

Segmentasi Supplier Berdasarkan Data Transaksi Menggunakan K-Means dengan Pendekatan QRFM Termodifikasi dan Validasi Multi-Platform

Nasywa Dhiya Luthfiyyah¹, Baenil Huda², Fitria Nurapriani³, Elfina Novalia⁴

Sistem Informasi, Fakultas Ilmu Komputer, Universitas Buana Perjuangan, Karawang

si22.nasywaluthfiyyah@mhs.ubpkarawang.ac.id¹, baenil88@ubpkarawang.ac.id²,

fitria.apriani@ubpkarawang.ac.id³, elfinanovalia@ubpkarawang.ac.id⁴,

Abstrak: Penelitian ini bertujuan untuk melakukan segmentasi supplier berdasarkan data transaksi menggunakan pendekatan data mining dengan algoritma K-Means. Permasalahan yang diangkat adalah belum adanya pengelompokan supplier yang dapat mendukung pengambilan keputusan strategis dalam pengelolaan supplier. Metode yang digunakan adalah K-Means clustering dengan pendekatan QRFM (*Quantity, Recency, Frequency, Monetary*) sebagai pengembangan dari model RFM. Dataset yang digunakan berupa data estimasi invoice periode Januari 2024 hingga Desember 2025 yang telah melalui proses transformasi dan normalisasi menggunakan metode Min-Max. Evaluasi model dilakukan menggunakan *Davies-Bouldin Index* (DBI) dan divalidasi menggunakan dua platform, yaitu RapidMiner dan Google Colab. Hasil penelitian menunjukkan bahwa model RFM menghasilkan nilai DBI terbaik sebesar 0,357, sedangkan QRFM sebesar 0,377. Sementara itu, hasil validasi pada Google Colab menunjukkan nilai DBI sebesar 0,642 dan Silhouette Score sebesar 0,580. Meskipun nilai DBI pada RapidMiner lebih optimal, kedua platform menghasilkan pola kluster yang konsisten. Hasil clustering menghasilkan tiga kelompok supplier, yaitu supplier prioritas, supplier menengah, dan supplier berisiko. Penelitian ini menunjukkan bahwa pendekatan QRFM mampu memberikan insight yang lebih komprehensif dibandingkan dengan RFM dalam analisis segmentasi supplier.

Kata Kunci: K-Means, QRFM, Clustering, Supplier, DBI.

Abstract: This study aims to segment suppliers based on transaction data using a data mining approach with the K-Means algorithm. The problem addressed is the lack of supplier clustering that can support strategic decision-making in supplier management. The method used is K-Means clustering with the QRFM (*Quantity, Recency, Frequency, Monetary*) approach as an extension of the RFM model. The dataset consists of estimated invoice data from January 2024 to December 2025, which has undergone transformation and normalization using the Min-Max method. Model evaluation was performed using the *Davies-Bouldin Index* (DBI) and validated on two platforms: RapidMiner and Google Colab. The results show that the RFM model yields the best DBI value of 0.357, while the QRFM model yields 0.377. Meanwhile, validation results on Google Colab show a DBI value of 0.642 and a Silhouette Score of 0.580. Although the DBI value on RapidMiner is more optimal, both platforms produce consistent clustering patterns. The clustering results yielded three supplier groups: priority suppliers, medium-risk suppliers, and high-risk suppliers. This study demonstrates that the QRFM approach provides more comprehensive insights compared to RFM in supplier segmentation analysis.

Keywords: K-Means, QRFM, Clustering, Supplier, DBI.

1. Pendahuluan

Dalam era digital saat ini, perkembangan teknologi informasi telah mendorong organisasi untuk memanfaatkan data sebagai salah satu aset utama dalam mendukung proses pengambilan keputusan. Data transaksi yang tersimpan secara historis, khususnya dalam bentuk invoice, menyimpan informasi penting terkait aktivitas operasional yang dapat dianalisis lebih lanjut untuk memperoleh insight yang bernilai. Namun, dalam praktiknya, pemanfaatan data tersebut sering kali masih terbatas pada kebutuhan administratif dan pelaporan, sehingga potensi informasi yang terkandung di dalamnya belum dimanfaatkan secara optimal (Ariati et al., 2024; Panjaitan et al., 2025).

Algoritma K-Means merupakan salah satu metode clustering yang paling banyak digunakan karena memiliki keunggulan dalam kesederhanaan, efisiensi, serta kemampuannya dalam mengolah data dalam jumlah besar. Algoritma ini bekerja dengan membagi data ke dalam sejumlah cluster berdasarkan kedekatan jarak antar data terhadap pusat cluster atau centroid. Meskipun demikian, kualitas hasil clustering sangat dipengaruhi oleh pemilihan atribut yang digunakan sebagai dasar pengelompokan (Kristanti et al., 2024; Panjaitan et al., 2025).

Dalam penelitian ini, digunakan pendekatan QRFM, yaitu *Quantity, Recency, Frequency, dan Monetary*, sebagai pengembangan dari model RFM yang umum digunakan dalam analisis perilaku pelanggan. Penambahan atribut Quantity bertujuan memberikan dimensi tambahan dalam mengukur intensitas transaksi, sehingga diharapkan dapat menghasilkan segmentasi yang lebih representatif dibandingkan dengan model RFM konvensional. Dengan menggunakan pendekatan QRFM, setiap supplier dapat dianalisis berdasarkan jumlah transaksi, waktu terakhir transaksi, frekuensi transaksi, serta nilai transaksi yang dihasilkan (Fadhillah et al., 2025; Maresti et al., 2025).

Selain pemilihan atribut, tahap preprocessing data seperti normalisasi juga memiliki peran penting dalam meningkatkan kualitas clustering. Normalisasi dilakukan untuk menyamakan skala antar atribut sehingga tidak terjadi dominasi nilai tertentu dalam proses perhitungan jarak. Tanpa normalisasi, atribut dengan skala besar dapat mempengaruhi hasil clustering secara signifikan dan menghasilkan cluster yang kurang optimal. Untuk mengevaluasi kualitas hasil clustering, digunakan metode Davies-Bouldin Index (DBI) yang merupakan salah satu metrik evaluasi internal dalam clustering. DBI mengukur tingkat kesamaan antar cluster dan tingkat pemisahan antar cluster, di mana nilai DBI yang lebih kecil menunjukkan kualitas clustering yang lebih baik. Oleh karena itu, DBI digunakan sebagai acuan dalam menentukan hasil clustering yang optimal dalam penelitian ini (Hananto et al., 2021; Triantara et al., 2025).

Selain itu, penelitian ini juga melakukan validasi hasil clustering menggunakan dua platform yang berbeda, yaitu RapidMiner dan Google Colab. Penggunaan dua platform ini bertujuan untuk memastikan konsistensi hasil clustering serta membandingkan performa masing-masing platform dalam mengimplementasikan algoritma K-Means. Perbedaan implementasi, parameter, maupun proses pengolahan data pada masing-masing platform dapat mempengaruhi hasil akhir clustering, sehingga diperlukan analisis perbandingan untuk memperoleh hasil yang lebih akurat dan dapat dipertanggungjawabkan.

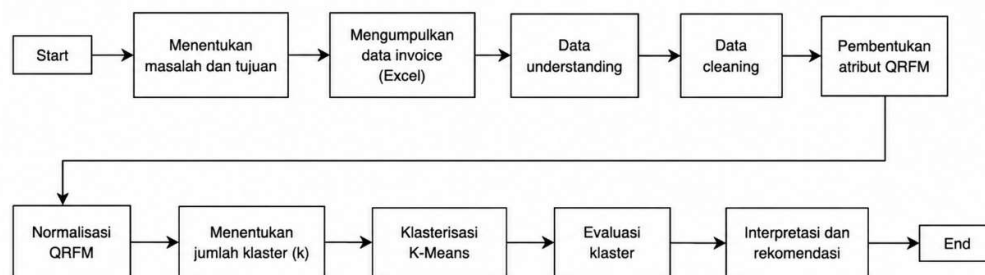
2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan teknik data mining berbasis unsupervised learning untuk melakukan segmentasi supplier

berdasarkan pola transaksi invoice. Pendekatan ini dipilih karena mampu mengidentifikasi pola tersembunyi dalam data tanpa memerlukan variabel target. Metode yang digunakan adalah algoritma K-Means Clustering, yaitu metode pengelompokan non-hierarkis Yang mengelompokkan data berdasarkan tingkat kemiripan karakteristik antar objek menggunakan ukuran jarak tertentu (Baenil Huda & Bayu Priyatna, 2020; Permana et al., 2023).

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari sistem pengadaan perusahaan dalam bentuk file Microsoft Excel dengan periode pengamatan Januari 2024 hingga Desember 2025. Dataset awal terdiri atas 3272 transaksi yang melibatkan 173 supplier. Setiap transaksi kemudian diagregasi menjadi data per supplier sehingga diperoleh 173 data yang digunakan pada tahap analisis. Sebelum proses clustering dilakukan, data terlebih dahulu diperiksa untuk mendeteksi *missing values* dan duplikasi. Data yang tidak lengkap pada atribut utama dihapus atau disesuaikan, sedangkan data duplikat dihilangkan agar tidak menimbulkan bias dalam hasil pengelompokan. Adapun kriteria inklusi dalam penelitian ini adalah data invoice yang memiliki atribut identitas supplier, tanggal transaksi, jumlah barang, dan nilai transaksi secara lengkap, sedangkan data di luar rentang waktu penelitian atau tidak memenuhi kelengkapan atribut dikeluarkan dari analisis (Permana et al., 2023).

Tahapan penelitian disusun secara konsisten mengikuti alur data selection, data understanding, data preprocessing, transformation, data mining, dan evaluation. Data selection dilakukan untuk memilih data yang sesuai dengan tujuan penelitian berdasarkan periode pengamatan dan kelengkapan atribut. Berikutnya, data understanding digunakan untuk memahami struktur, sebaran, dan karakteristik awal dataset sebelum dilakukan pengolahan lebih lanjut. Pada tahap data preprocessing, data diperiksa untuk menangani missing values, duplikasi, dan ketidakkonsistenan data. Setelah itu, tahap transformation dilakukan dengan membentuk data agregat berbasis supplier menggunakan pendekatan QRFM. Data yang telah dibentuk kemudian dianalisis pada tahap data mining menggunakan algoritma K-Means, dan hasil clustering dievaluasi menggunakan Davies-Bouldin Index serta Silhouette Score. Alur tahapan penelitian tersebut disajikan pada Gambar 1.



Gambar 1. Flowchart Alur Penelitian (Permana et al., 2023)

Penelitian dimulai dari proses pengumpulan informasi data hingga interpretasi hasil clustering. Setiap tahapan saling terhubung secara sistematis untuk memastikan kualitas analisis yang dihasilkan. Lalu dilakukan proses pemilihan data (*data selection*) yang relevan dengan tujuan penelitian, kemudian dilanjutkan dengan data understanding untuk memahami struktur, distribusi, serta karakteristik data. Pada tahap ini dilakukan eksplorasi awal terhadap data untuk mengidentifikasi pola transaksi, frekuensi invoice, serta potensi permasalahan seperti data tidak konsisten atau nilai ekstrem.

Tahap berikutnya adalah transformasi data menggunakan pendekatan QRFM (*Quantity, Recency, Frequency, Monetary*). Pada tahap ini, data transaksi diubah

menjadi data agregat berbasis supplier untuk menggambarkan perilaku transaksi. Variabel Quantity dihitung sebagai rata-rata nilai transaksi yang diperoleh dari perbandingan antara Monetary dan Frequency, Recency menunjukkan selisih waktu dari transaksi terakhir, Frequency merupakan jumlah transaksi, dan Monetary adalah total nilai transaksi. Definisi masing-masing variabel disajikan pada Tabel 1. (Asana et al., 2025; Hartati, 2020),

Tabel 1. Definisi QRFM

Atribut	Deskripsi
Quantity	rata-rata transaksi, dihitung dari perbandingan <i>Monetary</i> dan <i>Frequency</i> ($Q = M/F$)
Recency	waktu sejak transaksi terakhir
Frequency	jumlah transaksi per-supplier
Monetary	total nilai transaksi

Untuk mengatasi perbedaan skala antar variabel, dilakukan proses normalisasi data menggunakan metode Min-Max Scaling sehingga seluruh atribut berada dalam rentang yang sama.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Proses ini bertujuan untuk menghindari dominasi variabel tertentu dalam perhitungan jarak pada algoritma K-Means.

3. Hasil dan Pembahasan

3.1. Data Penelitian

Data yang digunakan dalam penelitian ini adalah data transaksi invoice dari pemasok yang diperoleh dari sistem pengadaan perusahaan selama periode Januari 2024 hingga Desember 2025. Dataset awal masih berupa data mentah yang mencakup beberapa atribut utama seperti identitas pemasok, tanggal transaksi, jumlah barang, dan nilai transaksi. Setiap baris data menggambarkan satu transaksi invoice yang kemudian diolah menjadi data agregat berdasarkan pemasok dengan menggunakan pendekatan QRFM. Proses agregasi ini bertujuan untuk membentuk karakteristik perilaku transaksi masing-masing pemasok sehingga dapat dimanfaatkan dalam proses pengelompokan (*clustering*).

Row No.	SUPPLIER ↑	Q	R	F	M
1	SUPPLIER PGA 001	56525399.875	58	24	1356609597
2	SUPPLIER PGA 002	17300120.135	60	74	1280208890
3	SUPPLIER PGA 003	8585363.111	138	9	77268268
4	SUPPLIER PGA 004	2902514.118	62	51	148028220
5	SUPPLIER PGA 005	20135400	83	11	221489400
6	SUPPLIER PGA 006	456662.500	149	8	3653300
7	SUPPLIER PGA 007	1145150	77	3	3435450
8	SUPPLIER PGA 008	55024331.643	61	28	1540681286
9	SUPPLIER PGA 009	1238645.370	61	46	56977687
10	SUPPLIER PGA 011	17250000	245	2	34500000

Gambar 2. Data Sebelum Normalisasi

Setelah dilakukan normalisasi menggunakan metode Min-Max, seluruh atribut berada dalam rentang 0 hingga 1 sehingga setiap variabel memiliki kontribusi yang seimbang dalam proses clustering.

Row...	SUPPLIER ↑	Q	R	F	M
1	SUPPLIER PGA 001	0.286	0.061	0.164	0.834
2	SUPPLIER PGA 002	0.086	0.064	0.521	0.787
3	SUPPLIER PGA 003	0.042	0.165	0.057	0.046
4	SUPPLIER PGA 004	0.013	0.066	0.357	0.090
5	SUPPLIER PGA 005	0.101	0.093	0.071	0.135
6	SUPPLIER PGA 006	0.001	0.179	0.050	0.001
7	SUPPLIER PGA 007	0.004	0.086	0.014	0.001
8	SUPPLIER PGA 008	0.278	0.065	0.193	0.947
9	SUPPLIER PGA 009	0.005	0.065	0.321	0.034
10	SUPPLIER PGA 011	0.086	0.304	0.007	0.020

Gambar 3. Data Setelah Normalisasi

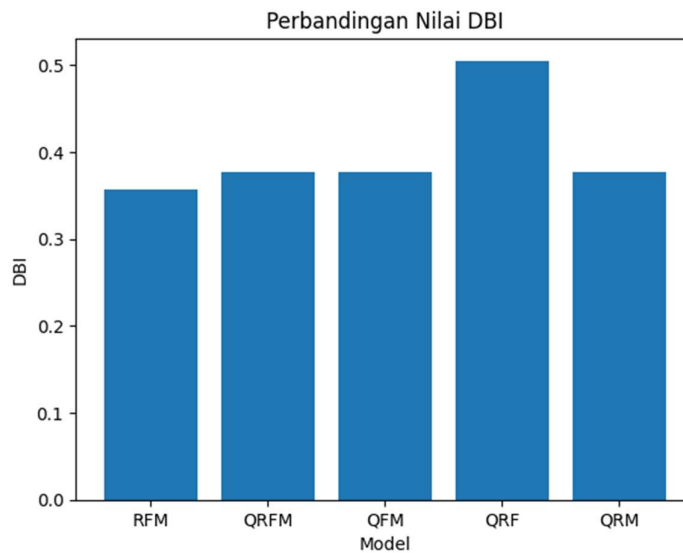
3.2. Evaluasi Model

Evaluasi dilakukan menggunakan *Davies-Bouldin Index* (DBI) untuk menentukan kualitas cluster yang dihasilkan.

Tabel 2. Perbandingan Nilai DBI

Model	DBI
RFM	0,357
QRFM	0,377
QFM	0,377
QRF	0,505
QRM	0,377

Hasil menunjukkan bahwa model RFM memiliki nilai DBI paling kecil, sehingga menghasilkan cluster terbaik secara matematis.

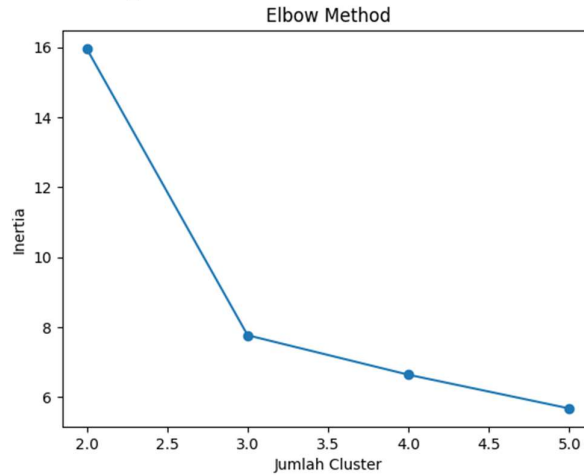


Gambar 4. Perbandingan Nilai DBI

Model RFM dan pengembangannya mampu memberikan insight terhadap perilaku transaksi (Carudin, 2021; Syahra et al., 2025).

3.3. Penentuan Jumlah Cluster

Penentuan jumlah cluster dilakukan menggunakan metode Elbow untuk menemukan jumlah cluster optimal.



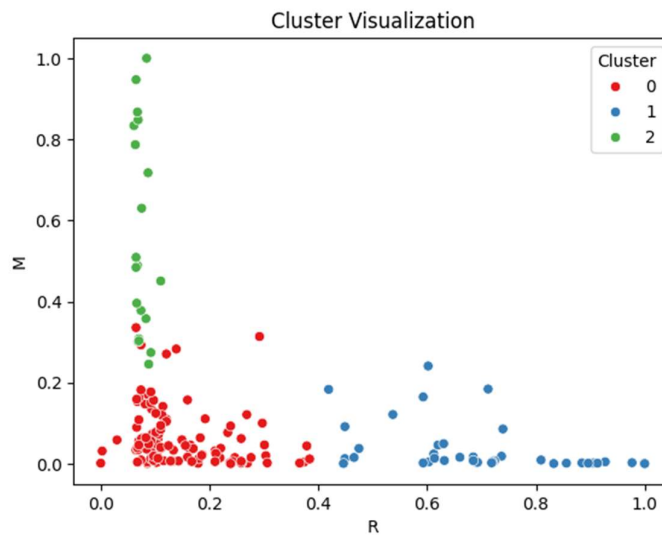
Gambar 5. Grafik Elbow Cluster

3.4. Hasil Clustering

Berdasarkan hasil clustering menggunakan K-Means, diperoleh tiga cluster utama (Maresti et al., 2025). Hasil menunjukkan terbentuknya tiga cluster.

Tabel 3. Distribusi Cluster

Cluster	Jumlah Anggota
Cluster 1	115
Cluster 2	38
Cluster 3	20



Gambar 6. Visualisasi Cluster

3.5. Analisis Centroid

Nilai centroid digunakan untuk memahami karakteristik masing-masing cluster.

Tabel 4. Nilai Centroid

Cluster	Q	R	F	M
C1	0.057	0.144	0.089	0.063
C2	0.103	0.697	0.013	0.036
C3	0.074	0.074	0.571	0.556

3.6. Interpretasi Cluster

Berdasarkan nilai centroid, segmentasi memberikan pembagian karakteristik yang berbeda (Hendrawan & Gunadi, 2025), masing-masing cluster dapat diinterpretasikan sebagai berikut:

- Cluster 3** merupakan supplier prioritas karena memiliki nilai *Frequency* dan *Monetary* yang tinggi.
- Cluster 1** merupakan supplier menengah dengan aktivitas transaksi yang cukup stabil.
- Cluster 2** merupakan supplier berisiko karena memiliki frekuensi transaksi yang rendah.

3.7. Silhouette Cluster

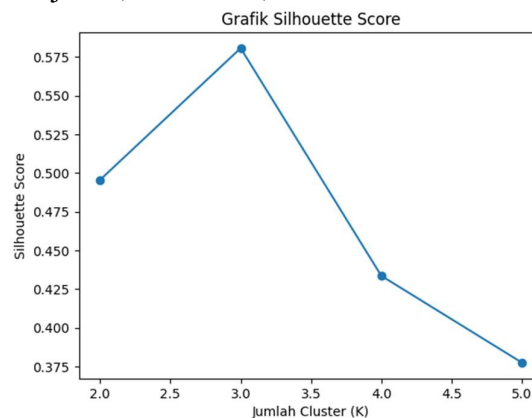
Analisis kluster dalam penelitian ini dievaluasi menggunakan metode *Silhouette Coefficient* untuk menentukan jumlah kluster (K) yang paling optimal dalam merepresentasikan struktur data. metode ini mengukur seberapa mirip sebuah objek dengan klusternya sendiri (kohesi) dibandingkan dengan kluster lainnya.

1. Analisis Perbandingan Nilai Silhouette Score

Berdasarkan hasil pengujian terhadap nilai $K=2$ hingga $K=5$, diperoleh fluktuasi nilai Silhouette Score sebagaimana tersaji pada Gambar 7 (Grafik Silhouette Score). Data menunjukkan bahwa nilai rata-rata silhouette tertinggi dicapai pada saat $K=3$, dengan skor sebesar 0,5806.

Secara rinci, perbandingan skor untuk setiap nilai K adalah sebagai berikut:

- Pada $K = 2$, skor tercatat sebesar 0,4955
- Pada $K = 3$, terjadi peningkatan signifikan menjadi 0,5806, yang mengindikasikan struktur kluster yang lebih kuat dan stabil.
- Pada $K = 4$ dan $K = 5$, nilai skor mengalami penurunan tajam masing-masing menjadi 0,4336 dan 0,3775.



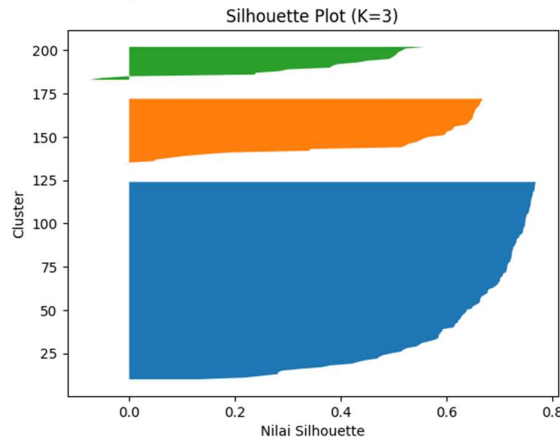
Gambar 7. Grafik Silhouette Score

Penurunan nilai setelah $K=3$ menunjukkan bahwa penambahan jumlah kluster justru memperlemah ikatan antar-anggota dalam satu kelompok dan menciptakan tumpang tindih (overlap) antar-kluster. Oleh karena itu,

berdasarkan grafik elbow pada skor silhouette ini, K=3 ditetapkan sebagai jumlah klaster terbaik karena memiliki nilai yang paling mendekati 1.

2. Analisis Visual Silhouette Plot (K = 3)

Untuk memvalidasi konsistensi hasil pada K = 3, dilakukan analisis mendalam melalui Gambar 8 (Silhouette Plot K=3). Grafik ini memberikan gambaran visual mengenai distribusi ketebalan dan kedalaman setiap klaster:



Gambar 8. Silhouette Plot K=3

Melalui integrasi data statistik dan visualisasi plot, penelitian ini menyimpulkan bahwa pembagian data ke dalam tiga klaster (K=3) menghasilkan konfigurasi yang paling optimal. Model ini mampu menyeimbangkan antara kepadatan internal klaster dan jarak antar-klaster secara efektif, sehingga interpretasi data pada tahap selanjutnya akan menjadi lebih akurat.

3.8. Validasi Model

Validasi dilakukan dengan membandingkan hasil clustering menggunakan RapidMiner dan Google Colab. Hasil menunjukkan bahwa kedua platform menghasilkan pola cluster yang konsisten, meskipun terdapat sedikit perbedaan nilai evaluasi (Fadhillah et al., 2025).

Tabel 5. Perbandingan Hasil Clustering

Aspek	RapidMiner	Google Colab
Metode	K-Means	K-Means
Atribut	QRFM	QRFM
Jumlah Cluster	3	3
Nilai DBI	0,377	0,642
Pola Cluster	Konsisten	Konsisten
Keterangan	Cluster terbentuk dengan baik	Hasil lebih optimal

Berdasarkan Tabel 5, hasil clustering menggunakan RapidMiner dan Google Colab menunjukkan jumlah cluster yang sama, yaitu tiga cluster, dengan pola segmentasi yang konsisten. Namun demikian, terdapat perbedaan nilai evaluasi, di mana RapidMiner menghasilkan nilai DBI yang lebih kecil, yaitu 0,377, dibandingkan dengan Google Colab sebesar 0,642. Hal ini menunjukkan bahwa hasil clustering pada RapidMiner lebih optimal.

4. Kesimpulan dan Saran

Berdasarkan hasil penelitian, metode K-Means dengan pendekatan QRFM berhasil membentuk tiga cluster supplier yang menunjukkan perbedaan karakteristik transaksi secara jelas, yaitu supplier prioritas, supplier menengah, dan supplier berisiko. Pembagian ini mengindikasikan bahwa perusahaan dapat menggunakan hasil segmentasi untuk menetapkan strategi pengelolaan yang berbeda pada tiap kelompok, misalnya mempertahankan supplier prioritas, memantau supplier menengah, dan mengevaluasi supplier berisiko secara lebih intensif. Evaluasi menggunakan Davies-Bouldin Index menunjukkan bahwa model RFM memperoleh nilai terbaik sebesar 0,357, sedangkan QRFM sebesar 0,377, sehingga secara matematis RFM sedikit lebih optimal. Meskipun demikian, QRFM tetap memberikan keunggulan analitis karena mempertimbangkan dimensi Quantity sehingga interpretasi perilaku transaksi supplier menjadi lebih kaya dan lebih representatif.

Hasil validasi pada RapidMiner dan Google Colab menunjukkan pola cluster yang konsisten, yang ditunjukkan oleh kesamaan komposisi anggota cluster dan kemiripan karakteristik centroid antarplatform. Perbedaan nilai DBI, yaitu 0,377 pada RapidMiner dan 0,642 pada Google Colab, mengindikasikan bahwa kualitas pemisahan cluster pada RapidMiner lebih baik, tetapi tidak berarti struktur segmentasi berubah secara substansial. Dengan demikian, konsistensi hasil lebih tepat diartikan sebagai kesamaan struktur segmentasi dan kecenderungan karakteristik cluster, bukan kesamaan nilai evaluasi yang identik. Oleh karena itu, pendekatan QRFM tetap layak digunakan sebagai dasar segmentasi supplier karena mampu menghasilkan pengelompokan yang stabil sekaligus informatif untuk mendukung pengambilan keputusan.

Berdasarkan hasil penelitian, pengembangan selanjutnya disarankan untuk membandingkan QRFM dengan metode segmentasi lain seperti Hierarchical Clustering, DBSCAN, atau pendekatan berbasis Gaussian Mixture Model agar dapat diketahui metode yang paling sesuai untuk karakteristik data supplier. Penelitian berikutnya juga disarankan menambahkan variabel lain seperti lead time, tingkat retur, atau performa pengiriman agar hasil segmentasi tidak hanya mencerminkan pola transaksi, tetapi juga kualitas hubungan dengan supplier. Selain itu, perlu dilakukan pengujian sensitivitas terhadap parameter algoritma K-Means, terutama pemilihan centroid awal dan jumlah cluster, untuk meningkatkan kestabilan hasil clustering. Pengembangan dashboard interaktif yang menampilkan hasil segmentasi secara visual juga direkomendasikan agar hasil penelitian lebih mudah digunakan dalam pengambilan keputusan operasional.

Daftar Pustaka

- Ariati, I., Norsa, R. N., Akhsan, L., & Heikal, J. (2024). Segmentasi Pelanggan Menggunakan K-Means Clustering Studi Kasus Pelanggan Uht Milk Greenfield. *Cerdika Jurnal Ilmiah Indonesia*, 4(6), 429–432.
- Asana, I. M. D. P., Radhitya, M. L., Yogantara, D. N., Sudiantara, I. G., & Hartawan, I. P. N. (2025). Customer Segmentation Using RFM Model and Fuzzy C-Means at PT SNS 21 Bali. *Jurnal Krisnadana (Jurnal Komputer, Sistem Kendali & Jaringan)*, 5(1), 45–56. <https://doi.org/10.58982/krisnadana.v5i1.929>
- Baenil Huda, & Bayu Priyatna. (2020). Penggunaan Aplikasi Content Management System (CMS) Untuk Pengembangan Bisnis Berbasis E-Commerce. *Systematics*, 1(2), 81–88.
- Carudin. (2021). Pemanfaatan Data Transaksi Untuk Dasar Membangun Strategi Berdasarkan Karakteristik Pelanggan Dengan Algoritma K-Means Clustering dan Model RFM. *Jurnal Teknologi Terpadu*, 7(1), 15–22.

- <https://media.neliti.com/media/publications/493730-water-ph-and-turbidity-control-system-in-0a553e14.pdf>
- Fadhillah, M. F., Suyoso, A. L. A., & Puspitasari, I. (2025). *Segmentasi Pelanggan dengan Algoritma Clustering Berdasarkan Atribut Recency, Frequency dan Monetary (RFM)*. 5(1), 48–56.
- Hananto, A. L., Assiroj, P., Priyatna, B., Nurhayati, Ahmad, F., Rahman, A. Y., & Hilabi, S. S. (2021). Analysis Of Drug Data Mining With Clustering Technique Using K-Means Algorithm Analysis Of Drug Data Mining With Clustering Technique Using K-Means Algorithm. *IOP Publishing*. <https://doi.org/10.1088/1742-6596/1908/1/012024>
- Hartati, A. S. (2020). Penerapan Algoritma K-Means Dalam Pemetaan Pelanggan Potensial Menggunakan Model Recency Frequency Monetary. *Jurnal Ekonomi*, 2(1), 41–49.
- Hendrawan, M. B., & Gunadi, G. (2025). Segmentasi Layanan Jasa Kantor Akuntan Publik Menggunakan Algoritma K-Means Clustering. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 7(2), 268–276. <https://doi.org/10.47233/jteksis.v7i2.1912>
- Kristanti, B. T., Junaidi, A., & Mandyartha, E. P. (2024). Implementasi K-Means Clustering Dalam Segmentasi Pelanggan Berdasarkan Usia, Pendapatan, Dan Model Rfm (Studi Kasus: Lantikya Store Jombang). *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(3). <https://doi.org/10.23960/jitet.v12i3.4677>
- Maresti, F. A., Rahayu, W. I., Lustin, M. B. C., & Pakpahan, T. H. (2025). Implementasi K-Means untuk Melakukan Segmentasi Produk Berdasarkan Data Transaksi Retail. *Jurnal Ilmiah Sains Dan Teknologi*, 9(1), 20–32. <https://doi.org/10.47080/saintek.v9i1.3856>
- Panjaitan, T. P., Dwi Akbar, A., Rahmah, S. N., Ernadi, S. C., Sitorus, U. V., Fatoni, M. A., & Inayah, F. (2025). Klasterisasi Data Penjualan Menggunakan Algoritma K-Mean Dengan RapidMiner. *Jurnal of Data Science Methods and Applications*, 01(01), 25–32. <https://doi.org/10.30873/jodmapps.v1i1.pp25-32>
- Permana, M. D., Hananto, A. L., Novalia, E., Huda, B., & Tukino. (2023). Klasterisasi Data Jamaah Umrah pada Tanurmutmainah Tour. *Jurnal KomtekInfo*, 10, 15–20. <https://doi.org/10.35134/komtekinfo.v10i1.332>
- Syahra, Y., Fadlil, A., & Yuliansyah, H. (2025). Customer Segmentation Using RFM and K-Means Clustering to Support CRM in Retail Industry. *Sinkron*, 9(3), 1120–1131. <https://doi.org/10.33395/sinkron.v9i3.14907>
- Triantara, I. K. Y., Raharja, M. A., & Sarasvananda, I. B. G. (2025). Segmentasi Pelanggan Berbasis RFMT Menggunakan K-Means dan Hierarchical Clustering. *Jurnal Nasional Teknologi Informasi Dan Aplikasinya*, 3(4), 917–924.