasi hasil pengelompokan dan menentukan jumlah *cluster* yang optimal untuk digunakan.

METODE PENELITIAN

Hasil akhir penelitian dapat diperoleh dengan mengikuti langkah-langkah dari awal proses penelitian, sehingga dapat ditarik kesimpulan dari keseluruhan proses penelitian, seperti pada Gambar 2.



Gambar 2. Tahapan Penelitian (Sumber: Data Penelitian, 2022)

1. *Selection*, pemilihan data diperlukan untuk memilah data, karena data yang ada tidak semuanya dapat dipergunakan.

2. *Pre-processing/cleaning*, menghapus dan mengoreksi redundansi, inkonsistensi, dan kesalahan dari data yang dipilih dilakukan pada tahap ini.
3. *Transformation*, tahap ini adalah transformasi dataset sehingga dapat digunakan dalam proses penambangan data. Tahap ini termasuk proses kreatif yang sangat bergantung pada pola pencarian *database* atau jenis informasi yang akan dicari.
4. *Data mining*, tahapan ini merupakan tahapan utama dari proses KDD, dimana data yang terpilih akan dicari informasi yang bernilai dengan menggunakan metode dan teknik data *mining*. Dalam data *mining*, ada berbagai pendekatan, strategi, dan algoritma untuk dipilih, tergantung pada tujuan atau informasi yang ingin dicari.
5. *Interpretation/Evaluation*, data *mining* menghasilkan pengetahuan dan informasi yang harus disajikan secara mudah dimengerti oleh mereka yang membutuhkannya.

HASIL DAN PEMBAHASAN

4.1 Analisis Data

Pada penelitian ini tahapan pertama adalah pengumpulan data, data yang digunakan adalah data transaksi pengambilan barang di lokasi penelitian. Data tersebut selanjutnya akan diproses data yang relevan maupun data yang inkonsisten dari *database*, karena tidak semua data yang ada dibutuhkan dalam proses data *mining*. Setelah data siap diseleksi selanjutnya adalah pembersihan data/*cleaning* data, yaitu proses menjadikan data agar berkualitas dengan membuang data yang inkonsisten, memeriksa duplikasi data dan

memperbaiki kesalahan seperti kesalahan penulisan.

1. Penentuan jumlah cluster

Berdasarkan hasil *Davies Bouldin Index* (DBI) jumlah kluster optimal yang dipakai pada penelitian ini berjumlah 5 kluster, langkah pertama adalah penentuan titik pusat cluster secara acak seperti ditabel berikut:

Tabel 1. Titik pusat cluster

Cluster	Nama Item	Tahun 2020	Tahun 2021
C0	Supra Clamp	6173	5303
C1	Plastic	9371	22288
C2	Disc	11912	14177
C3	Spring & Flat Washer	10225	11098
C4	Kaca & Kapur	6200	5374

(Sumber: Data Penelitian, 2022)

2. Menghitung jarak ke pusat kluster

Langkah kedua setelah diperoleh titik pusat kluster adalah melakukan perhitungan jarak *Euclidean*, jarak terkecil akan dipisahkan kelompoknya. Selanjutnya akan diperoleh titik *centroid* baru untuk dijadikan sebagai acuan perhitungan berikutnya hingga nilai antara *centroid* sebelum dan sesudah bernilai sama. Berikut perhitungan jarak *Euclidean* pada iterasi 1.

$$d_{1.0} = \sqrt{(532 - 6.173)^2 + (608 - 5.303)^2} = 7.339,20$$

$$d_{1.1} = \sqrt{(532 - 9.371)^2 + (608 - 22.288)^2} = 23.412,61$$

$$d_{1.2} = \sqrt{(532 - 11.912)^2 + (608 - 14.177)^2}$$

$$= 17.709,38$$

$$d_{1.3} = \sqrt{(532 - 10.225)^2 + (608 - 11.098)^2} = 14.282,66$$

$$d_{1.4} = \sqrt{(532 - 6.200)^2 + (608 - 5.374)^2} = 7.405,47$$

$$d_{2.0} = \sqrt{(6 - 6.173)^2 + (28 - 5.303)^2} = 8.115,26$$

$$d_{2.1} = \sqrt{(6 - 9.371)^2 + (28 - 22.288)^2} = 24.149,76$$

$$d_{2.2} = \sqrt{(6 - 11.912)^2 + (28 - 14.177)^2} = 18.491,81$$

$$d_{2.3} = \sqrt{(6 - 10.225)^2 + (28 - 11.098)^2} = 15.065,62$$

$$d_{2.4} = \sqrt{(6 - 6.200)^2 + (28 - 5.374)^2} = 8.182,01$$

dan seterusnya sampai data ke-135. Hasil perhitungan tersebut digunakan untuk menentukan titik *centroid* baru.

3. Menentukan titik *centroid* baru

Setelah mendapatkan anggota setiap kluster selanjutnya pusat *cluster* baru dihitung berdasarkan data anggota tiap-tiap kluster sesuai dengan rumus pusat anggota kluster. Sehingga didapatkan hasil seperti pada Tabel 2:

Tabel 2. Pusat kluster baru

Cluster	Tahun 2020	Tahun 2021
C0	456,90	493,41
C1	9.371,00	22.288,00
C2	17.158,25	17.030,75
C3	10.225,00	11.098,00

C4	4.479,50	6.388,25
----	----------	----------

(Sumber: Data Penelitian, 2022)

Untuk iterasi ke-2, ulangi langkah ke-2 (kedua) seperti iterasi ke-1 yaitu menghitung masing-masing jarak ke pusat kluster dengan menggunakan pusat kluster baru.

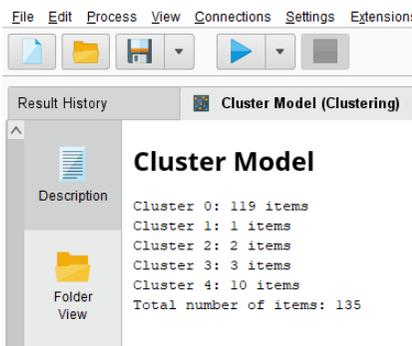
Pada pengolahan data ini peneliti melakukan perhitungan jarak *Euclidean*

Tabel 3. Jumlah item per cluster

Cluster	Jumlah Item	Keterangan
Cluster 0	119 item	sangat jarang digunakan
Cluster 1	1 item	jarang digunakan
Cluster 2	2 item	penggunaan sedang
Cluster 3	3 item	sering digunakan
Cluster 4	10 item	sangat sering digunakan

(Sumber: Data Penelitian, 2022)

Hasil pengujian dengan menggunakan *software RapidMiner* sejalan dengan perhitungan manual yang menggunakan rumus *Euclidean Distance*. Hasil pengujian dengan *software RapidMiner* dapat dilihat di Gambar 3 berikut:



Gambar 3. Cluster Model
(Sumber: Data Penelitian, 2022)

Untuk mengetahui jumlah *cluster* yang optimal dalam pengolahan data dengan menggunakan metode *k-means* perlu

sampai pada iterasi ke-3, karena iterasi ke-2 dan iterasi ke-3 memiliki anggota yang sama maka tidak perlu dilakukan iterasi/perulangan lagi. Hasil *clustering* telah mencapai stabil dan *konvergen*. Berikut hasil akhir perhitungan manual *Euclidean Distance* pada iterasi ke-3.

dilakukan evaluasi *clustering*, caranya adalah mencari nilai DBI terkecil untuk mengetahui optimasi jumlah klusternya. Dibawah ini adalah langkah untuk mencari nilai DBI dengan menggunakan 5 cluster data.

- a. *Sum of square within cluster (SSW)*
Semakin kecil nilai SSW maka hasil *clustering* semakin baik, perhitungan SSW ini berdasarkan nilai *centroid* yang didapatkan dari iterasi ke-3.

Tabel 4. Pusat *centroid* iterasi ke-3

Cluster	Tahun 2020	Tahun 2021
C0	263,59	290,80
C1	9.371,00	22.288,00
C2	21.909,50	19.705,50
C3	11.679,67	13.270,00
C4	4.366,40	5.181,40

(Sumber: Data Penelitian, 2022)

Sebelum mencari nilai SSW, langkah pertama adalah mencari jarak data ke *cluster* seperti dibawah ini:

$$d(x_1, c_1) = \sqrt{(532 - 263,58)^2 + (608 - 290,80)^2} = 415,12$$

$$d(x_2, c_1) = \sqrt{(6 - 263,58)^2 + (28 - 290,80)^2} = 367,99$$

$$d(x_3, c_1) = \sqrt{(347 - 263,58)^2 + (444 - 290,80)^2} = 174,44$$

$$d(x_4, c_1) = \sqrt{(25 - 263,58)^2 + (58 - 290,80)^2} = 333,35$$

$$d(x_5, c_1) = \sqrt{(1 - 263,58)^2 + (52 - 290,80)^2} = 354,93$$

dan selanjutnya hingga data ke-135. Setelah didapatkan hasil perhitungan jarak data ke *cluster* diatas, langkah selanjutnya mencari nilai rata-rata jarak data ke *cluster* berdasarkan *cluster* data itu sendiri. Hasil perhitungan SSW dapat dilihat di Tabel 5 berikut:

Tabel 5. Hasil Perhitungan SSW

SSW cluster 0	443,11
SSW cluster 1	0

SSW cluster 2	1.643,23
SSW cluster 3	1.769,83
SSW cluster 4	1.645,17

(Sumber: Data Penelitian, 2022)

b. *Sum of Square Between (SSB)*

$$SSB_{0,0} = \sqrt{(263,59 - 263,59)^2 + (290,8 - 290,8)^2} = 0$$

$$SSB_{0,1} = \sqrt{(263,59 - 9.371)^2 + (290,8 - 22.288)^2} = 23.808,02$$

$$SSB_{0,2} = \sqrt{(263,59 - 21.909,5)^2 + (290,8 - 19.705,5)^2} = 29.077,07$$

$$SSB_{0,3} = \sqrt{(263,59 - 11.679,67)^2 + (290,8 - 13.270)^2} = 17.285,44$$

$$SSB_{0,4} = \sqrt{(263,59 - 4.336,4)^2 + (290,8 - 5.181,4)^2} = 6.383,65$$

Dari perhitungan diatas didapatkan matriks seperti pada Tabel 6 berikut:

Tabel 6. Hasil Perhitungan SSB

SSB	0	1	2	3	4
0	0	23.808,02	29.077,07	17.285,44	6.383,65
1	23.808,02	0	12.801,69	9.308,83	17.823,63
2	29.077,07	12.801,69	0	12.085,74	22.775,20
3	17.285,44	9.308,83	12.085,74	0	10.904,55
4	6.383,65	17.823,63	22.775,20	10.904,55	0

(Sumber: Data Penelitian, 2022)

c. *Ratio*

$$R_{0,1} = \frac{443.11 + 0}{23.808,02} = 0,02$$

$$R_{0,2} = \frac{443.11 + 1.643,23}{29.077,07} = 0,07$$

$$R_{0,3} = \frac{443.11 + 1.769,83}{17.285,44} = 0,13$$

$$R_{0,4} = \frac{443.11 + 1.645,17}{17.285,44} = 0,33$$

$$R_{1,2} = \frac{6.383,65}{0 + 1.643,23} = 0,13$$

$$R_{1,2} = \frac{0 + 1.769,83}{12.801,69} = 0,13$$

$$R_{1,3} = \frac{0 + 1.769,83}{9.308,83} = 0,19$$

$$R_{1,4} = \frac{0 + 1.645,17}{17.823,63} = 0,09$$

$$R_{2,3} = \frac{1.643,23 + 1.769,83}{12.085,74} = 0,28$$

$$R_{2,4} = \frac{1.643,23 + 1.645,17}{22.775,20} = 0,14$$

$$R_{3,4} = \frac{1.769,83 + 1.645,17}{10.904,55} = 0,31$$

Berikut matriks perhitungan ratio:

Tabel 7. Hasil Perhitungan *Ratio*

R	0	1	2	3	4	R Max
0	0	0,02	0,07	0,13	0,33	0,33
1	0,02	0	0,13	0,19	0,09	0,19
2	0,07	0,13	0	0,28	0,14	0,28
3	0,13	0,19	0,28	0	0,31	0,31
4	0,33	0,09	0,14	0,31	0	0,33

(Sumber: Data Penelitian, 2022)

d. Perhitungan DBI

$$DBI = \frac{1}{5}(0,33 + 0,19 + 0,28 + 0,31 + 0,33) = 0,288$$

Semakin kecil nilai DBI yang didapatkan (non-negatif ≥ 0) maka semakin baik kluster yang diperoleh dari pengelompokan *k-means* yang digunakan. Setelah mendapatkan nilai DBI untuk 5 kluster, selanjutnya membandingkan nilai DBI terkecil dengan jumlah kluster lainnya. Dengan memanfaatkan *software RapidMiner*, peneliti melakukan pengujian jumlah cluster 2-10 seperti pada Tabel 8 berikut:

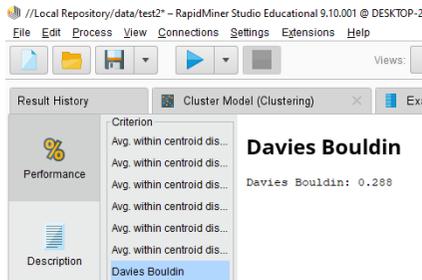
5	0.288
6	0.416
7	0.558
8	0.513
9	0.445
10	0.377

(Sumber: Data Penelitian, 2022)

Dari hasil pengujian, diketahui jumlah *cluster* yang paling optimal adalah 5 *cluster* dengan DBI 0,288. Sehingga pengelompokan yang dilakukan dengan metode *k-means* diatas, pengelompokan 5 *cluster* adalah yang *cluster* yang paling optimal.

Tabel 8. Hasil Pengujian DBI

Cluster	Nilai DBI
2	0.343
3	0.471
4	0.426



Gambar 4. Davies Bouldin 5 cluster
(Sumber: Data Penelitian, 2022)

SIMPULAN

Dari hasil pengujian yang telah dilakukan, maka diperoleh beberapa kesimpulan:

1. Dengan menggunakan algoritma *k-means clustering* didapatkan kluster paling optimal yaitu 5 cluster data dengan nilai Davies Bouldin Index (DBI) 0.288.
2. Algoritma *k-means clustering* dapat diterapkan di PT Lautan Lestari Shipyard, dengan mengolah data transaksi pengambilan barang seperti nama item dan *qty* barang. Dari item yang paling banyak keluar inilah selanjutnya diatur penempatannya, sehingga memudahkan *storemen* untuk mendapatkan barang tersebut.
3. Barang yang sangat sering digunakan dapat dilihat di cluster 4 dengan jumlah 10 item barang yaitu bulu roll, cable JZ, filter, isolasi, kaca dan kapur, lampu, paku dan buat, *racor*, *shackle* dan *supra clamp*. Item yang masuk dalam cluster 4 ini direkomendasikan untuk ditempatkan di rak terdekat, sehingga memudahkan dalam proses pengambilan barang.

DAFTAR PUSTAKA

Arhami, M., & Nasir, M. (2020). Data Mining Algoritma dan Implementasi.

DATA MINING Algoritma Dan Implementasi, 2.

Buulolo, E. (2020). *Data Mining Untuk Perguruan Tinggi*. Deepublish.

Ginantra, N. L. W. S. R., Arifah, F. N., Wijaya, A. H., Septarini, R. S., Ahmad, N., Ardiana, D. P. Y., Effendy, F., Iskandar, A., Hazriani, H., & Sari, I. Y. (2021). *Data Mining dan Penerapan Algoritma*. Yayasan Kita Menulis.
<https://books.google.co.id/books?id=v0gtEAAAQBAJ>

Jollyta, D., Siddik, M., Mawengkang, H., & Efendi, S. (2021). *Teknik Evaluasi Cluster Solusi Menggunakan Python Dan Rapidminer*. Deepublish.
<https://books.google.co.id/books?id=3rcgEAAAQBAJ>

Lestari, D. R. (2019). *Implementasi Data Mining Untuk Pengaturan Layout Swalayan Delimas Lestari Kencana Dengan Menggunakan Dengan Menggunakan Algoritma K-Means Clustering*. 6(2), 112–117.

Muningsih, E., Maryani, I., & Handayani, V. R. (2021). Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa. *Jurnal Sains Dan Manajemen*, 9(1), 95–100.
<https://ejournal.bsi.ac.id/ejurnal/index.php/evolusi/article/view/10428/4839>

Raja Gukguk, A. W. S., & Sitohang, S. (2021). *Jurnal Comasie PENERAPAN DATA MINING DALAM PEMILIHAN LAPTOP*. 5, 63–70.

Santoso, B., & Azis, A. I. S. (2020). *Machine Learning & Reasoning Fuzzy Logic Algoritma, Manual, Matlab, & Rapid Miner*. Deepublish.
<https://books.google.co.id/books?id>

- =4j_YDwAAQBAJ
- Vulandari, R. T. (2017). *Data Mining: Teori dan Aplikasi Rapidminer*.
- Wanto, A., Siregar, M. N. H., Windarto, A. P., Hartama, D., Ginantra, N. L. W. S. R., Napitupulu, D., Negara, E. S., Lubis, M. R., Dewi, S. V., & Prianto, C. (2020). *Data Mining: Algoritma dan Implementasi*. Yayasan Kita Menulis.
- Witanto, S., Ratnawati, D. E., & Anam, S. (2019). Pengelompokan Fungsi Aktif Senyawa Data SMILES (Simplified Molecular Input Line Entry System) Menggunakan Metode K-Means Dengan Inisialisasi Pusat Klaster Menggunakan Metode Heuristic O (N LogN). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(1), 702–707.

	<p>Biodata</p> <p>Penulis pertama, Nober Six Salvanius Mendrofa, merupakan mahasiswa Prodi Teknik Informatika Universitas Putera Batam.</p>
	<p>Biodata</p> <p>Penulis kedua, Sunarsan Sitohang, S.Kom., M.TI., merupakan Dosen Prodi Teknik Informatika Universitas Putera Batam.</p>