

## Analisis RFM dan K-Means Clustering untuk Segmentasi Pelanggan pada PT. Sanutama Bumi Arto

Wiendha Arteka Silamantha<sup>1</sup>, Kristophorus Hadiono<sup>2</sup>  
<sup>1,2</sup>Universitas Stikubank, Indonesia

E-mail : <sup>1</sup>wiendhaartekasilamantha@mhs.unisbank.ac.id,  
<sup>2</sup>kristophorus.hadiono@edu.unisbank.ac.id

### Abstract

*This research aims to segment customers at PT. Sanutama Bumi Arto by applying RFM (Recency, Frequency, Monetary) analysis combined with the K-Means Clustering algorithm. RFM analysis is used to identify customer purchasing characteristics based on recency, frequency of purchases and total purchase value (monetary). Then, the K-Means algorithm is used to group customers into different segments based on the similarity of RFM characteristics. This research uses customer transaction data from PT. Sanutama Bumi Arto. The research results show that there are two customer clusters with different characteristics, namely customers with low purchasing levels and customers with high purchasing levels. Customer clusters with high purchasing levels have higher recency, frequency and monetary values compared to customer clusters with low purchasing levels. Cluster evaluation was carried out using the Silhouette Score (0.44), WSS (972.19) and BSS (1112.73) metrics, which shows that clustering has good performance. It is hoped that the results of this research can provide valuable insight for PT. Sanutama Bumi Arto in understanding customer behavior and developing more effective marketing strategies.*

**Keywords:** customer segmentation, RFM analysis, K-Means, clustering

### Abstrak

*Penelitian ini bertujuan untuk melakukan segmentasi pelanggan pada PT. Sanutama Bumi Arto dengan menerapkan analisis RFM (Recency, Frequency, Monetary) yang dikombinasikan dengan algoritma K-Means Clustering. Analisis RFM digunakan untuk mengidentifikasi karakteristik pembelian pelanggan berdasarkan (recency), frekuensi pembelian (frequency), dan nilai total pembelian (monetary). Kemudian, algoritma K-Means digunakan untuk mengelompokkan pelanggan ke dalam segmen-segmen yang berbeda berdasarkan kemiripan karakteristik RFM. Penelitian ini menggunakan data transaksi pelanggan PT. Sanutama Bumi Arto. Hasil penelitian menunjukkan bahwa terdapat dua cluster pelanggan dengan karakteristik yang berbeda, yaitu pelanggan dengan tingkat pembelian yang rendah dan pelanggan dengan tingkat pembelian yang tinggi. Cluster pelanggan dengan tingkat pembelian yang tinggi memiliki nilai recency, frequency, dan monetary yang lebih tinggi dibandingkan dengan cluster pelanggan dengan tingkat pembelian yang rendah. Evaluasi cluster dilakukan dengan metrik Silhouette Score(0,44), WSS(972,19) dan BSS(1112,73), yang menunjukkan bahwa clustering berhasil dilakukan dengan baik. Hasil penelitian ini diharapkan dapat memberikan wawasan yang berharga bagi PT. Sanutama Bumi Arto dalam memahami perilaku pelanggan dan mengembangkan strategi pemasaran yang lebih efektif.*

**Kata Kunci:** segmentasi pelanggan, analisis RFM, K-Means, clustering

## 1. Pendahuluan

PT. Sanutama Bumi Arto adalah perusahaan terkemuka yang bergerak di bidang penjualan pompa air, untuk melayani kebutuhan pelanggan yang tersebar di berbagai wilayah di Indonesia. Seiring dengan pertumbuhan usahanya, telah terkumpul data

transaksi pelanggan yang meliputi informasi pembelian, preferensi produk, dan riwayat interaksi. Namun, data yang terakumulasi ini belum dimanfaatkan secara optimal untuk memahami karakteristik dan perilaku pelanggan secara mendalam.

Pemahaman mendalam tentang pelanggan merupakan kunci untuk mengembangkan strategi pemasaran yang efektif dan meningkatkan kepuasan pelanggan[1], [2]. Salah satu bentuk pemahaman mendalam tentang pelanggan adalah dengan melakukan segmentasi pelanggan. Segmentasi pelanggan menjadi pendekatan yang memungkinkan perusahaan untuk mengelompokkan pelanggan ke dalam kelompok-kelompok yang memiliki karakteristik dan kebutuhan yang serupa[3], [4]. Berdasarkan informasi yang diperoleh mengenai segmen pelanggan secara spesifik, maka PT. Sanutama Bumi Arto dapat merancang penawaran produk, program loyalitas, dan komunikasi pemasaran yang lebih personal dan relevan.

*Clustering* merupakan salah satu metode yang umum digunakan untuk segmentasi pelanggan. *Clustering* sendiri merupakan sebuah teknik dalam pembelajaran mesin yang digunakan untuk mengelompokkan obyek-obyek data ke dalam beberapa kelompok (*cluster*) berdasarkan kemiripan karakteristiknya. Tujuannya adalah untuk menemukan struktur yang tersembunyi dalam data dan mengelompokkan objek-objek yang serupa ke dalam satu kelompok yang sama, sementara objek-objek yang berbeda ditempatkan pada kelompok yang berbeda. *Clustering* merupakan salah satu metode *unsupervised learning*, artinya algoritma belajar untuk mengelompokkan data tanpa adanya label atau target yang telah ditentukan sebelumnya.

*K-Means clustering* adalah metode yang umum digunakan untuk segmentasi pelanggan[5]. Dalam metode ini, pelanggan dikelompokkan ke dalam sejumlah *cluster* yang telah ditentukan sebelumnya ( $K$ ). Algoritma ini bekerja dengan cara mencari pusat *cluster* (*centroid*) dan kemudian menetapkan setiap pelanggan ke *cluster* terdekat berdasarkan jarak antara pelanggan dan *centroid*[6]. *K-Means* dipilih karena kemudahan implementasinya, kecepatan komputasi yang efisien[7], dan kemampuannya dalam menangani dataset yang besar[8]. Selain itu, *cluster* yang dihasilkan *K-Means* mudah diinterpretasikan, sehingga memudahkan perusahaan dalam memahami karakteristik setiap segmen pelanggan.

Model RFM (*Recency, Frequency, Monetary*) adalah metode yang banyak digunakan untuk menganalisis perilaku pembelian pelanggan. Model ini mengelompokkan pelanggan ke dalam segmen-segmen berdasarkan seberapa baru pembelian terakhir mereka (*Recency*), seberapa sering mereka membeli (*Frequency*), dan seberapa banyak uang yang mereka habiskan (*Monetary*)[9].

Setelah nilai RFM dihitung, algoritma *K-Means* digunakan untuk mengelompokkan pelanggan ke dalam sejumlah *cluster* yang telah ditentukan.

Penelitian terkait segmentasi pelanggan telah banyak dilakukan sebelumnya dengan menggunakan model RFM yang dikombinasikan dengan *K-Means* untuk menganalisis segmentasi pelanggan pada perusahaan properti PT. Pollux Aditama Kencana. Hasil penelitian menunjukkan bahwa model gabungan ini lebih unggul daripada hanya menggunakan model RFM saja. Penelitian ini berhasil mengidentifikasi tiga kelompok pelanggan dengan tingkat loyalitas yang berbeda (tinggi, sedang, dan rendah) serta memberikan rekomendasi strategi pemasaran yang sesuai untuk masing-masing kelompok[10].

Kombinasi model RFM dan *K-Means Clustering* juga digunakan untuk untuk segmentasi pelanggan pada UKM Gemilang Kencana, produsen manisan buah carica. Penelitian ini berhasil mengidentifikasi tiga kelompok pelanggan utama, yaitu pelanggan umum (*typical customer*), superstar, dan pelanggan tidak aktif (*dormant customer*), masing-masing dengan karakteristik pembelian yang berbeda. Hasil segmentasi ini memberikan wawasan berharga bagi UKM untuk merancang strategi pemasaran yang lebih efektif[11].

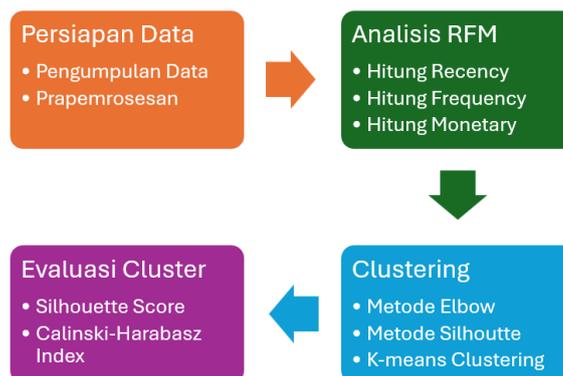
Penelitian selanjutnya membahas segmentasi pelanggan pada PT. Coversuper Indonesia Global, perusahaan aksesoris otomotif, dengan menggunakan kombinasi model RFM dan K-Means *Clustering*. Hasilnya adalah empat segmen pelanggan dengan karakteristik yang berbeda. Penelitian ini menunjukkan bahwa model gabungan RFM dan K-Means dapat menghasilkan segmentasi pelanggan yang akurat dan bermanfaat bagi perusahaan dalam menentukan strategi pemasaran yang tepat [12].

Penelitian-penelitian ini menunjukkan bahwa analisis RFM yang dikombinasikan dengan K-Means *Clustering* merupakan pendekatan yang efektif dan sering digunakan dalam segmentasi pelanggan. Berpijak dari hasil penelitian sebelumnya penelitian ini dilakukan dengan menerapkan analisis RFM dan K-Means *Clustering* pada data transaksi pelanggan PT. Sanutama Bumi Arto.

Fokus dari penelitian ini adalah pengelompokan pelanggan berdasarkan karakteristik pembelian pompa air. Hal inilah yang membedakan penelitian ini dari penelitian sebelumnya yang umumnya berfokus pada produk ritel atau FMCG (*Fast-Moving Consumer Goods*). Selain itu, penelitian ini berusaha menggali lebih dalam karakteristik setiap segmen pelanggan untuk merancang strategi pemasaran yang lebih spesifik dan sesuai dengan kebutuhan pelanggan di industri pompa air.

## 2. Metodologi Penelitian

Penelitian ini dilakukan dengan menerapkan pendekatan kuantitatif dan memanfaatkan data transaksi pelanggan tahun 2019 PT. Sanutama Bumi Arto. Tahapan penelitian yang digunakan pada penelitian ini seperti ditunjukkan pada Gambar 1.



**Gambar 1.** Tahapan Penelitian

Tahap awal penelitian ini adalah pengumpulan dan prapemrosesan data. Pada tahap ini dilakukan pengumpulan data. Data pelanggan yang relevan, meliputi informasi transaksi seperti tanggal pembelian, kode pelanggan, nama pelanggan dan total nilai pembelian dikumpulkan. Data yang terkumpul akan diproses melalui tahap prapemrosesan untuk membersihkan data dari nilai yang hilang atau tidak relevan, serta melakukan transformasi dan standardisasi data untuk memastikan kesesuaian dengan analisis selanjutnya. Contoh data transaksi penjualan seperti ditunjukkan pada Tabel 1.

**Table 1.** Contoh Data Transaksi Penjualan

No	Tanggal	Kode Pelanggan	Nama Pelanggan	Total Pembelian
1	8/9/2019	CS458	Agung Teknik	7250000
2	3/29/2019	CS438	TB Jaya Teknik	3444000
3	2/7/2019	CS396	TB Alex Putra	750000
4	2/6/2019	CS394	TB Harapan	1550000
5	1/12/2019	CS293	TB Sedulur	5500000

Setelah tahap praproses selesai, tahap selanjutnya data disiapkan untuk analisis RFM (*Recency, Frequency, Monetary*). Tahap ini meliputi perhitungan nilai *Recency* (waktu sejak pembelian terakhir), *Frequency* (frekuensi pembelian), dan *Monetary* (total nilai pembelian) untuk setiap pelanggan. Contoh data hasil perhitungan RFM dapat dilihat pada Tabel 2.

**Table 2.** Contoh Perhitungan RFM

No	Kode Pelanggan	Recency	Frequency	Monetary
1	CS458	34	14	37145000
2	CS438	35	21	13966000
3	CS396	64	4	8232500
4	CS394	62	3	2130000
5	CS293	133	4	7005000

Hasil analisa RFM digunakan untuk menilai pelanggan berdasarkan ketiga metrik tersebut (*Recency, Frequency, Monetary*). Setiap pelanggan akan diberi skor RFM berdasarkan peringkat mereka dalam setiap metrik. Hasil skor RFM merupakan dasar untuk segmentasi pelanggan selanjutnya. Contoh data skor RFM ditunjukkan pada Tabel 3.

**Table 3.** Contoh Data Skor RFM

No	Kode Pelanggan	R Score	F Score	M Score	RFM Score
1	CS458	4	5	5	455
2	CS438	4	5	4	454
3	CS396	3	3	4	334
4	CS394	3	2	2	322
5	CS293	2	3	4	234

Algoritma K-Means diterapkan untuk melakukan pengelompokan pelanggan ke dalam segmen-segmen yang berbeda berdasarkan skor RFM mereka. Jumlah *cluster* (K) yang optimal akan ditentukan menggunakan metode *Elbow* dan *Silhouette*. Metode *Elbow* melibatkan plot jumlah *cluster* terhadap inersia (jumlah kuadrat jarak antara titik data dan centroid *cluster*). Titik "siku" pada plot ini menunjukkan jumlah *cluster* yang optimal.

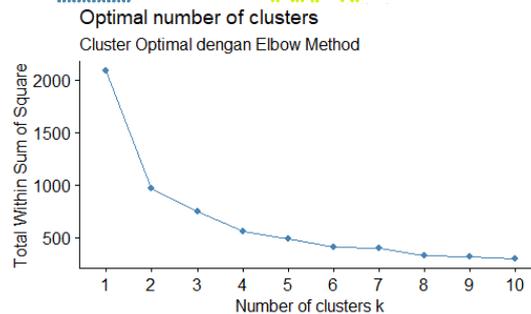
*Silhouette analysis* digunakan untuk memvalidasi kualitas *cluster* dengan mengukur kesamaan setiap titik data dengan *cluster* yang ditetapkan.

Setelah proses *clustering* selesai, kualitas *cluster* yang terbentuk akan dievaluasi lebih lanjut menggunakan *Silhouette score*, *Within-Cluster Sum of Squares* (WSS), dan *Between-Cluster Sum of Squares* (BSS). *Silhouette score* mengukur seberapa mirip setiap pelanggan dengan clusternya sendiri dibandingkan dengan *cluster* lain, memberikan indikasi kepadatan dan pemisahan *cluster*. WSS mengukur kepadatan *cluster* dengan menghitung total jarak kuadrat dalam *cluster*, sedangkan BSS mengukur pemisahan *cluster* dengan menghitung jarak kuadrat antara pusat *cluster*. Evaluasi ini memastikan bahwa *cluster* yang terbentuk memiliki koherensi internal yang tinggi dan perbedaan yang jelas antar *cluster*, sesuai dengan tujuan segmentasi pelanggan. Tahap akhir penelitian adalah melakukan analisis deskriptif terhadap setiap segmen pelanggan yang dihasilkan untuk memahami karakteristik dan kebutuhan masing-masing segmen. Analisis ini akan mencakup distribusi skor RFM, rata-rata nilai transaksi, dan preferensi produk.

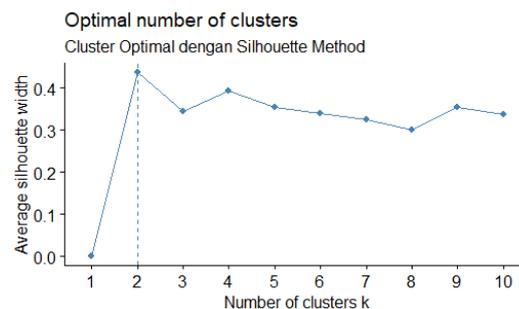
### 3. Hasil dan Pembahasan

Sebelum proses *clustering* k-means dilakukan, penentuan jumlah *cluster* optimal perlu dilakukan terlebih dahulu untuk mendapatkan hasil pengelompokan yang terbaik. Dua metode yang digunakan dalam menentukan jumlah *cluster* optimal adalah metode *Elbow* dan *Silhouette*. Berdasarkan hasil visualisasi pada Gambar 2

(*Elbow method*) dan Gambar 3 (*Silhouette method*), dapat disimpulkan bahwa jumlah *cluster* optimal untuk segmentasi pelanggan adalah 2.



**Gambar 2.** Metode *Elbow*



**Gambar 3.** Metode *Silhouette*

Setelah menentukan jumlah *cluster* optimal, proses selanjutnya adalah penerapan *clustering k-means* pada data RFM. Jumlah *cluster* optimal yang diperoleh dari proses sebelumnya (2 *cluster*) digunakan, dan nilai *nstart* diatur dengan nilai 25 untuk menjalankan algoritma sebanyak 25 kali sebagai inisialisasi pusat *cluster* yang berbeda. Hasil pusat *cluster* ditampilkan pada Tabel 4.

Tabel 4 memberikan informasi bahwa *cluster* pertama memiliki nilai rata-rata *R\_Score* sebesar 1.93, *F\_Score* sebesar 1.72, dan *M\_Score* sebesar 1.88. Hal tersebut mengindikasikan bahwa pelanggan dalam kelompok ini cenderung jarang melakukan pembelian, dengan nilai transaksi yang relatif rendah, dan tidak terlalu baru dalam melakukan pembelian terakhir. Pelanggan dalam *cluster* ini dapat dikategorikan sebagai pelanggan dengan aktivitas rendah dan kontribusi finansial yang lebih sedikit terhadap perusahaan.

Sebaliknya, *cluster* kedua menunjukkan nilai rata-rata *R\_Score* sebesar 3.84, *F\_Score* sebesar 4.02 dan *M\_Score* sebesar 3.88. Ini mengindikasikan bahwa pelanggan dalam kelompok ini melakukan pembelian lebih sering, dengan nilai transaksi yang lebih tinggi, dan melakukan transaksi pembelian yang lebih baru dibandingkan dengan *cluster* pertama. Pelanggan dalam *cluster* ini dapat dianggap sebagai pelanggan yang sangat aktif dan memberikan kontribusi finansial yang signifikan bagi perusahaan.

**Table 4.** Hasil Pusat *Cluster*

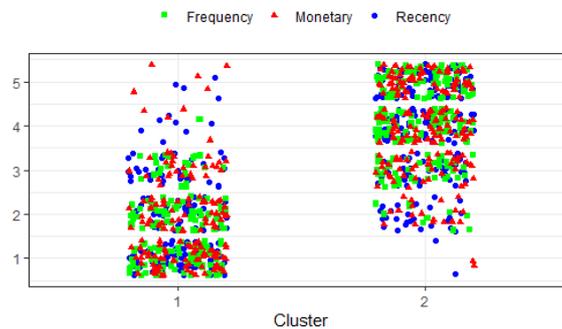
No	Cluster	R Score	F Score	M Score
1	1	1,93	1,72	1,88
2	2	3,84	4,02	3,88

Hasil perhitungan K-Means berdasarkan perhitungan metode RFM (*Recency*, *Frequency*, dan *Monetary*) divisualisasikan seperti Gambar 4. Dari Gambar 4 dapat dijelaskan bahwa pada sumbu x, terdapat dua *cluster*, yaitu *cluster 1* dan *cluster 2*. Sumbu y menunjukkan nilai RFM, dengan warna dan bentuk yang berbeda untuk setiap

komponen: *Frequency* ditandai dengan persegi hijau, *Monetary* dengan segitiga merah, dan *Recency* dengan lingkaran biru.

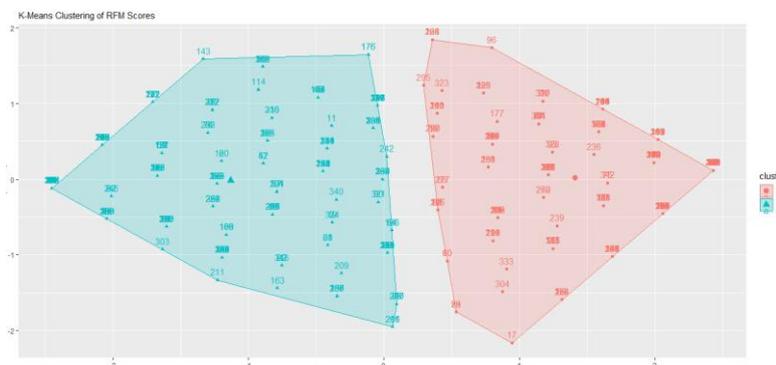
Data pada *cluster* 1 menunjukkan bahwa sebagian besar pelanggan dalam *cluster* ini memiliki nilai RFM yang lebih rendah. Nilai *Recency* berkisar antara 1 hingga 3, *Frequency* sebagian besar berada pada nilai 1 dan 2, dan *Monetary* juga berkisar antara 1 hingga 2, dengan beberapa pengecualian yang memiliki nilai lebih tinggi. Hal ini mencerminkan bahwa pelanggan dalam *cluster* ini tidak sering melakukan pembelian, nilai pembeliannya rendah, dan pembelian terakhir mereka sudah cukup lama.

Sedangkan data pada *cluster* 2 memiliki nilai RFM yang lebih tinggi. Nilai *Recency* sebagian besar berada pada nilai 4 hingga 5, *Frequency* berkisar antara 3 hingga 5, dan *Monetary* juga menunjukkan nilai yang lebih tinggi. Hal ini mengindikasikan bahwa pelanggan dalam *cluster* tersebut lebih sering melakukan pembelian, dengan nilai transaksi yang lebih tinggi, dan pembelian terakhir mereka terjadi lebih baru.



**Gambar 4.** Distribusi RFM

Untuk mendapatkan Gambaran lebih detail, kemudian dilakukan visualisasi pengelompokan seperti pada Gambar 5 dimana *cluster* 1 memiliki warna merah, sedangkan *cluster* 2 berwarna biru muda dan *scatter plot* seperti pada Gambar 5.



**Gambar 5.** Visualisasi Warna Cluster

Pada visualisasi dengan *scatter plot* didapatkan hasil bahwa :

***Recency vs. Frequency***

- a) *Cluster* dengan warna merah cenderung memiliki nilai frekuensi yang lebih tinggi dibandingkan dengan *cluster* biru.
- b) *Cluster* biru memiliki nilai *recency* yang lebih kecil, menunjukkan pelanggan di *cluster* ini lebih baru bertransaksi.

***Recency vs. Monetary***

- a) *Cluster* biru memiliki nilai *recency* yang lebih kecil, dan nilai *monetary* yang lebih tinggi. Ini menunjukkan bahwa pelanggan di *cluster* biru lebih baru dan memiliki pengeluaran yang lebih tinggi.

b) *Cluster* merah memiliki nilai *monetary* yang lebih rendah, tetapi rentang *recency* yang lebih luas.

**Frequency vs. Monetary**

Kedua *cluster* menunjukkan korelasi positif antara *frequency* dan *monetary*, meskipun *cluster* biru memiliki pelanggan dengan transaksi frekuensi dan pengeluaran yang lebih tinggi. Analisis *Density Plot* dari Gambar 6., menunjukkan bahwa :

**Recency**

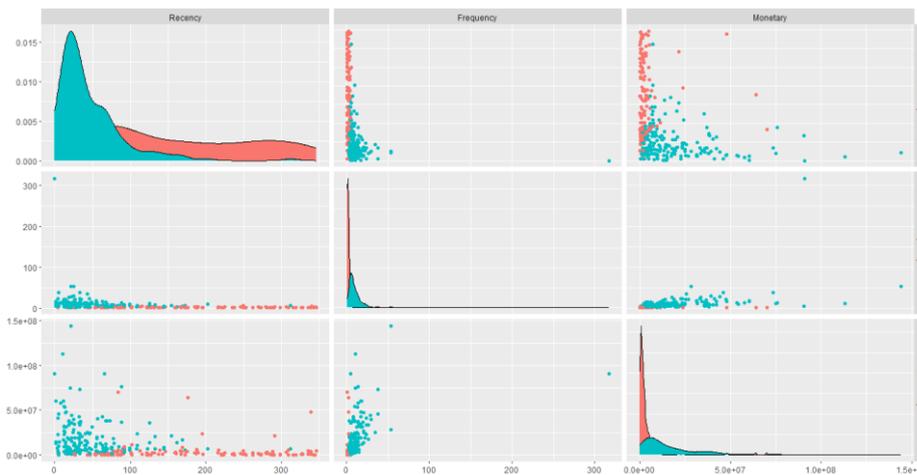
- a) Distribusi *cluster* biru terkonsentrasi pada nilai *recency* yang lebih rendah, menunjukkan pelanggan yang lebih baru bertransaksi.
- b) Distribusi *cluster* merah lebih merata dengan rentang *recency* yang lebih besar.

**Frequency**

Distribusi *cluster* merah terkonsentrasi pada nilai *frequency* yang lebih rendah, sedangkan *cluster* biru menunjukkan distribusi yang lebih merata pada nilai *frequency* yang lebih tinggi.

**Monetary**

- a) *Cluster* biru memiliki puncak yang lebih tinggi pada nilai *monetary*, menunjukkan bahwa pelanggan di cluster ini cenderung berbelanja lebih banyak.
- b) *Cluster* merah memiliki distribusi *monetary* yang lebih rendah dan lebih merata.



**Gambar 6.** Visualisasi *Scatter Plot*

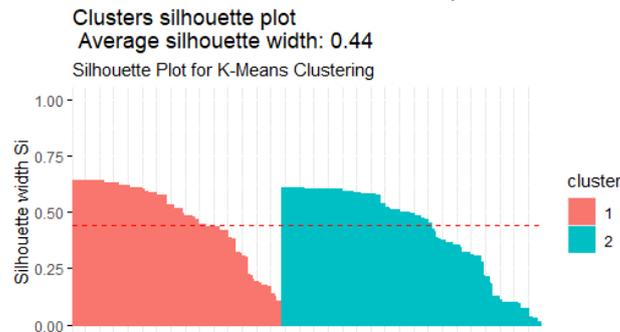
Dari Gambar 5 dan Gambar 6 dapat dilakukan analisis visualisasi seperti yang telah dijelaskan sebelumnya. Hasil analisis tersebut dapat digunakan sebagai acuan bagi perusahaan untuk mengembangkan strategi pemasaran yang berbeda untuk masing-masing *cluster*. Bagi *cluster* pelanggan kurang aktif yang masuk dalam *cluster* 1, dapat diterapkan strategi reaktivasi. Strategi tersebut bertujuan mengajak pelanggan untuk kembali melakukan transaksi dengan menawarkan diskon khusus. Sedangkan untuk pelanggan loyal seperti pada *cluster* 2, dapat ditawarkan program loyalitas pelanggan dengan penawaran eksklusif untuk meningkatkan pengalaman pelanggan.

Untuk mengevaluasi kualitas *clustering* dengan *k-means*, digunakan 3 metrik yaitu *silhouette score*, *Within-Cluster Sum of Squares(WSS)* dan *Between-Cluster Sum of Squares(BSS)*. Hasil evaluasi ditunjukkan pada tabel 5, dan visualisasi *silhouette score* pada Gambar 7.

**Table 5.** Hasil Metrik Evaluasi

No	Metode	Nilai
1	Sillhoutte Score	0,44
2	WSS	972,19
3	BSS	1112,73

*Silhouette score* mengukur seberapa mirip setiap pelanggan dengan *clusternya* sendiri dibandingkan dengan *cluster* lain. Nilai *silhouette* berkisar antara -1 hingga 1. Nilai 0.44 menunjukkan bahwa rata-rata, pelanggan cukup cocok dengan *clusternya* sendiri dan cukup berbeda dari pelanggan di *cluster* lain. Ini mengindikasikan bahwa *clusterisasi* cukup baik, tetapi mungkin masih ada beberapa pelanggan yang ambigu dan bisa saja masuk ke *cluster* lain.



**Gambar 7.** Visualisasi *Silhouette Score*

WSS mengukur total variasi di dalam setiap *cluster*. Nilai WSS yang rendah menunjukkan bahwa pelanggan dalam setiap *cluster* memiliki karakteristik RFM yang relatif serupa. Dalam hal ini, nilai 972.19 menunjukkan variasi yang cukup rendah di dalam *cluster*, yang berarti pelanggan dalam setiap *cluster* cenderung memiliki perilaku pembelian yang mirip.

BSS mengukur total variasi antara *cluster* yang berbeda. Nilai BSS yang tinggi menunjukkan bahwa *cluster-cluster* terpisah dengan baik dan memiliki karakteristik yang berbeda satu sama lain. Nilai 1112.73 menunjukkan bahwa kedua *cluster* pelanggan yang dihasilkan memiliki perbedaan yang signifikan dalam hal perilaku pembelian mereka.

#### 4. Kesimpulan

Berdasarkan hasil analisis RFM dan *clusterisasi* dengan penerapan metode K-Means, dengan nilai *silhouette score* 0.44, WSS 972.19, dan BSS 1112.73 dan jumlah *cluster* 2, dapat disimpulkan bahwa model cukup baik dalam mengelompokkan pelanggan ke dalam dua kelompok yang berbeda dan bermakna. Nilai *silhouette score* yang cukup tinggi menunjukkan bahwa objek-objek dalam *cluster* memiliki kemiripan satu sama lain dan berbeda dari objek-objek dalam *cluster* lain. Nilai WSS yang rendah menunjukkan bahwa objek dalam *cluster* berdekatan satu sama lain, sedangkan nilai BSS yang tinggi menunjukkan bahwa *cluster* terpisah jauh satu sama lain.

*Cluster* pertama menunjukkan karakteristik pelanggan dengan tingkat pembelian yang rendah, ditandai dengan nilai *recency*, *frequency*, dan *monetary* yang lebih rendah. Sementara itu, *cluster* kedua menunjukkan karakteristik pelanggan dengan tingkat pembelian yang tinggi, ditandai dengan nilai *recency*, *frequency*, dan *monetary* yang lebih tinggi. Hasil analisis ini memberikan landasan yang kuat bagi PT. Sanutama Bumi Arto untuk memahami perilaku pelanggan dan mengembangkan strategi pemasaran yang lebih efektif. Dengan membedakan pelanggan berdasarkan karakteristik pembelian mereka, perusahaan dapat merancang program pemasaran yang lebih personal dan relevan untuk setiap segmen pelanggan, sehingga dapat meningkatkan kepuasan pelanggan dan pada akhirnya meningkatkan penjualan.

Penelitian selanjutnya, dapat diarahkan pada pengembangan model segmentasi pelanggan yang lebih kompleks dengan mempertimbangkan lebih banyak variabel, seperti demografi pelanggan, preferensi produk, dan saluran pembelian.

## Daftar Pustaka

- [1] N. L. Rane, A. Achari, and S. P. Choudhary, “Enhancing customer loyalty through quality of service: Effective strategies to improve customer satisfaction, experience, relationship, and engagement,” *International Research Journal of Modernization in Engineering Technology and Science*, vol. 5, no. 5, pp. 427–452, 2023.
- [2] M. Ilmi and C. A. Zulkarnain, “Strategi Pemasaran Melalui Promosi, Harga Dan Kualitas Pelayanan Untuk Meningkatkan Kepuasan Konsumen Pada Pt Garuda Indonesia (Persero) Tbk,” *Jurnal Ilmiah Manajemen Ekonomi Dan Akuntansi (JIMEA)*, vol. 1, no. 1, pp. 66–73, 2023.
- [3] K. Tabianan, S. Velu, and V. Ravi, “K-means clustering approach for intelligent customer segmentation using customer purchase behavior data,” *Sustainability*, vol. 14, no. 12, p. 7243, 2022.
- [4] F. Yoseph, N. H. Ahamed Hassain Malim, M. Heikkilä, A. Brezulianu, O. Geman, and N. A. Paskhal Rostam, “The impact of big data market segmentation using data mining and clustering techniques,” *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 5, pp. 6159–6173, 2020.
- [5] R. Siagian, P. Sirait, and A. Halim, “Penerapan Algoritma K-Means dan K-Medoids untuk Segmentasi Pelanggan pada Data Transaksi E-Commerce,” *SISTEMASI: Jurnal Sistem Informasi*, vol. 11, no. 2, 2022.
- [6] M. Schott, “K-Means Clustering Algorithm for Machine Learning ,” Capital One Tech on Medium.
- [7] E. Patimah, E. Ermatita, and N. Chamidah, “Analisis Cluster Kepuasan Pengguna Terhadap Layanan Shopee Menggunakan Algoritma K-Means,” *Informatik: Jurnal Ilmu Komputer*, vol. 17, no. 3, pp. 209–217, 2021.
- [8] C. Sreedhar, N. Kasiviswanath, and P. Chenna Reddy, “Clustering large datasets using K-means modified inter and intra clustering (KM-I2C) in Hadoop,” *J Big Data*, vol. 4, no. 1, 2017, doi: 10.1186/s40537-017-0087-2.
- [9] A. Apichottanakul, M. Goto, K. Piewthongngam, and S. Pathumnakul, “Customer behaviour analysis based on buying-data sparsity for multi-category products in pork industry: A hybrid approach,” *Cogent Eng*, vol. 8, no. 1, 2021, doi: 10.1080/23311916.2020.1865598.
- [10] G. Arseta and H. D. Purnomo, “Analisa Segmentasi Customer Pada Perusahaan Bisnis Properties Menggunakan Model RFM (Kasus PT. Pollux Aditama Kencana),” 2023.
- [11] B. E. Adiana, I. Soesanti, and A. E. Permanasari, “Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model dan Teknik Clustering,” *Jurnal Terapan Teknologi Informasi*, vol. 2, no. 1, pp. 23–32, 2018, doi: 10.21460/jutei.2017.21.76.
- [12] A. T. Widiyanto and A. Witanti, “Segmentasi Pelanggan Berdasarkan Analisis RFM Menggunakan Algoritma K-Means Sebagai Dasar Strategi Pemasaran (Studi Kasus PT Coversuper Indonesia Global),” *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 1, no. 1, pp. 204–215, 2021.