

RFM Analysis dan Segmentasi Pelanggan Enterprise Business Kantor Pos Indonesia Menggunakan Affinity Propagation

Dafa Zain Musyafah¹, Amri Muhaimin^{2*}, Dwi Arman Prasetya³

^{1,2,3}Program Studi Sains Data, Universitas Pembangunan Nasional “Veteran” Jawa Timur, Jl. Rungkut Madya, Surabaya 60294, Indonesia

*Korespondensi Penulis. E-mail: 122083010095@student.upnjatim.ac.id,

2arman.prasetya.sada@upnjatim.ac.id, 3amri.muhamin.stat@upnjatim.ac.id, Telp: +6282335057373

Abstrak

Pelanggan Enterprise Business merupakan segmen strategis yang berkontribusi signifikan terhadap pendapatan Kantor Pos Indonesia, namun keberagaman pola transaksi antar pelanggan korporat menyulitkan perumusan strategi layanan yang tepat sasaran. Penelitian ini bertujuan menerapkan RFM Analysis (Recency, Frequency, Monetary) sebagai tahap rekayasa fitur berbasis data transaksi, yang selanjutnya diintegrasikan dengan algoritma Affinity Propagation (AP) untuk menghasilkan segmentasi pelanggan yang bermakna secara bisnis. Data yang digunakan adalah data transaksi internal Kantor Pos Indonesia periode Januari–Februari 2025 yang mencakup 34.999 observasi dari 1.325 pelanggan korporat unik. Setelah tahap preprocessing meliputi pembersihan data, penghapusan duplikat, penanganan outlier menggunakan metode IQR, dan normalisasi Min-Max diperoleh 1.032 pelanggan yang memenuhi kriteria analisis. Fitur RFM dikonstruksi melalui agregasi per ID Pelanggan: Recency (R) sebagai jumlah hari unik bertransaksi dari variabel Daily Trx, Frequency (F) sebagai total transaksi, dan Monetary (M) sebagai total revenue. Penentuan parameter AP dilakukan melalui grid search pada kombinasi nilai preference dan damping factor. Algoritma AP secara adaptif membentuk 10 klaster tanpa memerlukan penetapan jumlah klaster di awal. Hasil evaluasi menunjukkan Silhouette Score sebesar 0,3766 (kategori cukup baik), Davies-Bouldin Index sebesar 1,0486, dan Calinski-Harabasz Index sebesar 738,73. Klaster yang terbentuk mencerminkan spektrum perilaku yang beragam: Klaster 6 (Pelanggan Dorman, 32,36%) dan Klaster 9 (Pelanggan Tidak Aktif, 21,32%) mendominasi dengan lebih dari separuh total pelanggan, mengindikasikan kebutuhan mendesak akan strategi aktivasi. Sebaliknya, Klaster 7 (Pelanggan Bernilai Tinggi, 5,43%) dan Klaster 1 (Pelanggan Sangat Aktif, 3,39%) menghasilkan rata-rata revenue tertinggi masing-masing sebesar Rp7.525.038 dan Rp5.508.842, menegaskan pentingnya program retensi yang terfokus untuk kedua segmen ini guna mempertahankan kontribusi pendapatan terbesar bagi Kantor Pos Indonesia.

Kata kunci: RFM Analysis, Segmentasi Pelanggan, Affinity Propagation, Enterprise Business, Kantor Pos Indonesia

Abstract

Enterprise Business customers are a strategic segment that significantly contributes to the revenue of Pos Indonesia; however, the diversity of transaction patterns among corporate customers makes it difficult to formulate targeted service strategies. This study aimed to apply RFM Analysis (Recency, Frequency, Monetary) as a feature engineering stage derived from customer transaction data, subsequently integrated with the Affinity Propagation (AP) algorithm to produce business-meaningful customer segmentation. The dataset comprised 34,999 internal transaction records from 1,325 unique corporate customers of Pos Indonesia covering January–February 2025. After preprocessing including data cleaning, duplicate removal, IQR-based outlier handling, and Min-Max normalization 1,032 customers met the analysis criteria. RFM features were constructed per customer ID: Recency (R) as the number of unique transaction days, Frequency (F) as total transactions, and Monetary (M) as total revenue. AP parameters were determined via grid search across preference percentiles and damping factors. The AP algorithm adaptively formed 10 clusters without requiring a predefined cluster count. Evaluation results yielded a Silhouette Score of 0.3766 (satisfactory range), a Davies-Bouldin Index of 1.0486, and a Calinski-Harabasz Index of 738.73. The resulting clusters reflected diverse behavioral spectra: Cluster 6 (Dormant Customers, 32.36%) and Cluster 9 (Inactive

Customers, 21.32%) dominated quantitatively, together comprising over half of all customers and signaling an urgent need for activation strategies. Conversely, Cluster 7 (High-Value Customers, 5.43%) and Cluster 1 (Highly Active Customers, 3.39%) generated the highest average revenue of Rp7,525,038 and Rp5,508,842 respectively, underscoring the critical importance of focused retention programs for these segments.

Keyword: RFM Analysis, Customer Segmentation, Affinity Propagation, Enterprise Business, Pos Indonesia

PENDAHULUAN

Bisnis adalah sebuah kata yang sering kali kita dengar di setiap percakapan orang-orang yang sedang melakukan berbagai macam kegiatan (Mustika et al., 2020). Enterprise business merupakan segmen krusial dalam dunia bisnis, yang berfokus pada layanan bagi pelanggan korporat, agensi, dan organisasi besar. Segmen ini memainkan peran strategis dalam mendukung pertumbuhan perusahaan karena transaksinya biasanya bervolume tinggi, berfrekuensi tinggi, dan berkontribusi signifikan terhadap pendapatan. Setiap instansi dituntut untuk beradaptasi dengan tren baru, dimana hampir semua aspek kini berbasis digital dan memanfaatkan internet sebagai penopang utama (Jannah & Sadiqin, 2021). Oleh karena itu, mengelola pelanggan enterprise memerlukan pendekatan yang berbeda dibandingkan mengelola pelanggan ritel, baik dari segi strategi pemasaran, layanan, maupun manajemen hubungan pelanggan.

Enterprise Business merupakan segmen pelanggan korporat yang memegang peran strategis dalam mendukung pertumbuhan pendapatan perusahaan jasa. Segmen ini mencakup perusahaan besar, instansi pemerintah, dan organisasi yang membutuhkan layanan distribusi, logistik, dan transaksi keuangan dalam skala besar dan berkelanjutan. PT Pos Indonesia (Persero) sebagai Badan Usaha Milik Negara (BUMN) penyedia jasa pos dan logistik terbesar di Indonesia mengelola pelanggan Enterprise Business dengan volume dan frekuensi transaksi yang tinggi (Septyawati, 2024). Keberagaman pola transaksi antar pelanggan korporat yang meliputi frekuensi pengiriman harian, nilai pendapatan, berat kiriman, dan jenis layanan yang digunakan menciptakan tantangan dalam memahami karakteristik masing-masing pelanggan secara mendalam. Tanpa pemahaman yang baik tentang segmentasi pelanggan, perusahaan berisiko mengalokasikan sumber daya secara tidak efisien dan kesulitan mengembangkan strategi pemasaran yang tepat sasaran (Jannah & Sadiqin, 2021).

Klasterisasi pelanggan berbasis data transaksi merupakan pendekatan *unsupervised learning* yang telah terbukti efektif dalam mengidentifikasi pola tersembunyi pada data pelanggan yang kompleks dan heterogen (Maimon & Rokach, 2021). Namun, klasterisasi langsung atas variabel transaksi mentah sering kali menghasilkan segmen yang kurang bermakna secara bisnis karena tidak mempertimbangkan dimensi perilaku pelanggan secara terintegrasi. Untuk mengatasi keterbatasan tersebut, RFM Analysis (*Recency, Frequency, Monetary*) digunakan sebagai tahap rekayasa fitur sebelum proses klasterisasi. Model RFM adalah metode yang populer dan efektif untuk segmentasi pelanggan berdasarkan riwayat data transaksi yang telah digunakan selama lebih dari 50 tahun dalam pemasaran berbasis data (Cheng & Chen, 2009). Metode ini mengkuantifikasi perilaku pelanggan melalui tiga dimensi: *Recency* (kebaruan transaksi), *Frequency* (frekuensi transaksi), dan *Monetary* (nilai moneter transaksi). Penelitian Doğan, Ayçin, & Bulut (2018) membuktikan bahwa segmentasi eksisting yang hanya berbasis pengeluaran (*Monetary*) tidak memadai, dan integrasi RFM dengan algoritma klasterisasi menghasilkan segmen yang 60% berbeda dan jauh lebih representatif. Hosseini, Maleki, & Gholamian, (2010) juga menegaskan bahwa penggunaan fitur RFM sebagai masukan klasterisasi menghasilkan segmen yang lebih informatif dan dapat diinterpretasikan secara bisnis dibandingkan penggunaan variabel transaksi mentah secara langsung.

Berbagai algoritma klasterisasi telah diterapkan pada konteks segmentasi pelanggan berbasis RFM. Algoritma K-Means merupakan yang paling umum digunakan karena efisiensinya (Parikh & Abdelfattah, 2020), namun memiliki kelemahan mendasar yaitu sensitif terhadap inisialisasi pusat

klaster secara acak dan memerlukan penetapan jumlah klaster k sebelum proses dimulai. Hal ini berpotensi menghasilkan klaster yang tidak stabil ketika diterapkan pada data transaksi pelanggan korporat yang distribusinya tidak merata dan memiliki pola yang kompleks (Tzortzis et al., 2003). Sebagai alternatif, Affinity Propagation (AP) yang diperkenalkan oleh Frey & Dueck (2007) menawarkan pendekatan berbasis *message passing* pada matriks *similarity* untuk menentukan eksemplar dan jumlah klaster secara otomatis tanpa memerlukan inisialisasi pusat klaster (Dueck, 2009). Keunggulan AP terletak pada kemampuannya menangkap struktur data yang kompleks dan menghasilkan klaster yang lebih stabil karena eksemplar dipilih langsung dari data asli, bukan dari rata-rata terhitung. Penelitian Mustafa & Jacksi (2023) menunjukkan bahwa AP lebih unggul dalam fleksibilitas pembentukan klaster dibandingkan K-Means, terutama untuk data dengan distribusi yang tidak seragam. Meski penelitian Muhajir & Sari (2018) dan Asriny (2019) telah mengembangkan K-Affinity Propagation (K-AP) varian AP dengan kendali jumlah klaster untuk berbagai domain, penerapan AP murni yang diintegrasikan dengan RFM Analysis dalam konteks segmentasi pelanggan korporat sektor logistik masih sangat terbatas dalam literatur.

Kesenjangan ini menjadi urgensi penelitian yang dituangkan dalam artikel ini. Penelitian ini bertujuan untuk menerapkan RFM Analysis sebagai tahap rekayasa fitur berbasis data transaksi pelanggan Enterprise Business Kantor Pos Indonesia, yang selanjutnya diintegrasikan dengan algoritma Affinity Propagation untuk menghasilkan segmentasi pelanggan yang bermakna secara bisnis. Pilihan ini memberikan keunggulan eksplorasi: AP mampu menemukan jumlah segmen yang paling alami sesuai dengan distribusi perilaku transaksi pelanggan, tanpa bias akibat penetapan k di awal. Dataset yang digunakan mencakup 34.999 observasi pelanggan korporat dengan variabel Daily Trx, Transaksi, dan Revenue. Kualitas segmentasi dievaluasi menggunakan tiga metrik validasi internal: Silhouette Score, Davies–Bouldin Index (DBI), dan Calinski–Harabasz Index (CHI), untuk memberikan gambaran kualitas klaster yang komprehensif. Hasil penelitian diharapkan dapat memberikan landasan berbasis data bagi Kantor Pos Indonesia dalam merancang strategi pemasaran, program retensi pelanggan bernilai tinggi, dan strategi aktivasi pelanggan berkontribusi rendah.

METODE

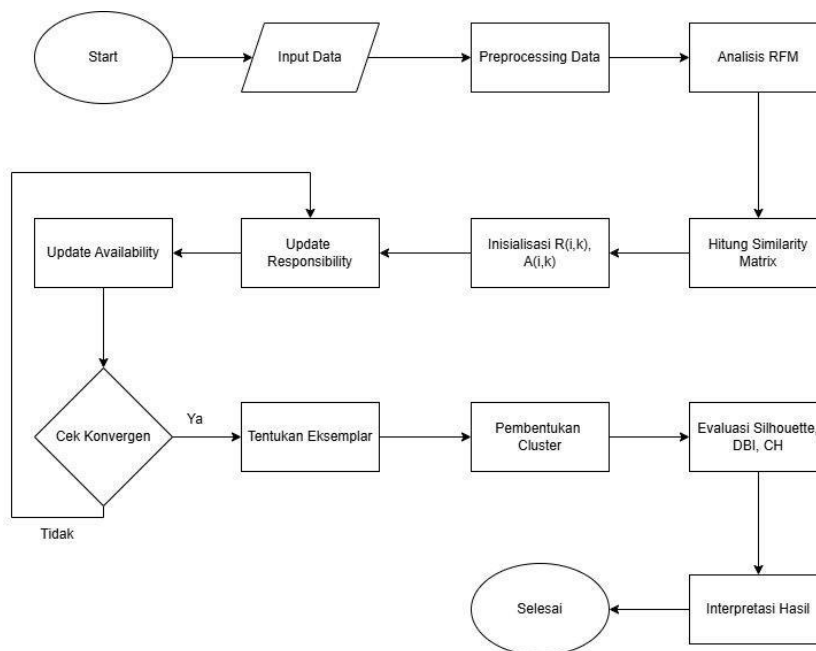
Penelitian ini merupakan penelitian kuantitatif berbasis data sekunder dengan pendekatan komputasional. Sumber data dalam penelitian ini adalah data transaksi internal pelanggan Enterprise Business Kantor Pos Indonesia yang diperoleh secara langsung dari unit bisnis perusahaan. Dataset mencakup 34.999 baris observasi yang masing-masing merepresentasikan satu aktivitas transaksi harian pelanggan korporat. Dataset memuat tujuh variabel: ID Pelanggan (identitas unik pelanggan korporat), Nama Pelanggan Korporat, Daily Trx (tanggal transaksi harian), Transaksi (jumlah transaksi), Revenue (total nilai pendapatan dalam satuan rupiah). Dari lima variabel tersebut, tiga variabel utama Daily Trx, Transaksi, dan Revenue digunakan sebagai basis konstruksi fitur RFM.

Tabel 1. Data Enterprise Business

ID Pelanggan	Nama Pelanggan Korporat	Daily Trx	Transaksi	Revenue
PERCTEMPRI NA05631B	PT TEMPRINA MEDIA GRAFIKA	1/13/2025	1	118,931,241
PERCTEMPRI NA05631B	PT TEMPRINA MEDIA GRAFIKA	1/14/2025	3	2,170,260,950
PERCTEMPRI NA05631B	PT TEMPRINA MEDIA GRAFIKA	1/21/2025	1	14,147,000
...	...	...	...	...

ID Pelanggan	Nama Pelanggan Korporat	Daily Trx	Transaksi	Revenue
BANKBSI05830C	BSI KCP MATARAM MAJAPAHIT	2/27/2025	1	6,000
LNKPP08836A	KPP Pratama Praya	2/4/2025	1	6,000

Data yang digunakan diperoleh dari sistem internal PT Pos Indonesia (Persero), mencakup data transaksi pelanggan Enterprise Business selama periode Januari hingga Februari 2025. Seluruh proses analisis dilakukan menggunakan bahasa pemrograman Python dengan library pendukung pandas, scikit-learn, dan matplotlib. Alur penelitian secara keseluruhan disajikan pada Gambar 1 berikut.



Gambar 1. Alur penelitian RFM Analisis dan Affinity Propagation

Sesuai dengan alur pada Gambar 1, penelitian ini dilaksanakan melalui tujuh tahap berurutan. Setiap tahap dirancang agar output satu tahap menjadi input tahap berikutnya, sehingga menghasilkan pipeline analisis yang sistematis dari data mentah hingga interpretasi segmentasi.

- Preprocessing data. Preprocessing merupakan tahap awal yang dilakukan untuk membersihkan dan mempersiapkan data sebelum digunakan dalam proses analisis lebih lanjut (Mirantika et al., 2023). Preprocessing mencakup: (1) seleksi variabel yang relevan untuk analisis RFM; (2) konversi tipe data, khususnya kolom Daily Trx yang berformat tanggal (datetime) dikonversi secara tepat agar dapat dihitung jumlah hari uniknya; (3) penanganan nilai hilang (*missing values*) dengan menghapus observasi yang memiliki nilai kosong pada variabel utama; serta (4) penghapusan baris duplikat.
- Analisis RFM. Setelah data bersih, konstruksi fitur RFM dilakukan melalui agregasi per ID Pelanggan. Recency (R) dihitung sebagai jumlah hari unik (*nunique*) dari kolom Daily Trx, merepresentasikan intensitas kehadiran harian pelanggan selama periode analisis nilai R

yang lebih besar mengindikasikan pelanggan yang lebih rutin dan aktif. Frequency (F) merupakan penjumlahan total nilai kolom Transaksi per pelanggan, mencerminkan seberapa sering pelanggan melakukan transaksi. Monetary (M) diperoleh dari penjumlahan total Revenue per pelanggan, merepresentasikan kontribusi finansial pelanggan terhadap pendapatan perusahaan.

- c) Perhitungan matriks *similarity*. Matriks *similarity* antar pelanggan dihitung sebagai negatif dari kuadrat jarak Euclidean antar pasang titik data pada ruang fitur RFM yang telah ternormalisasi. Nilai *similarity* yang mendekati nol menunjukkan dua pelanggan yang sangat mirip, sementara nilai yang sangat negatif menunjukkan perbedaan yang besar.
- d) Inisialisasi dan iterasi Affinity Propagation. Algoritma Affinity Propagation bekerja melalui mekanisme *message passing* yang dikirimkan secara iteratif pada matriks *similarity*. Terdapat dua jenis pesan yang saling diperbarui pada setiap iterasi: (i) *Responsibility* $r(i, k)$, yang mengukur seberapa cocok titik data k menjadi eksemplar bagi titik data i dibandingkan eksemplar kandidat lainnya; dan (ii) *Availability* $a(i, k)$, yang mencerminkan seberapa layak titik data i memilih k sebagai eksemplarnya dengan mempertimbangkan dukungan titik data lain terhadap k . Kedua pesan diinisialisasi bernilai nol, kemudian diperbarui secara bergantian pada setiap iterasi menggunakan parameter *damping factor* untuk menjaga stabilitas konvergensi.
- e) Cek konvergensi dan penentuan eksemplar. Proses iterasi berlanjut hingga penugasan eksemplar tidak berubah selama sejumlah iterasi berturut-turut, atau hingga batas maksimum iterasi tercapai. Setelah konvergen, eksemplar ditentukan berdasarkan kombinasi nilai *responsibility* dan *availability*. Setiap titik data lainnya kemudian ditetapkan ke eksemplar terdekat sehingga terbentuk kluster-kluster dengan masing-masing satu eksemplar sebagai representasi pusatnya.
- f) Evaluasi kualitas kluster. Proses evaluasi dilakukan untuk menilai apakah pola, informasi, atau pengetahuan yang dihasilkan sesuai dengan fakta empiris dan hipotesis yang telah dirumuskan sebelumnya (Marlina et al., 2018). Kualitas hasil segmentasi diukur menggunakan tiga metrik validasi internal yang saling melengkapi. Pertama, *Silhouette Score* mengukur seberapa mirip suatu titik data dengan klusternya sendiri dibandingkan kluster terdekat lainnya; nilai mendekati +1 menunjukkan pemisahan kluster yang baik. Kedua, *Davies-Bouldin Index (DBI)* mengukur rasio rata-rata dispersi intra-kluster terhadap jarak antar pusat kluster; nilai DBI yang lebih kecil menandakan kluster yang lebih kompak dan terpisah. Ketiga, *Calinski-Harabasz Index (CHI)* mengukur rasio varians antar-kluster terhadap varians intra-kluster; nilai CHI yang lebih besar mengindikasikan kluster yang lebih padat dan terpisah satu sama lain.
- g) Interpretasi hasil. Tahap akhir mencakup analisis profil setiap kluster berdasarkan nilai rata-rata dan median ketiga dimensi RFM. Setiap kluster diberi label segmen bisnis berdasarkan peringkat gabungan skor R, F, dan M misalnya *Pelanggan Bernilai Tinggi*, *Pelanggan Aktif*, *Pelanggan Potensial*, hingga *Pelanggan Dorman*.

HASIL DAN PEMBAHASAN

Visualisasi Hasil Analisis

Bagian ini menyajikan hasil dan pembahasan segmentasi pelanggan Enterprise Business Kantor Pos Indonesia yang dilakukan melalui integrasi RFM Analysis (Recency, Frequency, Monetary) dengan algoritma Affinity Propagation (AP). Proses analisis diawali dengan tahap preprocessing data yang mencakup pembersihan data, transformasi variabel, serta penghapusan outlier untuk meningkatkan kualitas data sebelum dilakukan klusterisasi. Dari total 1.325 pelanggan

korporat hasil agregasi awal, sebanyak 1.032 pelanggan memenuhi kriteria analisis dan digunakan sebagai objek segmentasi.

Berdasarkan fitur RFM yang telah dikonstruksi dan dinormalisasi, algoritma Affinity Propagation membentuk sejumlah klaster secara otomatis tanpa memerlukan penentuan jumlah klaster di awal. Segmentasi yang dihasilkan bertujuan untuk mengidentifikasi kelompok pelanggan dengan karakteristik transaksi yang berbeda sehingga dapat memberikan pemahaman yang lebih mendalam mengenai pola aktivitas, intensitas transaksi, dan kontribusi pendapatan masing-masing segmen pelanggan. Hasil analisis selanjutnya disajikan secara bertahap, meliputi profil dan distribusi klaster yang terbentuk, visualisasi karakteristik setiap segmen berdasarkan dimensi RFM, interpretasi bisnis dari masing-masing kelompok pelanggan, serta evaluasi kualitas klaster menggunakan metrik Silhouette Score, Davies–Bouldin Index (DBI), dan Calinski–Harabasz Index (CHI).

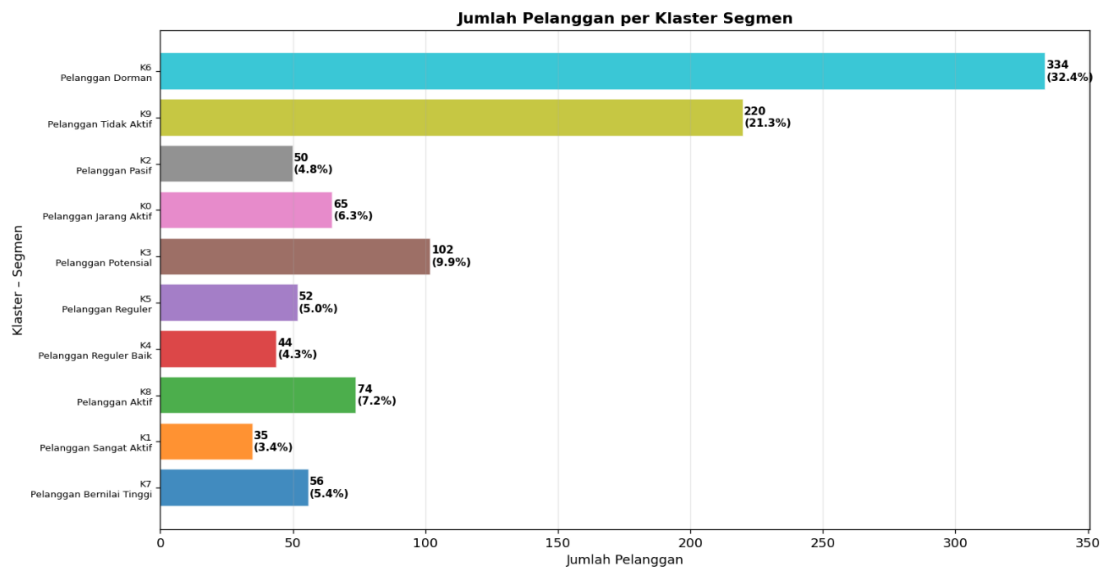
Tabel 2. Profil Klaster

Klaster	N_Pelanggan	R_mean	F_mean	M_mean	%_Pelanggan
0	65	14.02	226.02	1,955,765.14	6.30
1	35	26.74	741.40	5,508,842.14	3.39
2	50	6.58	62.24	3,646,185.08	4.84
3	102	28.04	113.97	1,995,141.18	9.88
4	44	41.50	173.89	2,947,664.68	4.26
5	52	19.94	142.90	4,952,361.27	5.04
6	334	4.18	20.98	401,798.59	32.36
7	56	36.68	422.62	7,525,038.14	5.43
8	74	32.27	323.78	3,965,586.49	7.17
9	220	15.37	49.95	866,478.41	21.32

Tabel tersebut menyajikan profil 10 klaster pelanggan yang terbentuk secara otomatis melalui algoritma Affinity Propagation berdasarkan karakteristik Recency (R), Frequency (F), dan Monetary (M). Klaster 6 merupakan segmen terbesar dengan 334 pelanggan (32,36%), diikuti Klaster 9 dengan 220 pelanggan (21,32%). Kedua klaster tersebut memiliki nilai rata-rata R, F, dan M yang relatif rendah dibandingkan klaster lainnya, menunjukkan dominasi pelanggan dengan tingkat aktivitas transaksi dan kontribusi pendapatan yang terbatas. Temuan ini mengindikasikan bahwa lebih dari separuh pelanggan Enterprise Business berada pada kategori keterlibatan rendah sehingga memerlukan perhatian khusus dalam strategi aktivasi dan peningkatan loyalitas.

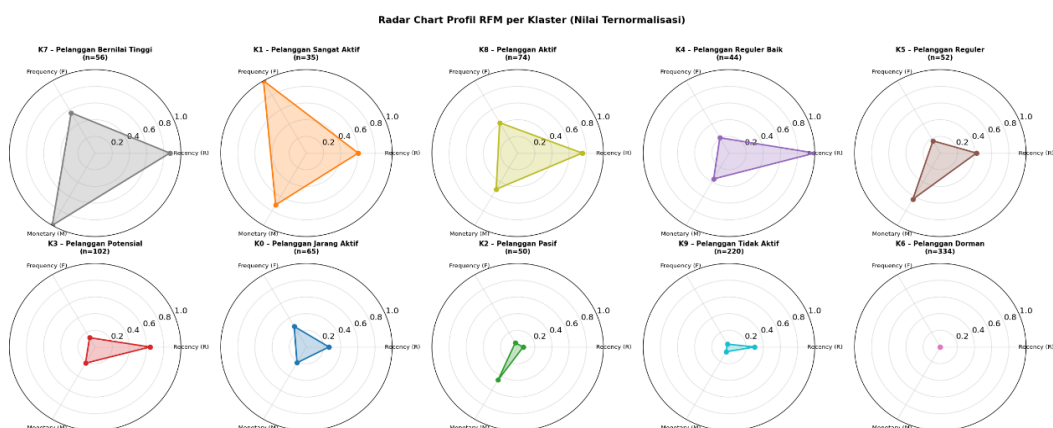
Sebaliknya, Klaster 7 dan Klaster 1 menunjukkan performa RFM yang paling menonjol. Klaster 7 memiliki nilai rata-rata Recency sebesar 36,68, Frequency sebesar 422,62, dan Monetary sebesar Rp7.525.038, sedangkan Klaster 1 mencatat nilai Frequency tertinggi sebesar 741,40 dengan Monetary sebesar Rp5.508.842. Meskipun kedua klaster tersebut hanya mencakup masing-masing 5,43% dan 3,39% dari total pelanggan, karakteristiknya menunjukkan kelompok pelanggan yang aktif dan memberikan kontribusi pendapatan yang signifikan bagi perusahaan. Kondisi ini

memperlihatkan adanya ketimpangan distribusi nilai pelanggan, di mana sebagian kecil pelanggan berkontribusi secara lebih besar dibandingkan mayoritas pelanggan lainnya.



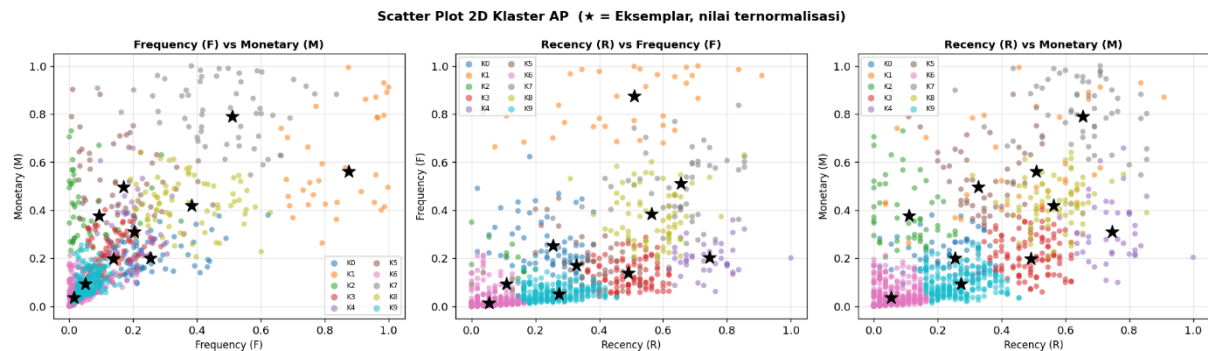
Gambar 2. Distribusi Jumlah Pelanggan per Klaster

Dari hasil analisis, Gambar 2 menunjukkan distribusi jumlah pelanggan pada setiap klaster hasil segmentasi. Klaster 6 (Pelanggan Dorman) merupakan segmen terbesar dengan 334 pelanggan (32,36%), diikuti Klaster 9 (Pelanggan Tidak Aktif) sebanyak 220 pelanggan (21,32%). Kedua segmen tersebut mencakup 53,68% dari total pelanggan, yang mengindikasikan bahwa sebagian besar pelanggan Enterprise Business memiliki tingkat aktivitas dan kontribusi yang relatif rendah. Kondisi ini menunjukkan perlunya strategi aktivasi dan reaktivasi untuk meningkatkan keterlibatan pelanggan. Sebaliknya, segmen bernilai tinggi yang terdiri atas Klaster 7 (Pelanggan Bernilai Tinggi), Klaster 1 (Pelanggan Sangat Aktif), dan Klaster 8 (Pelanggan Aktif) hanya mencakup sekitar 15,99% dari total pelanggan. Meskipun jumlahnya relatif kecil, segmen ini memiliki potensi kontribusi pendapatan yang lebih besar sehingga perlu menjadi fokus utama dalam strategi retensi pelanggan. Untuk memahami karakteristik setiap klaster berdasarkan ketiga dimensi RFM, Gambar 3 menyajikan profil RFM dalam bentuk radar chart.



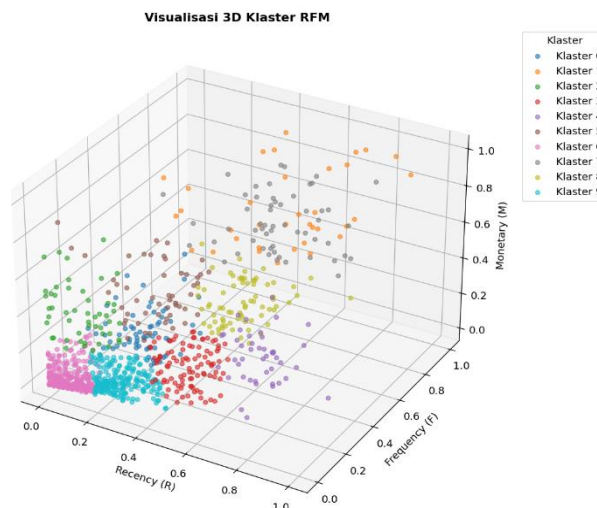
Gambar 3. Radar Chart Profil RFM per Klaster

Klaster K7 (Pelanggan Bernilai Tinggi) menampilkan bentuk radar yang relatif besar dan seimbang pada ketiga dimensi, mengindikasikan pelanggan yang aktif, sering bertransaksi, sekaligus berkontribusi besar terhadap pendapatan. Klaster K1 (Pelanggan Sangat Aktif) menunjukkan lonjakan ekstrem pada dimensi Frequency, mencirikan pelanggan yang sangat intensif dalam frekuensi transaksi harian meskipun nilai monetranya lebih rendah dari K7. Klaster K4 (Pelanggan Reguler Baik) menonjol pada dimensi Recency dengan nilai rata-rata tertinggi (41,50 hari aktif), menandakan konsistensi kehadiran transaksi yang paling tinggi. Sebaliknya, klaster K6 dan K9 pada radar bawah menampilkan bentuk yang sangat kecil di semua dimensi, mengonfirmasi keterlibatan yang sangat minim.



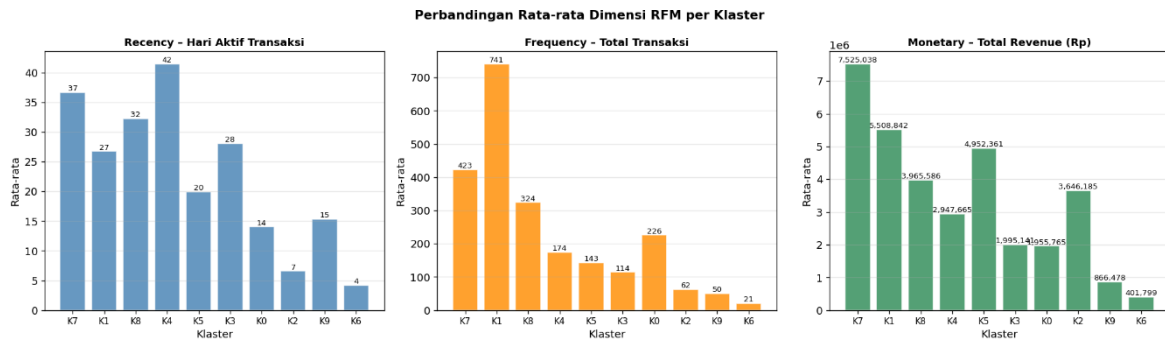
Gambar 4. Scatter Plot 2D: Tiga Pasang Dimensi RFM

Gambar 4 memetakan distribusi klaster pada tiga pasang dimensi RFM dalam ruang ternormalisasi secara dua dimensi. Tanda bintang menandai posisi eksemplar masing-masing klaster yang dipilih langsung dari data asli oleh algoritma AP. Pada pasangan Frequency vs Monetary (panel kiri), terlihat pemisahan yang cukup jelas antara klaster bernilai tinggi di kuadran kanan atas (K7 dan K1) dengan klaster dormant yang terkonsentrasi di kuadran kiri bawah (K6 dan K9). Pada pasangan Recency vs Frequency (panel tengah), K1 terlihat terpisah ke arah atas karena frekuensi transaksinya yang sangat tinggi, sementara K4 menonjol ke kanan karena Recency-nya yang dominan.



Gambar 5. Scatter Plot 3D

Visualisasi tiga dimensi (3D) memperlihatkan posisi seluruh 1.032 pelanggan sekaligus dalam ruang fitur R, F, dan M yang ternormalisasi. Visualisasi ini mengonfirmasi pola yang telah terlihat pada scatter plot 2D: sebagian besar pelanggan (K6 dan K9) terkonsentrasi di area berdensitas tinggi dekat titik asal dengan nilai R, F, dan M yang rendah, sementara kluster bernilai tinggi tersebar di area yang lebih jarang namun secara geometris terpisah dengan jelas. Posisi eksemplar masing-masing kluster tampak berada di titik sentral anggotanya, memvalidasi bahwa AP berhasil memilih representasi kluster yang tepat dari data asli.



Gambar 6. Bar Chart Perbandingan Rata-rata Dimensi RFM per Kluster

Pada dimensi Recency, K4 memiliki rata-rata hari aktif transaksi tertinggi (42 hari) yang mencerminkan konsistensi kehadiran harian paling tinggi, diikuti K7 (37 hari) dan K8 (32 hari). Pada dimensi Frequency, K1 jauh melampaui kluster lain dengan rata-rata 741,40 transaksi per periode hampir dua kali lipat K7 (422,62) menunjukkan pelanggan dengan intensitas operasional sangat tinggi. Pada dimensi Monetary, K7 mendominasi dengan rata-rata pendapatan Rp7,53 juta per pelanggan, diikuti K1 (Rp5,51 juta) dan K5 (Rp4,95 juta). Kluster K6 secara konsisten berada di posisi paling bawah pada ketiga dimensi, mengonfirmasi statusnya sebagai segmen dengan keterlibatan dan kontribusi finansial terendah.

Tabel 3. Kluster dan Segmen Setiap Pelanggan

ID Pelanggan	Nama Pelanggan Korporat	Kluster	Segmen
ADIRA07631A	PT ADIRA DINAMIKA MULTIFINANCE CAB MADIUN	9	Pelanggan Tidak Aktif
ADIRA07634A	ADIRA FINANCE PONOROGO	9	Pelanggan Tidak Aktif
AGROBBRBLPP05811A	BALAI BESAR RISET BUDIDAYA LAUT PERIKANAN	6	Pelanggan Dorman
...	...	...	...
UWM07600A	SLPKSBUniversitasWidyaMandala	4	Pelanggan Reguler Baik
WSUSA05805C	HENDY HARTADI SAMUDRA IDONG	2	Pelanggan Pasif
kesmahkota08861a	PT MAHKOTA ANUGRAH KARYA	6	Pelanggan Dorman

Tabel 3 memetakan seluruh 1.032 pelanggan Enterprise Business ke dalam 10 segmen bisnis berdasarkan peringkat skor RFM gabungan. Peringkat pertama ditempati Klaster 7 sebagai Pelanggan Bernilai Tinggi (56 pelanggan, 5,43%), dengan Klaster 1 sebagai Pelanggan Sangat Aktif (35 pelanggan, 3,39%) di posisi kedua. Di ujung lain, Klaster 6 sebagai Pelanggan Dorman (334 pelanggan, 32,36%) dan Klaster 9 sebagai Pelanggan Tidak Aktif (220 pelanggan, 21,32%) mendominasi secara kuantitatif. Pemberian label segmen yang berbasis data ini memungkinkan manajemen Kantor Pos Indonesia untuk mengidentifikasi secara tepat pelanggan mana yang memerlukan perhatian strategis berbeda. Kualitas klaster yang terbentuk dievaluasi menggunakan tiga metrik validasi internal sebagaimana disajikan pada Tabel 4.

Evaluasi Model

Tabel 4. Metrik Evaluasi

Metrik Evaluasi	Nilai
Silhouette Score	0.3766
Davies-Bouldin Index (DBI)	1.0486
Calinski-Harabasz Index (CHI)	738.73

Hasil evaluasi kualitas klaster menggunakan tiga metrik validasi internal menunjukkan Silhouette Score sebesar 0,3766 berada dalam kategori cukup baik (rentang 0,25 - 0,50), mengindikasikan bahwa klaster yang terbentuk memiliki pemisahan yang memadai meskipun terdapat beberapa tumpang tindih pada batas antar klaster. Davies-Bouldin Index (DBI) sebesar 1,0486 sedikit di atas ambang batas ideal ($<1,0$), namun masih dalam kisaran yang dapat diterima untuk klasterisasi berbasis data transaksional perusahaan. Calinski-Harabasz Index (CHI) sebesar 738,73 mengindikasikan klaster yang kompak secara intra dan terpisah secara antar-klaster. Secara keseluruhan, ketiga metrik menunjukkan bahwa konfigurasi AP dengan preference pada persentil ke-2 dan damping 0,7 menghasilkan segmentasi yang valid dan dapat diandalkan.

Hasil segmentasi ini memberikan landasan berbasis data yang konkret bagi Kantor Pos Indonesia dalam merancang strategi layanan yang terdiferensiasi. Segmen Pelanggan Bernilai Tinggi (K7) dan Pelanggan Sangat Aktif (K1) yang secara kolektif hanya mencakup 8,82% dari total pelanggan namun menghasilkan revenue rata-rata tertinggi, perlu diprioritaskan dalam program retensi eksklusif, penugasan account manager khusus, dan pengembangan layanan premium. Segmen menengah seperti Pelanggan Aktif (K8), Pelanggan Reguler Baik (K4), dan Pelanggan Reguler (K5) yang mencakup 16,47% dari total pelanggan, berpotensi ditingkatkan ke tier lebih tinggi melalui program loyalitas dan penawaran layanan nilai tambah. Sementara itu, segmen Pelanggan Dorman (K6) dan Pelanggan Tidak Aktif (K9) yang mendominasi secara numerik memerlukan kampanye reaktivasi berbasis insentif dengan biaya

SIMPULAN

Penelitian ini berhasil menerapkan RFM Analysis yang diintegrasikan dengan algoritma Affinity Propagation untuk menghasilkan segmentasi pelanggan Enterprise Business Kantor Pos Indonesia yang bermakna secara bisnis. Dari 1.032 pelanggan yang dianalisis, algoritma AP secara adaptif membentuk 10 klaster yang mencerminkan spektrum perilaku transaksi dari yang bernilai tinggi hingga dorman, tanpa memerlukan penetapan jumlah klaster di awal. Konfigurasi terpilih menghasilkan Silhouette Score = 0,3766, DBI = 1,0486, dan CHI = 738,73, yang secara keseluruhan mengindikasikan kualitas segmentasi yang cukup baik untuk data transaksional bisnis nyata. Dua klaster terbesar Pelanggan Dorman (K6, 32,36%) dan Pelanggan Tidak Aktif (K9, 21,32%)

menunjukkan bahwa lebih dari separuh pelanggan korporat berada pada segmen dengan keterlibatan rendah, sehingga strategi aktivasi menjadi prioritas utama. Sebaliknya, segmen Pelanggan Bernilai Tinggi (K7) dan Pelanggan Sangat Aktif (K1) yang hanya mencakup 8,82% dari total pelanggan menghasilkan rata-rata revenue tertinggi, menegaskan pentingnya program retensi yang terfokus dan terpersonalisasi. Penelitian selanjutnya dapat mengeksplorasi penambahan variabel temporal, perluasan periode data, serta perbandingan langsung antara Affinity Propagation dan K-Affinity Propagation pada dataset yang sama.

DAFTAR PUSTAKA

- Asriny, N. I. (2019). *Pengelompokkan Pekerja Paruh Waktu Dengan K-Affinity Propagation (K-Ap) Clustering Algorithm*.
- Cheng, C. H., & Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert Systems with Applications*, 36(3 PART 1), 4176–4184. <https://doi.org/10.1016/j.eswa.2008.04.003>
- Doğan, O., Ayçin, E., & Bulut, Z. A. (2018). Customer Segmentation by Using RFM Model and Clustering Methods: A Case Study in Retail Industry. *International Journal of Contemporary Economics and Administrative Sciences*, 8(1), 1–19. www.ijceas.com
- Dueck, D. (2009). *Affinity propagation: clustering data by passing messages*.
- Hosseini, S. M. S., Maleki, A., & Gholamian, M. R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications*, 37(7), 5259–5264. <https://doi.org/10.1016/j.eswa.2009.12.070>
- Jannah, S. L., & Sadiqin, A. (2021). Analisis Kualitas Pelayanan Terhadap Tingkat Kepuasan Pelanggan Pada PT Pos Indonesia Cabang Wonoayu Sidoarjo. *Jurnal Ekonomi, Manajemen, Bisnis Dan Sosial (EMBISS)*, 2(November), 37–48.
- Maimon, O., & Rokach, L. (2021). Data Mining and Knowledge Discovery. In *Urban Book Series*. https://doi.org/10.1007/978-981-15-8983-6_42
- Marlina, D., Lina, N., Fernando, A., & Ramadhan, A. (2018). Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi*, 4(2), 64. <https://doi.org/10.24014/coreit.v4i2.4498>
- Mirantika, N., Syamfithriani, T. S., & Trisudarmo, R. (2023). Implementasi Algoritma K-Medoids Clustering Untuk Menentukan Segmentasi Pelanggan. *Jurnal Nuansa Informatika*, 17(1), 196–204.
- Muhajir, M., & Sari, N. N. (2018). K-Affinity Propagation (K-AP) and K-Means Clustering for Classification of Earthquakes in Indonesia. *Proceeding - 2018 International Symposium on Advanced Intelligent Informatics: Revolutionize Intelligent Informatics Spectrum for Humanity, SAIN 2018*, 6–10. <https://doi.org/10.1109/SAIN.2018.8673344>
- Mustafa, A. A., & Jacksi, K. (2023). Affinity Propagation and K-Means Algorithm for Document Clustering Based on Semantic Similarity. *Science Journal of University of Zakho*, 11(2), 153–159. <https://doi.org/10.25271/sjuoz.2023.11.2.1148>
- Mustika, A., Mindari, E., Kartawinata, B. R., Rachmawati, M., Mahriani, E., Nugraheni, K. S., Khaerana, K., The, H. Y., Wulansari, A. S., Syahputra, S., Ambarwati, R., Wicaksono, G., Suyatno, A., & Solihat, W. (2020). Pengantar Bisnis. In *Angewandte Chemie International Edition*, 6(11), 951–952. (Vol. 3, Number 1). <https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf>
- Parikh, Y., & Abdelfattah, E. (2020). Clustering algorithms and RFM analysis performed on retail

transactions. *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 506–511.

Septyawati, N. (2024). *ANALISIS PENGEMBANGAN KARIR PENGARUHNYA TERHADAP KEPUASAN KERJA KARYAWAN PADA KANTOR PUSAT PT. POS INDONESIA (PERSERO) BANDUNG*. 2, 306–312.

Tzortzis, G., Likas, A., & Vlassis, N. (2003). The global kernel k-means clustering algorithm. *Proceedings of the International Joint Conference on Neural Networks*. <https://doi.org/10.1109/IJCNN.2008.4634069>