

**ANALISIS SEGMENTASI PELANGGAN TOKO ONLINE MARKETPLACE
BERDASARKAN RECENCY, FREQUENCY, MONETARY, TIME, SATISFACTION
(RFMTS) MENGGUNAKAN ALGORITMA K-MEDOIDS CLUSTERING**

Putri Angelina Windjaya¹, Bakti Siregar²

Universitas Matana Tangerang, Indonesia

Email: putriangelinaw@gmail.com, siregar.bakti@matanauniversity.ac.id

KATA KUNCI

Segmentasi
Pelanggan,
Clustering, K-
Medoids,
RFMT, RFMTS

ABSTRACT

The marketplace is an e-commerce website that applies traditional market concepts and implements them online. Due to the rapid development of the times, a business will seek various ways to maintain its business so that it continues to grow and generate a high income. One way is to build and maintain long-term relationships with customers. Therefore, it is necessary to implement a marketing strategy based on customer relationship management, all of which aims to increase turnover and profit while retaining customers. From these problems, this study applies customer segmentation to detect diversity among customers, so those segments represent potential customers to improve marketing strategies. This segmentation needs to consider the value of the customer's recency, frequency, monetary, time, and satisfaction (RFMTS) variables. Recency is the customer's length of time since the last payment. Frequency is how often customers make transactions. Monetary is the number of transactions made by the customer. Time is the time interval between two consecutive purchases by a customer. Satisfaction is the level of customer satisfaction based on the total rating and number of reviews. This study uses the K-Medoids Clustering algorithm to perform segmentation according to customer characteristics. This study uses primary data from a database of one of the online marketplace stores in Indonesia.

ABSTRAK

Marketplace merupakan website e-commerce yang menerapkan konsep pasar tradisional dan dilaksanakan secara online. Karena perkembangan zaman yang sangat pesat, suatu bisnis perlu mengaplikasikan berbagai cara untuk mempertahankan bisnisnya supaya terus berkembang dan menghasilkan pendapatan yang besar. Salah satu caranya yaitu dengan membangun dan mempertahankan hubungan jangka panjang dengan pelanggan. Oleh karena itu, sangatlah penting untuk menerapkan strategi pemasaran berdasarkan manajemen hubungan pelanggan (customer relationship management), yang bertujuan untuk meningkatkan omset dan profit seiring mempertahankan pelanggan. Dari permasalahan tersebut, penelitian ini menerapkan segmentasi pelanggan (customer segmentation) untuk mendeteksi perbedaan atau persamaan di antara pelanggan, sehingga mendapatkan segmen atau cluster mewakili pelanggan yang

berpotensi untuk meningkatkan strategi pemasaran. Segmentasi ini perlu mempertimbangkan nilai dari variabel recency, frequency, monetary, time, dan satisfaction (RFMTS) pelanggan. Recency merupakan lama waktu pelanggan sejak pembayaran terakhir. Frequency merupakan seberapa sering pelanggan melakukan transaksi. Monetary merupakan jumlah dari transaksi yang dilakukan pelanggan tersebut. Time merupakan interval waktu antara dua pembelian berturut-turut oleh pelanggan. Satisfaction merupakan tingkat kepuasan pelanggan berdasarkan total rating dan jumlah ulasan. Penelitian ini menggunakan algoritma K-Medoids Clustering untuk melakukan segmentasi sesuai karakteristik pelanggan. Penelitian ini menggunakan data primer dari database salah satu toko online marketplace di Indonesia.

PENDAHULUAN

Seiring perkembangan zaman, persaingan antar dunia usaha semakin intens dan sangat kompetitif. Hal ini sejalan dengan peningkatan teknologi dan internet yang memfasilitasi segala aktivitas bisnis. Internet telah menjadi sarana penunjang untuk berbelanja pada suatu *marketplace online* yang disebut dengan *e-commerce* (Prasetyo, Mustafid, & Hakim, 2020). Di Indonesia, *e-commerce* terus berkembang karena perilaku konsumtif dan gaya hidup di kalangan masyarakat menengah. Keberadaan *e-commerce* di Indonesia telah ada sejak tahun 1994 hingga saat ini. Berdasarkan institusi Otoritas Jasa Keuangan (OJK), dikatakan bahwa hasil survei pada tahun 2021 mengukuhkan Indonesia sebagai negara tertinggi di dunia yang menggunakan layanan *e-commerce* sebanyak 88,1% pengguna internet di Indonesia berbelanja secara *online* (Lestari & Agnia, 2023).

Jika perkembangan zaman sangat pesat, maka suatu bisnis perlu mengaplikasikan berbagai cara untuk mempertahankan bisnis tersebut supaya terus berkembang dan menghasilkan pendapatan yang besar. Salah satu cara untuk mempertahankannya yaitu dengan membangun hubungan pelanggan jangka panjang. Oleh sebab itu, sangat penting untuk menerapkan strategi pemasaran berdasarkan manajemen hubungan pelanggan (*customer relationship management*), yang bertujuan untuk meningkatkan *omset* dan *profit* seiring mempertahankan pelanggan (Mensouri, Azmani, & Azmani, 2022). Kunci dari bisnis *e-commerce* berkelanjutan adalah dengan memahami pelanggan untuk strategi pemasaran yang dipersonalisasi sesuai dengan kelompoknya (Zhou, Wei, & Xu, 2021). Dari permasalahan tersebut, penelitian ini menerapkan segmentasi pelanggan (*customer segmentation*) untuk mendeteksi keragaman di antara pelanggan, sehingga menghasilkan segmen atau *cluster* yang mewakili pelanggan berpotensi untuk meningkatkan strategi pemasaran.

Pada segmentasi data, pelanggan dikelompokkan menjadi beberapa kelompok, sehingga objek di setiap *cluster* memiliki kesamaan yang tinggi. Kesamaan tersebut dilihat dari nilai atribut yang relevan seperti jenis kelamin, usia, gaya hidup, penghasilan, pekerjaan, ataupun pembeliannya (Han, Pei, & Tong, 2022). Segmentasi pelanggan yang baik secara efektif mengalokasikan sumber daya pemasaran, memungkinkan bisnis menargetkan kelompok pelanggan tertentu, dan juga membantu membangun hubungan jangka panjang dengan pelanggan (Anitha & Patil, 2022). Segmentasi data juga disebut *clustering* di beberapa

Analisis Segmentasi Pelanggan Toko Online Marketplace Berdasarkan Recency, Frequency, Monetary, Time, Satisfaction (Rfmts) Menggunakan Algoritma K-Medoids Clustering.

aplikasi karena membagi kumpulan data besar menjadi kelompok-kelompok berdasarkan kesamaan (Han et al., 2022).

Pada penelitian ini, metode *clustering* yang digunakan adalah Algoritma K-Medoids berdasarkan nilai *recency*, *frequency*, dan *monetary* (RFM) yang sudah dikembangkan. Model RFM yang sudah dikembangkan tersebut bernama RFMTS dengan tambahan variabel T (*Time*) dan S (*Satisfaction*). Analisis menggunakan RFMTS bertujuan untuk meningkatkan hubungan pelanggan dan mengembangkan strategi pemasaran yang lebih efektif (Mensouri et al., 2022). Selain itu, mempertimbangkan kepuasan pelanggan membantu mengembangkan strategi pemasaran yang efektif dan memperkuat loyalitas pelanggan. Kepuasan pelanggan mencerminkan perbedaan antara kualitas layanan yang diharapkan dan yang dirasakan (Padma & Ahn, 2020).

Tujuan penelitian ini untuk menganalisis data penjualan dan mengetahui hasil segmentasi pelanggan menggunakan K-Medoids *clustering* berdasarkan nilai RFMT untuk menghasilkan segmen pelanggan yang berpotensi meningkatkan penjualan. Manfaat penelitian ini dapat memahami *Cluster Analysis* dengan *Algoritma K-Medoids clustering* berdasarkan nilai RFMTS. Dapat merekomendasikan kepada toko online marketplace untuk mengimplementasikan strategi pemasaran berdasarkan hasil segmentasi pelanggan yang diperoleh.

METODE PENELITIAN

Penelitian ini menggunakan data primer atau data penjualan dari salah satu toko online marketplace di Indonesia. Data tersebut didapatkan dari database toko yang memuat informasi pesanan pelanggan secara detail. Selain itu, untuk data ulasan didapatkan dengan cara web scraping pada laman toko. Data ulasan menyimpan informasi kepuasan pelanggan pada produk yang dibeli olehnya. Objek penelitian ini adalah pelanggan dan produk yang dibeli pelanggan pada toko online marketplace tersebut.

Pada pengolahan data, tipe data yang digunakan dalam penelitian adalah data kuantitatif atau data yang bersifat numerik/angka. Pengolahan data terdiri dari beberapa proses, yang pertama adalah pemilihan data. Pada pemilihan data yang dilakukan adalah menyeleksi data/variabel dari *database* yang sesuai dan akan digunakan untuk penelitian selanjutnya. Proses kedua adalah pembersihan data (Lucyana, Sari, Kurniawan, & Buchari, 2022). Di tahap ini data harus dipastikan sudah bersih, tidak ada duplikat, tidak ada yang salah, tidak ada data yang kosong (*NA*), ataupun inkonsisten. Proses ketiga adalah mentransformasi data supaya bentuk informasi data sesuai dengan model yang diinginkan. Setelah mentransformasi data, yang dilakukan selanjutnya adalah melakukan penanganan data *outlier*, dan pengecekan korelasi data. Jika data terdapat multikolinieritas perlu diperbaiki dengan menggunakan *Principal Component Analysis* (PCA). PCA juga berperan untuk melakukan penskalaan pada data agar data tidak terlampaui tinggi maupun terlampaui rendah

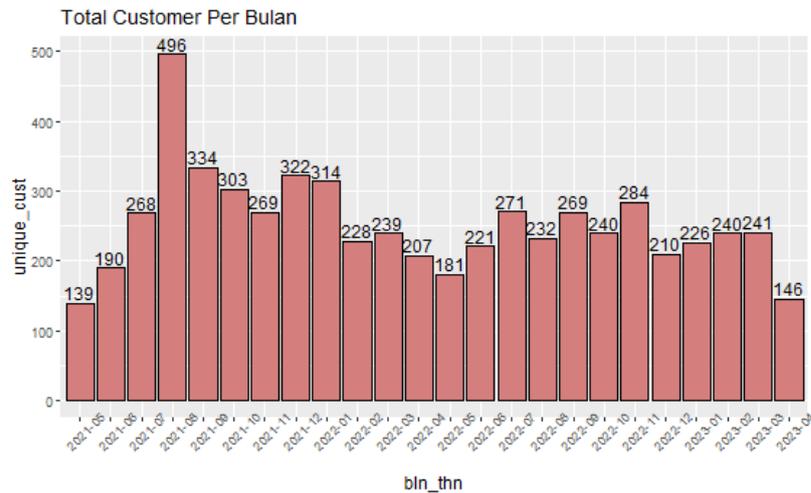
HASIL DAN PEMBAHASAN

Analisis Data Eksploratif

Setelah melakukan pembersihan data, tahap selanjutnya adalah analisis data eksploratif. Eksplorasi data dilakukan untuk memahami dan meringkas karakteristik utama data, sehingga dapat membantu dalam menentukan teknik atau langkah yang dilakukan selanjutnya (Rukajat,

2018). Variabel yang dieksplorasi adalah variabel Nama_Pembeli, Nomor_Invoice, dan Tanggal_Pembayaran. Dari variabel Nama_Pembeli dapat ditemukan total pelanggan untuk setiap bulan dan melihat pertumbuhan pelanggan baru per bulannya. Dari variabel Nomor_Invoice dapat ditemukan total pesanan per bulannya.

Total Pelanggan Per Bulan

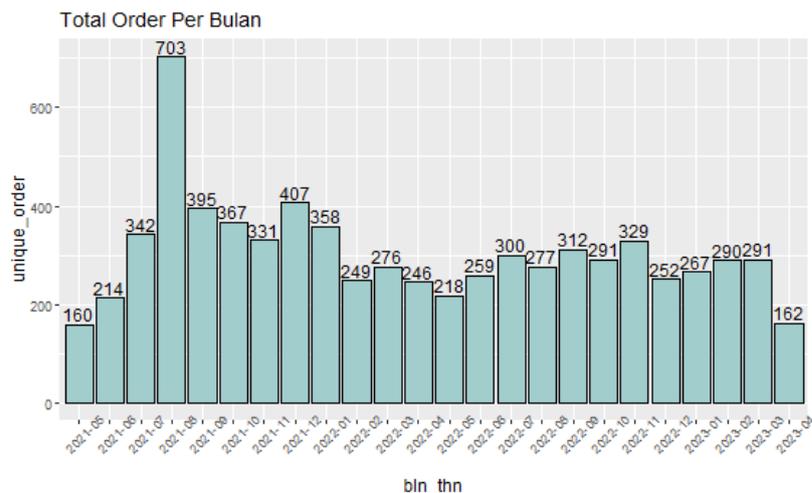


Gambar 1

Diagram Total Pelanggan Per Bulan

Gambar 1. menunjukkan grafik total pelanggan per bulan. Total pelanggan memiliki kenaikan per bulan sejak Mei 2021 hingga Agustus 2021. Lalu, tren tersebut memiliki penurunan sampai Mei 2022, walaupun ada kenaikan sedikit pada Desember 2021 dan Januari 2022. Setelah itu, total pelanggan per bulannya terus berada di antara 200-300 pelanggan, dan terjadi penurunan pada April 2023.

Total Pesanan Per Bulan



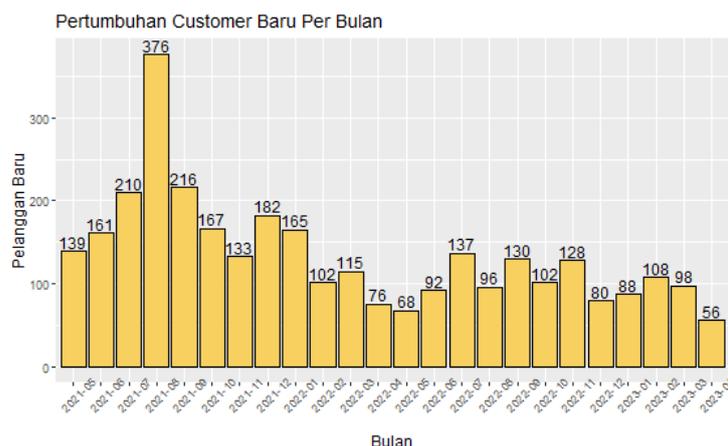
Gambar 2

Diagram Total Pesanan Per Bulan

Analisis Segmentasi Pelanggan Toko Online Marketplace Berdasarkan Recency, Frequency, Monetary, Time, Satisfaction (Rfmts) Menggunakan Algoritma K-Medoids Clustering.

Gambar 2. menunjukkan grafik total pesanan per bulan. Total pesanan per bulannya mengalami kenaikan mulai dari Mei 2021 hingga Agustus 2021. Setelah itu, tren mengalami penurunan hingga Mei 2022, walaupun terdapat sedikit kenaikan pada Desember 2021 dan Januari 2022. Lalu, total pesanan per bulannya terus berada di antara 200-330 pesanan, dan terjadi penurunan pada April 2023. Secara keseluruhan, tren pada Gambar 2. memiliki kemiripan dengan tren pada Gambar 1. Oleh sebab itu, hal ini menunjukkan bahwa banyaknya pelanggan per bulan seiring dengan banyaknya pesanan per bulan.

Pertumbuhan Pelanggan Baru Per Bulan



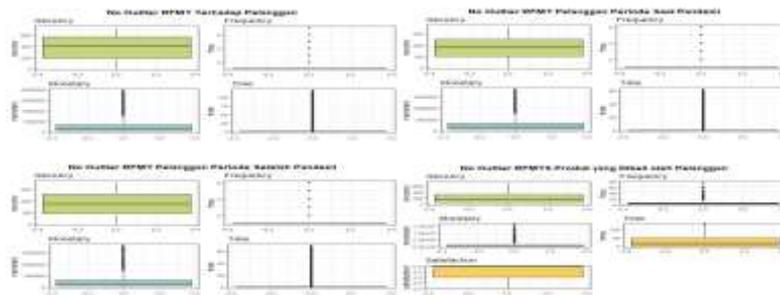
Gambar 3

Diagram Pertumbuhan Pelanggan Baru Per Bulan

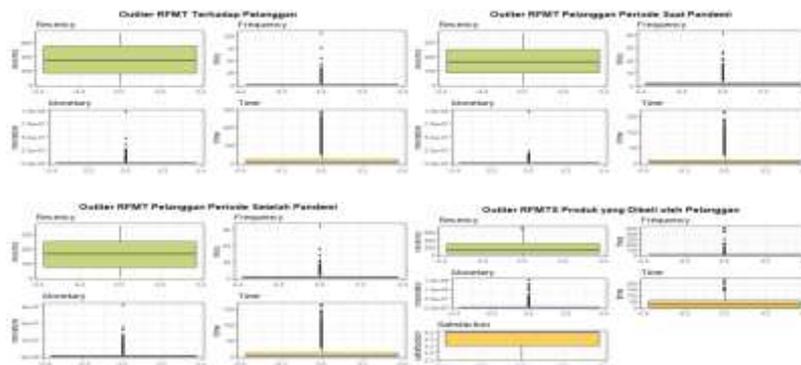
Gambar 3. menunjukkan grafik pertumbuhan pelanggan baru per bulan. Pertumbuhan pelanggan baru per bulannya mengalami kenaikan mulai dari Mei 2021 hingga Agustus 2021. Lalu grafik mengalami penurunan Mei 2022. Setelah itu, pertumbuhan pelanggan baru terus berada di antara 90-150 pelanggan baru, dan mengalami penurunan pada April 2023. Dari analisis data eksploratif, penelitian ini membagi data penjualan menjadi dua periode, yaitu periode Mei 2021 - April 2022 sebagai periode saat pandemi dan periode Mei 2022 – April 2023 sebagai periode setelah pandemi.

Analisis Data *Outlier*

Setelah melakukan analisis data eksploratif, maka langkah selanjutnya adalah analisis data *outlier*. Dalam analisis, *outlier* biasanya mendapatkan perlakuan khusus dengan cara memisahkannya dari data supaya tidak menyebabkan masalah dalam penentuan hasil akhir analisis (Sirait, 2022). Tidak semua *outlier* dapat dihilangkan karena terkadang *outlier* memiliki nilai penting dan berpengaruh dalam kumpulan objek data (Fitrayana & Saputro, 2022). Penanganan data *outlier* dilakukan pada data yang telah ditransformasi ke dalam bentuk.



Gambar 4
Boxplot Data Outlier setiap Data Transformasi



Gambar 5
Boxplot Data No Outlier Setiap Data Transformasi

RFMT dan RFMTS dan ditampilkan dengan metode *boxplot*.

Gambar 4. (a) menunjukkan *boxplot* untuk setiap variabel pada data transformasi RFMT pelanggan. Gambar 4. (b) menunjukkan *boxplot* untuk setiap variabel pada data transformasi RFMT pelanggan saat pandemi. Gambar 4. (c) menunjukkan *boxplot* untuk setiap variabel pada data transformasi RFMT pelanggan setelah pandemi. Gambar 4. (d) menunjukkan *boxplot* untuk setiap variabel pada data transformasi RFMTS produk yang dibeli pelanggan. Berdasarkan Gambar 4.4, dapat diketahui bahwa *outlier* yang digambarkan sebagai titik-titik hitam pada *boxplot* terlihat sangat banyak. Setelah dicari tahu, *outlier* tersebut merupakan *outlier* penting. Dikatakan *outlier* penting, karena bisa saja *outlier* tersebut merupakan pelanggan dengan frekuensi tinggi, artinya pelanggan yang sering berbelanja, pelanggan dengan *monetary* tinggi, artinya pelanggan yang total pembeliannya tinggi, ataupun pelanggan yang sudah lama berbelanja pada toko atau langganan tetap. Pelanggan inilah yang mempengaruhi kenaikan penjualan. Oleh sebab itu, *outlier* tersebut tidak bisa sembarang dihilangkan. Selain itu, karena data *outlier* terlalu banyak dan tidak bisa dihapus, penanganan akan dilakukan dengan cara mengambil 5% dari setiap data transformasi yang diidentifikasi sebagai data *outlier* dan dijadikan sebagai segmen khusus (Varlamis, Tserpes, Etemad, Júnior, & Matwin, 2019). Hal ini dilakukan supaya hasil *cluster* atau segmentasi lebih optimal.

Setelah memisahkan 5% dari data yang diidentifikasi sebagai *outlier*, tampilan *boxplot* nya pun ada perubahan seperti pada Gambar 5. yaitu *outlier* yang berada sangat jauh dari kebanyakan data sudah dipisahkan. Walaupun masih terdapat *outlier*, khususnya pada

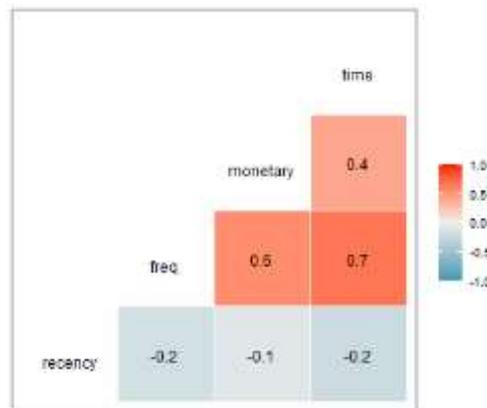
Analisis Segmentasi Pelanggan Toko Online Marketplace Berdasarkan Recency, Frequency, Monetary, Time, Satisfaction (Rfmts) Menggunakan Algoritma K-Medoids Clustering.

variabel *frequency* (F), *monetary* (M), dan *time* (T), tetapi langkah ini sudah mengatasi banyaknya *outlier* pada data.

Analisis Korelasi Data

Setelah melakukan analisis data outlier, selanjutnya dilakukan analisis korelasi data. Korelasi data merupakan kondisi antara dua variabel yang memiliki hubungan linier. Analisis korelasi data dapat mengetahui ada atau tidaknya hubungan linier antara variabel. Analisis korelasi data dilakukan ke data-data yang telah ditransformasi sebelumnya.

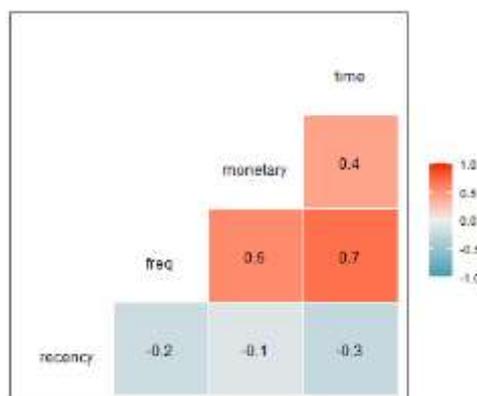
Korelasi Data RFMT Terhadap Pelanggan



Gambar 6
Matriks Korelasi RFMT Terhadap Pelanggan

Gambar 6. menunjukkan matriks korelasi setiap variabel pada data RFMT pelanggan. Korelasi yang tinggi diindikasikan dengan warna matriks yang memerah, sedangkan korelasi yang rendah diindikasikan dengan warna matriks yang biru. Terlihat bahwa terdapat korelasi yang tinggi pada variabel F, M, dan T. Hal ini menandakan ketika frekuensi pelanggan meningkat, *monetary* meningkat, umur pelanggan berlangganan juga meningkat.

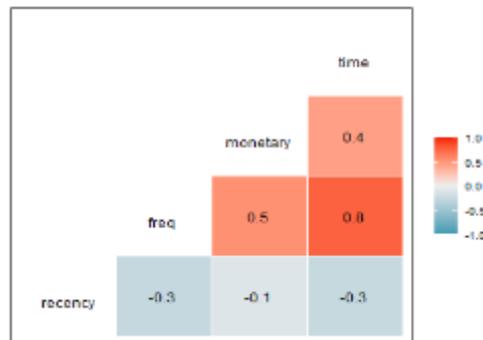
Korelasi Data RFMT Pelanggan Periode Saat Pandemi



Gambar 7
Matriks Korelasi RFMT Pelanggan Periode Saat Pandemi

Gambar 7. menunjukkan terdapat korelasi yang tinggi pada variabel F, M, dan T. Hal ini tentu sangat identik dengan Gambar 6. karena periode saat pandemi merupakan sebagian dari RFMT terhadap pelanggan di Gambar 6.

Korelasi Data RFMT Pelanggan Periode Setelah Pandemi



Gambar 8

Matriks Korelasi RFMT Pelanggan Periode Setelah Pandemi

Gambar 8. menunjukkan bahwa terdapat korelasi yang tinggi pada variabel F, M, dan T. Matriks korelasi setelah pandemi juga sangat identik dengan Gambar 6. dan Gambar 7.

Korelasi Data RFMTS Produk yang Dibeli Pelanggan



Gambar 9

Matriks Korelasi RFMTS Produk yang dibeli Pelanggan

Gambar 9. menunjukkan bahwa terdapat korelasi yang tinggi antara variabel F dan M. Selain variabel tersebut tidak ada korelasi yang signifikan. Berdasarkan analisis korelasi yang telah dilakukan di atas, dapat disimpulkan bahwa variabel yang memiliki hubungan linier dan sangat signifikan yaitu variabel F, M, dan T. Korelasi yang sangat tinggi ini perlu diatasi agar hasil dari segmentasi lebih baik dan optimal. Penanganan korelasi data akan menggunakan *Principal Component Analysis (PCA)*.

Analisis Segmentasi Pelanggan Toko Online Marketplace Berdasarkan Recency, Frequency, Monetary, Time, Satisfaction (Rfmts) Menggunakan Algoritma K-Medoids Clustering.

Principal Component Analysis (PCA)

Setelah melakukan analisis korelasi data, dapat diketahui bahwa setiap data transformasi memiliki korelasi data yang tinggi (Ningsih & Dukalang, 2019). Oleh sebab itu, langkah selanjutnya adalah penanganan korelasi data menggunakan PCA. Seperti yang telah diketahui bahwa data memiliki beberapa variabel dengan korelasi tinggi atau juga disebut multikolinieritas. Oleh sebab itu, untuk mereduksi dimensi data dan mengatasi multikolinieritas pada data akan dilakukan PCA. Seperti yang dilakukan sebelumnya, PCA diterapkan ke setiap data yang telah ditransformasi.

Tabel 1
Pratinjau PCA Setiap Data Transformasi

(a) PCA RFMT Pelanggan				(b) PCA RFMT Pelanggan Saat Pandemi			
PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
-0,8504	-0,9996	0,05078	0,08664	-0,77903	0,65677	-0,03031	-0,07390
-0,6946	0,47544	-0,0066	-0,0414	0,76823	-0,27412	0,23159	0,15685
-0,585	0,58942	-0,1521	-0,0086	-0,74154	0,09897	0,00222	-0,01184
0,19617	-0,6521	0,22313	-0,297	2,25291	0,66352	0,37237	-1,22678
0,81957	1,62981	0,29989	-0,1528	-0,46373	0,99660	-0,54177	-0,23976

(c) PCA RFMT Pelanggan Setelah Pandemi				(d) PCA RFMTS Produk yang dibeli Pelanggan				
PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4	PC5
-1,11771	-1,06511	0,42839	-0,03653	-1,55593	-1,04534	1,09248	-0,47298	-0,51152
-0,95083	-0,79136	0,22192	-0,01368	-0,53520	-0,06714	0,67468	0,93874	0,31783
1,30518	1,40322	0,62921	0,28723	0,33737	0,63981	0,48170	0,32766	0,21767
-1,01034	-1,04313	0,26785	-0,01404	-1,09470	-0,49721	0,97202	-0,74176	-0,60165
1,96110	0,13531	0,00775	0,73087	-0,64338	0,05326	-1,05393	1,13312	0,55949

Tabel 1. (a) merupakan tabel hasil PCA untuk data RFMT pelanggan. Tabel 1. (b) merupakan tabel hasil PCA untuk data RFMT pelanggan pada periode saat pandemi. Tabel 1. (c) merupakan tabel hasil PCA untuk data RFMT pelanggan pada periode setelah pandemi. Tabel 1. (d) merupakan tabel hasil PCA untuk data RFMTS produk yang dibeli pelanggan. Berdasarkan Tabel 1., data telah diterapkan PCA sehingga data tersebut dapat digunakan untuk tahap analisis berikutnya.

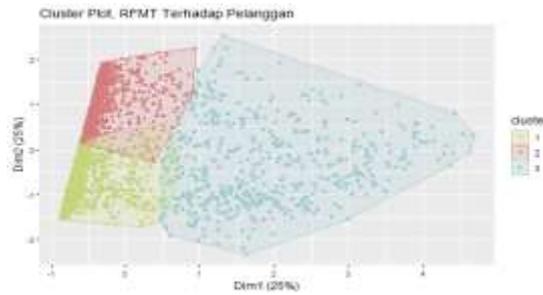
Penentuan *k Cluster* Optimal

Pada tahap ini dilakukan penentuan terhadap berapa banyak kluster yang dipilih supaya hasil segmentasi lebih optimal. Penentuan *cluster* menggunakan 3 metode yaitu metode *Elbow*, metode *Silhouette*, dan metode *Gap Statistic* (Izzadin, 2020).

K-Medoids Clustering

Setelah mengetahui banyaknya *k cluster* yang dipilih, maka langkah selanjutnya adalah melakukan segmentasi menggunakan metode K-Medoids. Metode K-Medoids diterapkan ke setiap data yang telah mendapatkan *k cluster* nya (Supriyadi, Triayudi, & Sholihati, 2021).

a. Clustering Data RFMT Pelanggan

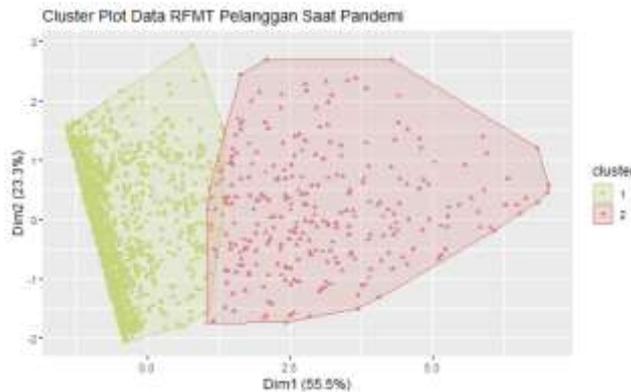


Gambar 10
Cluster Plot Data RFMT Terhadap Pelanggan

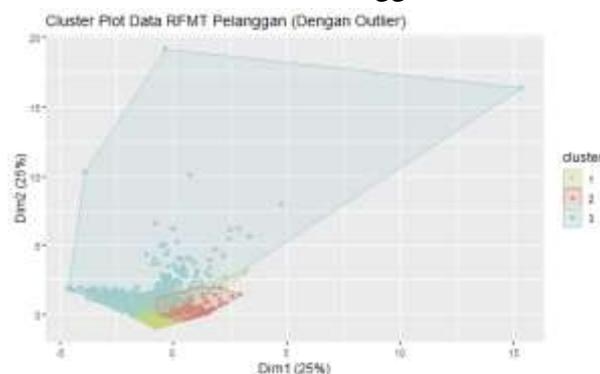
Gambar 10. menunjukkan hasil segmentasi untuk data RFMT terhadap pelanggan. Berdasarkan gambar tersebut, dapat diketahui bahwa segmen pada data cukup berdekatan atau bersebelahan dengan jumlah *cluster* sebanyak 3. Akan tetapi, seperti yang telah dijelaskan sebelumnya, jika penanganan *outlier* tidak dilakukan, maka hasil segmentasi akan terlihat seperti Gambar 11.

Gambar 11. menunjukkan hasil *clustering* yang kurang optimal. Hal ini terjadi karena terdapat data *outlier* yang membagi segmen menjadi kacau dan hasilnya tidak memuaskan. Oleh sebab itu, langkah penanganan *outlier* merupakan langkah penting.

b. Clustering Data RFMT Pelanggan Periode Saat Pandemi



Gambar 11
Cluster Plot Data RFMT Pelanggan Saat Pandemi

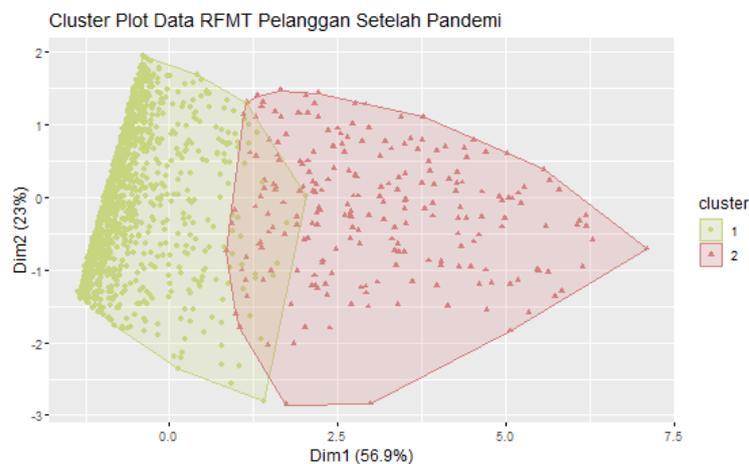


Gambar 12
Cluster Plot Data RFMT Pelanggan (Dengan Outlier)

Analisis Segmentasi Pelanggan Toko Online Marketplace Berdasarkan Recency, Frequency, Monetary, Time, Satisfaction (Rfmts) Menggunakan Algoritma K-Medoids Clustering.

Gambar 12. merupakan visualisasi dari hasil segmentasi data RFMT pelanggan untuk periode saat pandemi dengan jumlah *cluster* sebanyak 2.

c. Clustering Data RFMT Pelanggan Periode Setelah Pandemi

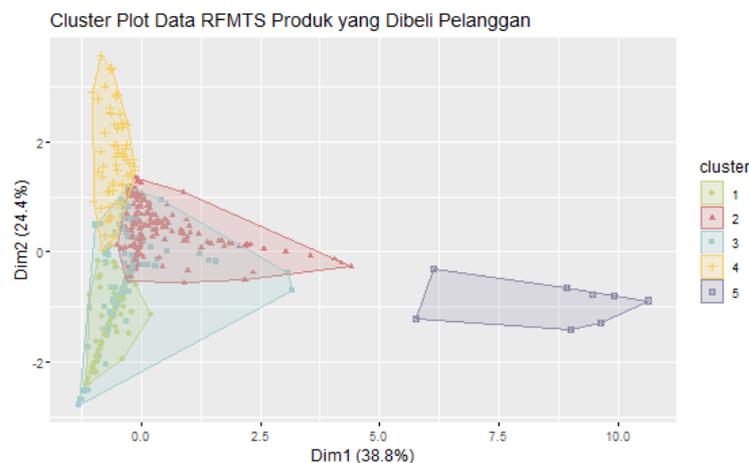


Gambar 13

Cluster Plot Data RFMT Pelanggan Setelah Pandemi

Gambar 13. merupakan hasil segmentasi data RFMT pelanggan untuk periode setelah pandemi dengan jumlah *cluster* sebanyak 2 juga. Jika diperhatikan Gambar 12. dan Gambar 13. seperti potongan dari Gambar 12., karena memang kedua data tersebut merupakan potongan dari data RFMT pelanggan.

d. Clustering Data RFMTS Produk yang Dibeli oleh Pelanggan



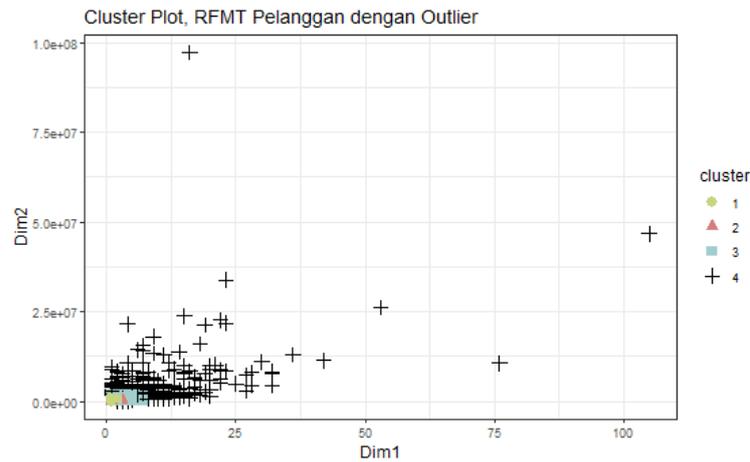
Gambar 14

Cluster Plot Data RFMTS Produk yang Dibeli Pelanggan

Gambar 14. menunjukkan hasil segmentasi dari data RFMTS produk yang dibeli oleh pelanggan dengan jumlah *cluster* sebanyak 5. Dari gambar tersebut, terlihat bahwa 4 *cluster* sangat berdekatan dan hanya ada 1 *cluster* yang berjauhan. Hal ini membuktikan kenapa *gap statistic* menaksir data terlalu rendah atau *underestimated*. Karena telah mendapatkan jumlah *cluster* dan hasil segmentasi, maka selanjutnya dapat dilakukan analisis setiap segmen data berdasarkan RFMT dan RFMTS. Selain itu, perlu dilakukan penambahan satu segmen khusus data *outlier*, ketika data terdapat *outlier* yang dipisahkan.

Analisis RFMT dan RFMTS

Pada bagian ini dilakukan analisis untuk setiap *cluster* berdasarkan nilai RFMT atau RFMTS, sehingga dapat mengetahui karakteristik pelanggan di setiap *cluster* tersebut. Nilai RFMT atau RFMTS setiap *cluster* diambil berdasarkan rata-rata setiap nilainya. Sebelum dilakukan analisis, perlu diketahui ketika *outlier* dimasukkan kembali ke dalam data dan menjadi *cluster* tersendiri, maka *plotting* nya akan terlihat pada Gambar 15.

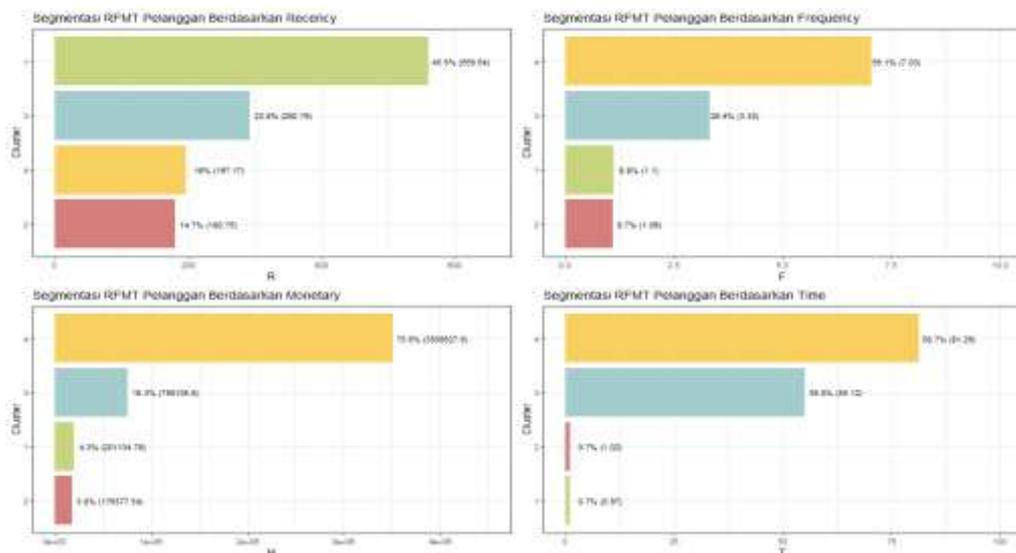


Gambar 15

Cluster Plot Data RFMT Pelanggan dengan Outlier

Berdasarkan Gambar 15 menunjukkan bahwa titik yang berwarna hitam merupakan data *outlier* yang dimasukkan kembali dan dijadikan satu *cluster* tersendiri. Setelah melakukan *clustering* menggunakan algoritma K-Medoids, maka selanjutnya dapat dilakukan analisis RFMT dan RFMTS untuk mengetahui karakteristik setiap *cluster*.

Analisis RFMT Terhadap Pelanggan



Gambar 16

Analisis RFMT Terhadap Pelanggan

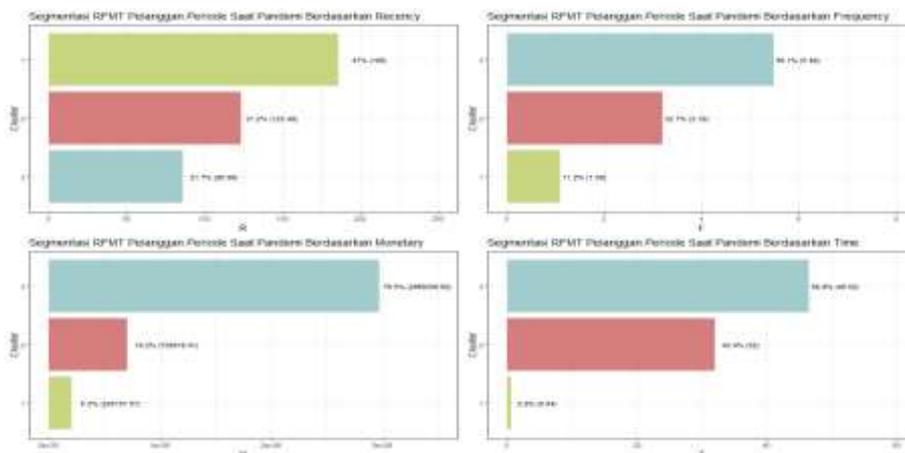
Dari Gambar 16, dapat diketahui karakteristik setiap *cluster* nya berdasarkan RFMT. Pertama, untuk *cluster 1* terlihat memiliki karakteristik R yang sangat tinggi, F yang rendah, M yang sangat rendah, dan T yang sangat rendah. Tinggi rendahnya suatu variabel ditentukan berdasarkan jarak nilai variabel tersebut ke rata-rata nilai suatu variabel. R yang sangat tinggi menandakan bahwa pelanggan di dalam *cluster* tersebut sudah sangat lama/tidak lagi melakukan transaksi di *marketplace* tersebut. F yang rendah menandakan bahwa pelanggan di dalam *cluster* tersebut jarang/ tidak sering melakukan transaksi berulang. M yang sangat rendah artinya pelanggan di dalam *cluster* ini hanya berbelanja sedikit produk/ jumlah uang yang dihabiskannya tergolong sedikit. Terakhir, T yang sangat rendah artinya pelanggan di dalam *cluster* tersebut bukan termasuk langganan karena rentang waktu antara tanggal awal pembelian dan tanggal terakhir pembelian sangatlah kecil atau mendekati nol. Dari penjelasan di atas dapat disimpulkan bahwa karakteristik dalam *cluster 1* adalah pelanggan yang sudah tidak aktif secara belanja dengan jumlah pembelian yang terbatas.

Cluster 2 memiliki R yang rendah, F yang rendah, M yang sangat rendah, dan T yang sangat rendah. R yang rendah menandakan bahwa pelanggan di dalam *cluster* tersebut baru-baru melakukan transaksi. F, M, dan T memiliki karakteristik yang mirip dengan *cluster 1*. Sehingga, dapat disimpulkan bahwa *cluster 2* merupakan pelanggan baru yang melakukan pembelian dalam jumlah terbatas.

Cluster 3 memiliki R yang sedang, F yang sedang, M yang rendah, dan T yang tinggi. R yang sedang menandakan bahwa pelanggan di dalam *cluster* tersebut tidak begitu lama / tergolong masih berbelanja di *marketplace* tersebut. F yang sedang menandakan bahwa pelanggan di dalam *cluster* tersebut masih termasuk sering melakukan transaksi berulang. M yang rendah mengartikan bahwa pelanggan di dalam *cluster* ini berbelanja dengan jumlah uang yang relatif rendah. T yang tinggi mengartikan bahwa pelanggan di dalam *cluster* ini masih tergolong langganan karena rentang waktu tanggal awal pembelian dan tanggal terakhir pembelian cukup tinggi. Dari penjelasan di atas, dapat dikatakan bahwa pelanggan di dalam *cluster 3* merupakan pelanggan yang setia dengan pembelian berulang namun jumlah uang untuk berbelanja yang relatif rendah.

Cluster 4 memiliki R yang rendah, F yang sangat tinggi, M yang sangat tinggi, dan T yang sangat tinggi. R yang rendah menandakan bahwa pelanggan di dalam *cluster* tersebut baru-baru melakukan transaksi. F yang sangat tinggi menandakan bahwa pelanggan di dalam *cluster* tersebut sangat sering melakukan berbelanja berulang. M yang sangat tinggi menandakan bahwa pelanggan di dalam *cluster* tersebut berbelanja dengan jumlah uang yang relatif tinggi. T yang sangat tinggi menandakan bahwa pelanggan tersebut tergolong langganan karena rentang waktu tanggal awal pembelian dan tanggal terakhir pembelian sangat tinggi. Dari penjelasan di atas, dapat disimpulkan bahwa pelanggan di dalam *cluster 4* merupakan pelanggan yang berharga dan loyal.

Analisis RFMT Terhadap Pelanggan Periode Saat Pandemi



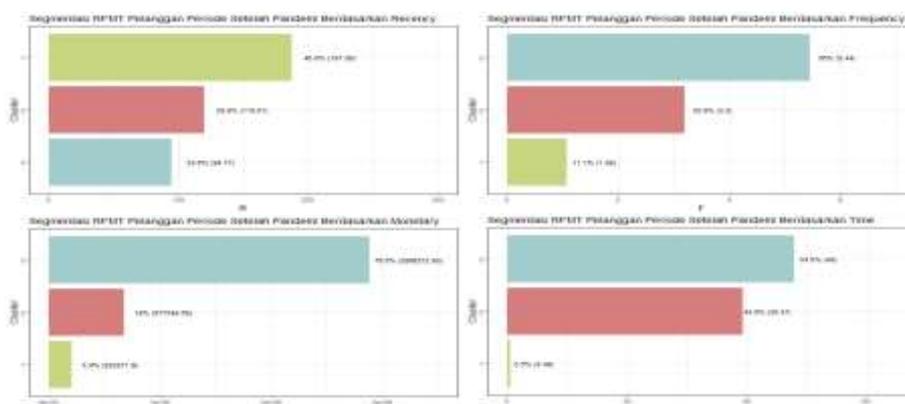
Gambar 17

RFMT Pelanggan Periode Saat Pandemi

Dari Gambar 17. dapat diidentifikasi karakteristik pelanggan setiap *cluster* berdasarkan RFMT. *Cluster 1* memiliki R yang tinggi, F yang rendah, M yang rendah, dan T yang sangat rendah. Dari hal tersebut menjelaskan bahwa pelanggan di *cluster 1* merupakan pelanggan yang tidak aktif.

Cluster 2 memiliki R yang sedang, F yang sedang, M yang sedang, dan T yang tinggi. Dari hal tersebut menandakan bahwa pelanggan di *cluster 2* merupakan pelanggan yang setia. Terakhir, *cluster 3* memiliki R yang rendah, F yang tinggi, M yang sangat tinggi, dan T yang sangat tinggi mengartikan bahwa pelanggan pada *cluster 3* merupakan pelanggan yang berharga dan loyal di periode saat pandemi.

Analisis RFMT Terhadap Pelanggan Periode Setelah Pandemi



Gambar 18

RFMT Pelanggan Periode Setelah Pandemi

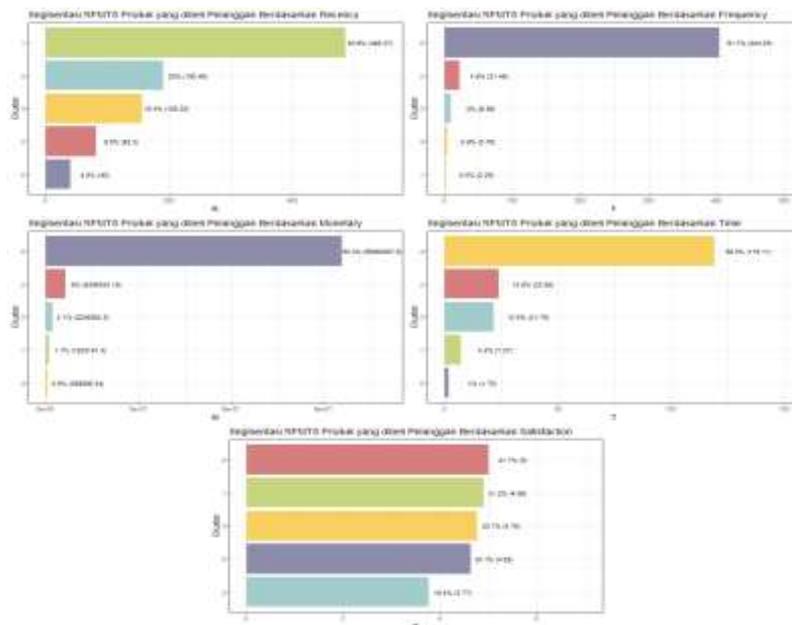
Dari Gambar 18. dapat diidentifikasi karakteristik pelanggan dari setiap *cluster* berdasarkan RFMT. *Cluster 1* memiliki R yang tinggi, F yang rendah, M yang rendah, dan T yang sangat rendah, sehingga dapat dikatakan pelanggan di dalam *cluster 1* merupakan pelanggan yang sudah tidak aktif. *Cluster 2* memiliki R yang sedang, F yang sedang, M yang sedang, dan T yang tinggi, sehingga dapat dikatakan pelanggan di dalam *cluster 2* merupakan

Analisis Segmentasi Pelanggan Toko Online Marketplace Berdasarkan Recency, Frequency, Monetary, Time, Satisfaction (Rfmts) Menggunakan Algoritma K-Medoids Clustering.

pelanggan yang setia. *Cluster 3* memiliki R yang rendah, F yang tinggi, M yang sangat tinggi, dan T yang sangat tinggi, sehingga dapat dikatakan pelanggan di dalam *cluster 3* merupakan pelanggan yang berharga atau loyal.

Dilihat dari karakteristik setiap *cluster* di periode saat pandemi dan setelah pandemi memiliki karakteristik yang sangat mirip. Jadi dapat disimpulkan berdasarkan karakteristik RFMT pelanggan, tidak ada perbedaan yang signifikan antara periode saat pandemi maupun setelah pandemi.

Analisis RFMTS Produk yang Dibeli Oleh Pelanggan



Gambar 19

RFMTS Produk yang Dibeli Pelanggan

Dari Gambar 19. dapat diketahui karakteristik produk dari setiap *cluster* berdasarkan RFMTS. *Cluster 1* memiliki R yang sangat tinggi, F yang sangat rendah, M yang sangat rendah, T yang sangat rendah, dan S yang tinggi. Dari hal tersebut dapat dikatakan bahwa produk di dalam *cluster 1* merupakan produk yang sudah lama tidak dibeli, memiliki harga yang tergolong rendah, tetapi pelanggan puas saat membeli produk ini. *Cluster 2* memiliki R yang rendah, F yang rendah, M yang rendah, T yang sedang, dan S yang sangat tinggi. Dari hal tersebut dapat dikatakan bahwa produk di dalam *cluster 2* merupakan produk yang masih sering dibeli pelanggan, dengan harga yang tergolong sedang, dan pelanggan sangat puas saat membeli produk ini. *Cluster 3* memiliki R yang sedang, F yang rendah, M yang rendah, T yang sedang, dan S yang sedang. Dari hal tersebut dapat dikatakan bahwa produk di dalam *cluster 3* merupakan produk yang sudah jarang dibeli pelanggan dengan harga tergolong rendah, tetapi pelanggan kurang puas saat membeli produk ini. *Cluster 4* memiliki R yang sedang, F yang sangat rendah, M yang sangat rendah, T yang sangat tinggi, dan S yang tinggi. Dari hal tersebut dapat dikatakan bahwa produk di dalam *cluster 4* merupakan produk yang sangat laku dibeli pelanggan dengan harga tergolong rendah, dan pelanggan sangat puas saat membeli produk tersebut. *Cluster 5* memiliki R yang rendah, F yang sangat tinggi, M yang sangat tinggi, T yang sangat rendah, dan S yang tinggi. Dari hal tersebut dapat dikatakan

bahwa produk di *cluster* 5 merupakan produk tertentu yang dibeli pelanggan dengan harga tergolong tinggi, tetapi biasanya pelanggan hanya membeli produk ini sekali atau dua kali saja, dan pelanggan puas saat membeli produk tersebut.

Pemetaan *Cluster* Berdasarkan Provinsi

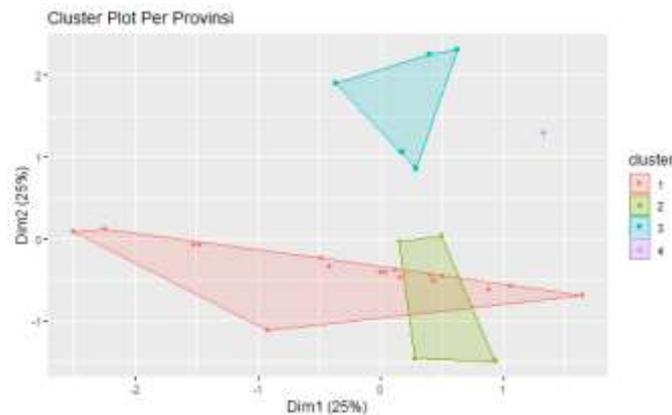
Selain melakukan analisis RFMT dan RFMTS berdasarkan pelanggan dan produk, penelitian ini juga melakukan segmentasi yang dilihat berdasarkan provinsi di Indonesia. Segmentasi ini dilakukan untuk melihat provinsi yang paling berpengaruh dengan meningkatnya penjualan toko hingga provinsi yang kurang berpengaruh dengan penjualan toko (Ayu, 2020). Langkah yang dilakukan untuk mendapatkan data RFMT terhadap provinsi sama seperti yang dilakukan pada subbab 4.4 hingga subbab 4.9. Tabel 4.8 menunjukkan data provinsi yang sudah ditransformasi ke dalam bentuk RFMT.

Tabel 2
Pratinjau Data RFMT Berdasarkan Provinsi

Provinsi	Total Customer	Recency	Frequency	Monetary	Time
Bali	95	21	77	40289006	8
Banten	1172	2	985	276994806	1
Bengkulu	19	222	13	3301828	32
Aceh	11	94	11	3086800	57
Daerah Istimewa Yogyakarta	102	38	91	39737428	7
DKI Jakarta	2859	1	2235	820589251	0
Gorontalo	1	329	1	88000	0
Jambi	47	23	42	7853500	15
Jawa Barat	2464	1	2103	592413715	0
Jawa Tengah	623	2	535	211700633	1

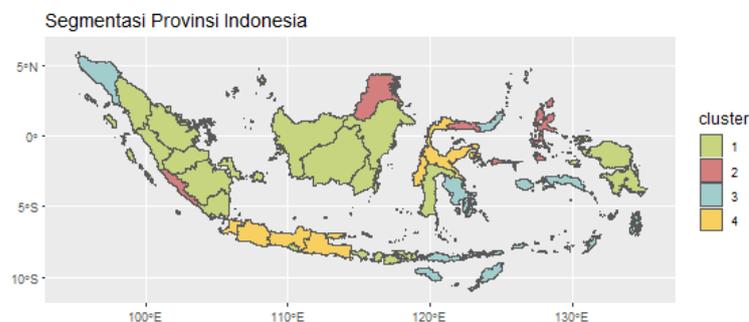
Dari data pada Tabel 2, dilakukan analisis data *outlier* dan memisahkan 5% *outlier* dari data. Sehingga dari total baris data sebanyak 33 baris menjadi 27 baris. Tabel 2. juga analisis korelasi data dan menerapkan PCA. Penentuan *k cluster* sesuai dengan data RFMT terhadap Pelanggan, yaitu sebanyak 4 *cluster*. Setelah mendapatkan *k cluster*, selanjutnya dapat menampilkan visualisasi segmentasi provinsi menggunakan K-Medoids *clustering*. Hasil segmentasi per provinsi dapat dilihat pada Gambar 4.25.

Analisis Segmentasi Pelanggan Toko Online Marketplace Berdasarkan Recency, Frequency, Monetary, Time, Satisfaction (Rfmts) Menggunakan Algoritma K-Medoids Clustering.



Gambar 20

Cluster Plot RFMT Berdasarkan Provinsi



Gambar 21

Segmentasi Berdasarkan Provinsi di Indonesia

Berdasarkan Gambar 4.25, *cluster* per provinsi ada sebanyak 4 *cluster*. Gambar 4.26 merupakan visualisasi negara Indonesia berdasarkan segmentasi per provinsi yang sudah didapatkan sebelumnya.

Dilihat dari Gambar 4.26, dapat diketahui karakteristik pelanggan yang ada di setiap provinsi berdasarkan *cluster* nya. *Cluster* 1 yang merupakan pelanggan yang sudah tidak aktif dengan pembelian terbatas kebanyakan berasal dari Sumatera Utara, Sumatera Barat, Sumatera Selatan, Riau, Lampung, Kepulauan Riau, Kepulauan Bangka Belitung, Jambi, Kalimantan Barat, Kalimantan Selatan, Kalimantan Tengah, Kalimantan Timur, Papua Barat, Sulawesi Selatan, Bali, dan Nusa Tenggara Barat. *Cluster* 2 yang merupakan pelanggan baru dengan jumlah pembelian terbatas kebanyakan berasal dari Bengkulu, Gorontalo, Kalimantan Utara, dan Maluku Utara. *Cluster* 3 yang merupakan pelanggan setia dengan pembelian berulang namun jumlah uang yang dikeluarkan untuk berbelanja relatif rendah, kebanyakan berasal dari Aceh, Maluku, Nusa Tenggara Timur, Sulawesi Tenggara, dan Sulawesi Utara. *Cluster* 4 yang merupakan pelanggan berharga dan loyal kebanyakan berasal dari Pulau Jawa, Sulawesi Barat, dan Sulawesi Tengah.

KESIMPULAN

Dari hasil penelitian yang dilakukan, beberapa kesimpulan penting dapat ditarik. Pertama, dalam menganalisis data penjualan dan segmentasi pelanggan menggunakan K-Medoids clustering berdasarkan nilai RFMT, ditemukan bahwa jumlah cluster yang paling sesuai adalah 4 cluster. Setiap cluster memiliki karakteristiknya sendiri, dengan cluster terbaik

adalah pelanggan berharga/loyal dan cluster terburuk adalah pelanggan yang sudah tidak aktif lagi. Segmen pelanggan yang berpotensi untuk meningkatkan penjualan adalah yang teridentifikasi sebagai cluster ke-3 dan ke-4, karena mereka merupakan pelanggan setia dengan pembelian berulang dan pelanggan loyal yang selalu melakukan pembelian dengan jumlah uang yang tinggi. Kedua, analisis yang dilakukan terhadap periode saat pandemi dan setelah pandemi menunjukkan bahwa jumlah cluster yang sesuai tetap sama, yakni 3 cluster, dengan karakteristik yang identik. Ketiga, hasil segmentasi produk menunjukkan bahwa kebanyakan pelanggan puas dengan produk yang mereka beli, meskipun terdapat segmen pelanggan kurang puas, yaitu cluster ke-3. Terakhir, dalam segmentasi pelanggan berdasarkan nilai RFMT untuk setiap provinsi di Indonesia, ditemukan bahwa pelanggan yang loyal cenderung berasal dari pulau Jawa, Sulawesi Barat, dan Sulawesi Tengah. Kesimpulan ini memberikan wawasan yang berharga bagi strategi pemasaran dan pengembangan produk di masa mendatang.

DAFTAR PUSTAKA

- Anitha, Palaksha, & Patil, Malini M. (2022). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University-Computer and Information Sciences*, 34(5), 1785–1792.
- Ayu, Widia Resti. (2020). *strategi pemasaran dalam meningkatkan penjualan pada toko bag corner ponorogo*. IAIN Ponorogo.
- Fitrayana, Pravasta Rama, & Saputro, Dewi Retno Sari. (2022). Algoritme Clustering Large Application (CLARA) untuk Menangani Data Outlier. *Prisma, Prosiding Seminar Nasional Matematika*, 5, 721–725.
- Han, Jiawei, Pei, Jian, & Tong, Hanghang. (2022). *Data mining: concepts and techniques*. Morgan Kaufmann.
- Izzadin, Fata Mukhammad. (2020). *Optimasi Jumlah Cluster K-Means Dengan Metode Elbow Dan Silhouette Pada Produktivitas Tanaman Pangan Di Provinsi Jawa Tengah Tahun 2018*.
- Lestari, Ayu, & Agnia, Hanum. (2023). Retail Business Strategy In Increasing Sales Through Innovation And Customer Experience. *Jurnal Ekonomi Teknologi Dan Bisnis (JETBIS)*, 2(9), 618–622.
- Lucyana, Nadya, Sari, Purwita, Kurniawan, Dedy, & Buchari, Muhammad Ali. (2022). ANALISIS PENYEBAB KECELAKAAN PESAWAT DI INDONESIA MENGGUNAKAN METODE K-MEANS. *Jusikom: Jurnal Sistem Komputer Musirawas*, 7(2), 89–105.
- Mensouri, Doae, Azmani, Abdellah, & Azmani, Monir. (2022). K-Means customers clustering by their RFMT and score satisfaction analysis. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 13(6).
- Ningsih, Setia, & Dukalang, Hendra H. (2019). Penerapan metode suksesif interval pada analisis regresi linier berganda. *Jambura Journal of Mathematics*, 1(1), 43–53.
- Padma, Panchapakesan, & Ahn, Jiseon. (2020). Guest satisfaction & dissatisfaction in luxury hotels: An application of big data. *International Journal of Hospitality Management*, 84, 102318.

Analisis Segmentasi Pelanggan Toko Online Marketplace Berdasarkan Recency, Frequency, Monetary, Time, Satisfaction (Rfmts) Menggunakan Algoritma K-Medoids Clustering.

Prasetyo, Stevanus Sandy, Mustafid, Mustafid, & Hakim, Arief Rachman. (2020). Penerapan fuzzy c-means kluster untuk segmentasi pelanggan e-commerce dengan metode recency frequency monetary (RFM). *Jurnal Gaussian*, 9(4), 421–433.

Rukajat, Ajat. (2018). *Pendekatan penelitian kualitatif (Qualitative research approach)*. Deepublish.

Sirait, Coki Remando. (2022). *LKP di Bah Butong Sumatera Utara Unit Usaha Teh PT. Perkebunan Nusantara IV*.

Supriyadi, Aceng, Triayudi, Agung, & Sholihati, Ira Diana. (2021). Perbandingan algoritma k-means dengan k-medoids pada pengelompokan armada kendaraan truk berdasarkan produktivitas. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 6(2), 229–240.

Varlamis, Iraklis, Tserpes, Konstantinos, Etemad, Mohammad, Júnior, Amílcar Soares, & Matwin, Stan. (2019). A Network Abstraction of Multi-vessel Trajectory Data for Detecting Anomalies. *EDBT/ICDT Workshops, 2019*.

Zhou, Jinfeng, Wei, Jinliang, & Xu, Bugao. (2021). Customer segmentation by web content mining. *Journal of Retailing and Consumer Services*, 61, 102588.