

# Implementasi Metode Clustering Partisi dalam Menentukan Segmentasi Pelanggan

Nita Mirantika<sup>1\*</sup> dan Estiko Rijanto<sup>2</sup>

<sup>1</sup>Magister Sistem Informasi, Universitas Komputer Indonesia, Jl. Dipati Ukur 112-116 Bandung

<sup>2</sup>Badan Riset dan Inovasi Nasional (BRIN) - Indonesia

\* Korespondensi: Email: [nita.mirantika@gmail.com](mailto:nita.mirantika@gmail.com), [estikorijanto@gmail.com](mailto:estikorijanto@gmail.com)

Diterima: 02-11-2023 ; Review: 09-11-2023; Disetujui: 12-07-2024

Cara sitasi: Mirantika, N., Rijanto, E., . 2024. Implementasi Metode Clustering Partisi dalam Menentukan Segmentasi Pelanggan, Jurnal Tata Kelola dan Kerangka Kerja TI, 10(1): 8-16

**ABSTRAK** – Segmentasi pelanggan termasuk bagian dari strategi pemasaran yang diperlukan perusahaan melalui pengelompokan pelanggan yang memiliki kesamaan karakteristik. PT XYZ yang bergerak di bidang peternakan memiliki pelanggan yang tersebar di berbagai wilayah dengan karakteristik yang beragam sehingga membutuhkan segmentasi pelanggan dalam membuat strategi pemasarannya. Penelitian ini bertujuan untuk melakukan segmentasi pelanggan dengan menggunakan metode clustering partisi yaitu algoritma K-Means dan algoritma K-Medoids serta membandingkan kinerja kedua algoritma tersebut. Model RFM (Recency-Frequency-Monetary) digunakan dalam pemilihan atribut penelitian. Metode Elbow digunakan dalam penentuan jumlah cluster optimum dan metode Davies-Bouldin Index (DBI) digunakan dalam evaluasi hasil clustering. Segmentasi pelanggan menggunakan algoritma K-Means menghasilkan tiga jenis pelanggan yaitu superstar, typical customer dan dormant customer. Sedangkan segmentasi pelanggan menggunakan algoritma K-Medoids menghasilkan empat jenis pelanggan yaitu superstar, typical customer, customer needing attention dan dormant customer. Hasil evaluasi clustering diperoleh nilai DBI K-Means jauh lebih kecil dari nilai DBI K-Medoids. Hasil ini menunjukkan bahwa pada metode clustering partisi, algoritma K-Means lebih baik daripada algoritma K-Medoids dalam menentukan segmentasi pelanggan.

**Kata Kunci** – Clustering Partisi, Algoritma K-Means, Algoritma K-Medoids, RFM, DBI

## 1. PENDAHULUAN

Segmentasi pelanggan adalah mengelompokkan pelanggan yang memiliki kesamaan karakteristik [1][2]. Segmentasi pelanggan diperlukan oleh perusahaan dalam membuat strategi pemasaran agar pemasaran dapat tepat sasaran [3][4]. PT XYZ merupakan perusahaan yang bergerak pada bidang peternakan dan memiliki pelanggan yang tersebar diberbagai daerah Propinsi Jawa Barat, Jawa Tengah dan Jawa Timur. Pelanggan PT XYZ yang datang dari berbagai daerah memiliki karakteristik yang beragam, hal ini menuntut PT XYZ dapat membuat keputusan yang tepat dalam strategi pemasaran yaitu dengan melakukan segmentasi pelanggan. Salah satu solusi dalam melakukan segmentasi pelanggan diantaranya melalui penerapan teknik *data mining clustering partisi* [5][6][7].

Metode *clustering partisi* terdiri dari beberapa algoritma diantaranya adalah algoritma K-Means dan algoritma K-Medoids. Kedua algoritma tersebut merupakan algoritma partisi yang paling terkenal

[8]. Algoritma K-Means merupakan algoritma *clustering partisi* yang sangat populer dan banyak dipelajari karena sederhana dan mudah diterapkan [7][9]. Nilai rata-rata digunakan pada algoritma K-Means untuk menjadi pusat klaster [10][11]. Algoritma K-Medoids mempunyai nama lain algoritma PAM (*Partitioning Around Medoid*) merupakan metode *clustering* yang cukup peka terhadap *outlier* [10]. Perwakilan objek (*medoid*) digunakan pada algoritma K-Medoids untuk menjadi pusat *cluster* [12]. Pemilihan atribut dalam penerapan segmentasi pelanggan dapat menggunakan model RFM (*Recency-Frequency-Monetary*). Penilaian perilaku pelanggan diterapkan pada model RFM berdasarkan tiga atribut yaitu *recency* untuk penilaian waktu transaksi pelanggan yang paling terakhir, *frequency* untuk penilaian jumlah total dari keseluruhan transaksi pelanggan, dan *monetary* untuk penilaian jumlah total uang yang dibelanjakan pelanggan [13][14]. Dengan mengetahui *history* transaksi pelanggan, perusahaan dapat melakukan penilaian pada tiga atribut tersebut

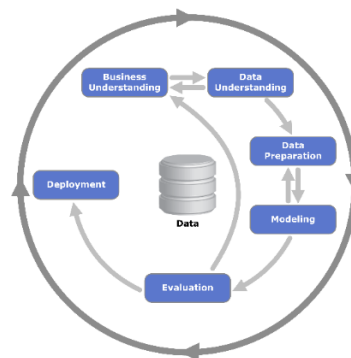
[1].

Penelitian mengenai segmentasi pelanggan menggunakan metode *clustering* partisi telah dilakukan oleh Shirole, dkk [5] yaitu segmentasi pelanggan menggunakan model RFM dan algoritma K-Means *clustering* yang menghasilkan empat *cluster* dari *dataset* e-commerce berdasarkan skor RFM. Hasil tersebut dapat membantu perusahaan untuk mengembangkan strategi pasar dan sebagai media promosi kepada pelanggan setia. Penelitian lain juga dilakukan oleh Sheshasaayee, dkk [7] yaitu implementasi analisis RFM berbasis teknik *clustering* untuk perilaku nasabah dalam transaksi online menggunakan algoritma K-Means. Penelitian tersebut menghasilkan dua *cluster* pelanggan dari *dataset* belanja online. Selanjutnya penelitian lain dilakukan oleh Aryuni, dkk [12] yaitu mengenai segmentasi nasabah pada bank XYZ dengan menerapkan algoritma K-Means dan K-Medoids. Penelitian mendapatkan hasil bahwa algoritma K-Means lebih unggul dari algoritma K-Medoids dilihat dari jarak intra *cluster*. Sedangkan dilihat dari nilai Davies-Bouldin Index, kinerja algoritma K-Means sedikit lebih unggul dari algoritma K-Medoids. Selain dalam hal segmentasi pelanggan, terdapat penelitian yang dilakukan untuk menganalisis perbandingan penerapan algoritma K-Means dan algoritma K-Medoids diantaranya dilakukan oleh Utomo [15] yaitu untuk *clustering* penyebaran wabah Covid-19 yang melanda Indonesia. Penelitian mendapatkan hasil bahwa berdasarkan nilai Davies Boulden index metode terbaik untuk mengelompokkan penyebaran wabah virus corona di Indonesia adalah metode K-Means. Penelitian lain yang dilakukan oleh Nirmal [16] dalam membandingkan algoritma K-Medoids dan algoritma K-Means diperoleh hasil bahwa algoritma K-medoids lebih baik dalam mengelompokkan data ketika terdapat outlier daripada algoritma kmeans. K-medoids lebih baik dalam skalabilitas untuk *dataset* yang lebih besar dan juga karena lebih efisien daripada K-means.

Berdasarkan penelitian terdahulu yang telah dipaparkan diatas, maka dilakukan penelitian ini untuk menghasilkan segmentasi pelanggan menggunakan metode *clustering* partisi yaitu algoritma K-Means dan algoritma K-Medoids serta untuk membandingkan kinerja kedua algoritma *clustering* partisi tersebut dalam melakukan segmentasi pelanggan di PT XYZ. Segmentasi pelanggan yang dihasilkan dapat menjadi masukan bagi PT XYZ dalam membuat rencana strategi pemasaran.

## 2. METODOLOGI PENELITIAN

Metode CRISP-DM menjadi acuan dalam melakukan penelitian ini. Metode CRISP-DM merupakan standar dalam pengembangan proyek *data mining* dan knowledge discovery [17]. Dalam prosesnya CRISP-DM mempunyai enam tahapan pengembangan *data mining* [18]. Gambar 1 menunjukkan enam tahapan proses CRISP-DM.



**Gambar 1.** Siklus proses CRISP-DM  
(Sumber : Chapman dkk, 2000)

Tahapan penelitian yang dilakukan berdasarkan metode CRISP-DM yaitu:

### 2.1 Business understanding

Tahap *business understanding* meliputi kegiatan untuk mengetahui tujuan dan kebutuhan perusahaan, kemudian diterjemahkan ke dalam *data mining* [18]. PT XYZ mempunyai pelanggan yang banyak di berbagai daerah dengan karakteristik yang beragam sehingga membutuhkan segmentasi pelanggan. Pada *data mining*, metode *clustering* dapat digunakan untuk melakukan segmentasi pelanggan [7].

### 2.2 Data understanding

Tahap *data understanding* meliputi kegiatan mengumpulkan data dan melakukan pemahaman yang mendalam tentang data [18]. Data transaksi penjualan PT XYZ digunakan pada penelitian ini yaitu periode Januari 2021 – Desember 2021 dengan jumlah transaksi sebanyak 2.584 transaksi dan 456 pelanggan.

### 2.3 Data preparation

Tahap *data preparation* meliputi kegiatan mendapatkan data akhir untuk dilakukan pemodelan. Pada tahap ini meliputi pemilihan atribut data, *record*, tabel, proses pembersihan dan transformasi data [18]. Dalam melakukan segmentasi pelanggan, pemilihan atribut menggunakan model RFM yang pada tahun 1994 diperkenalkan oleh Arthur Huges [13][14]. Model RFM melakukan proses penilaian berdasarkan tiga perilaku

pelanggan yaitu: penilaian berdasarkan waktu transaksi pelanggan yang paling terakhir terdiri dari tanggal, bulan dan tahun disebut *recency* (R), penilaian pelanggan berdasarkan jumlah total dari keseluruhan transaksi pelanggan disebut *frequency* (F) dan penilaian pelanggan berdasarkan jumlah total uang yang dibelanjakan oleh pelanggan disebut *monetary* (M) [22]. Tabel 1 menunjukkan *dataset* penelitian menggunakan tiga atribut RFM.

**Tabel 1.** *Dataset* Penelitian

Pelanggan	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
Pelanggan 1	29	2	67500000
Pelanggan 2	64	5	169500000
.....	.....	.....	.....
Pelanggan 455	32	5	172000000
Pelanggan 456	39	4	137500000

Setelah mendapatkan *dataset*, selanjutnya dilakukan pembersihan data. *Dataset* yang diperoleh mungkin tidak lengkap, tidak konsisten, mengandung kesalahan, dan terdapat atribut yang hilang atau atribut yang mengandung data duplikat. Hal ini dapat menurunkan kualitas hasil penelitian, oleh karena itu dilakukan pembersihan data [5]. Pembersihan data dilakukan dengan menggunakan software RStudio. *Dataset* yang sudah dibersihkan selanjutnya dilakukan transformasi data untuk meningkatkan akurasi model. Pada tahap transformasi data, supaya diperoleh skala data yang tidak terlalu jauh maka dilakukan normalisasi [2][13]. Dalam penelitian ini, proses normalisasi menggunakan metode Z-Index. Normalisasi Z-Index merupakan teknik dalam normalisasi data, dimana nilai rata-rata (mean) dan standar deviasi digunakan untuk menormalisasikan nilai atribut A [8]. Tabel 2 menunjukkan hasil normalisasi data.

**Tabel 2.** Normalisasi *Dataset* Penelitian

Pelanggan	<i>Zrecency</i>	<i>Zfrequency</i>	<i>Zmonetary</i>
Pelanggan 1	-0.985009110	-0.22474040	-0.177628220
Pelanggan 2	-0.621288195	-0.04086189	-0.024807743
.....	.....	.....	.....
Pelanggan 455	-0.953833031	-0.04086189	-0.021062143
Pelanggan 456	-0.881088848	-0.10215473	-0.072751422

## 2.4 Modeling

Tahap *modeling* meliputi kegiatan implementasi berbagai teknik pemodelan yaitu berupa algoritma *data mining* [18] dalam hal ini adalah membandingkan metode *clustering* partisi algoritma K-Means dengan algoritma K-Medoids.

### 2.4.1 Algoritma K-Means.

Pemodelan algoritma K-Means termasuk

*clustering* partisi yang sangat terkenal sehingga banyak dipelajari karena sederhana dan mudah diterapkan [7][9]. Nilai rata-rata digunakan pada algoritma K-Means untuk menjadi pusat kluster [10][11]. Algoritma K-Means mempunyai tahapan sebagai berikut:

- 1) Penentuan jumlah kluster optimum yang akan dibuat, dengan menggunakan metode Elbow.
- 2) Penginisialisasian pusat *cluster* awal (*centroid*) pada setiap *cluster* secara acak.
- 3) Penghitungan jarak setiap data pengamatan ke pusat kluster (*centroid*) terpilih dihitung menggunakan rumus Euclidian Distance, pada persamaan (1)

$$dist(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

- 4) Mengelompokkan setiap objek pengamatan berdasarkan jarak yang paling dekat dengan *centroid*.
- 5) *Centroid* baru diperbaharui menggunakan nilai rata-rata kluster
- 6) Mengulangi langkah tiga sampai dengan lima, sehingga diperoleh anggota setiap kluster tidak berubah.

### 2.4.2 Algoritma K-Medoids.

Pemodelan algoritma K-Medoids dengan nama lain algoritma PAM (*Partitioning Around Medoid*) merupakan salah satu metode *clustering* partisi yang diusulkan untuk meningkatkan algoritma K-Means dan menyelesaikan beberapa masalah, termasuk sensitivitas terhadap *outlier* [10][19]. Perwakilan objek (*medoid*) digunakan pada algoritma K-Medoids untuk menjadi pusat cluster. *Medoid* merupakan objek data yang paling sentral diantara titik *cluster* [10]. Tahapan dalam algoritma K-Medoids adalah:

- 1) Penentuan jumlah kluster optimum yang akan dibuat, dengan menggunakan metode Elbow.
- 2) Menginisialisasi pusat *cluster* awal (*medoid*) secara acak dari masing-masing *cluster*.
- 3) Penghitungan jarak setiap data pengamatan ke kluster paling dekat dihitung dengan rumus Euclidian Distance pada persamaan (1).
- 4) Kandidat *medoid* baru dihitung dengan acak dari data pengamatan pada masing-masing kluster.
- 5) Jarak tiap objek pada setiap kluster dengan calon *medoid* yang baru dihitung.
- 6) Total simpangan (S) dihitung yaitu mengurangi nilai total *distance* yang baru dengan total *distance* yang lama. Objek ditukar dengan data *cluster* sehingga didapatkan objek yang baru menjadi *medoid*, apabila nilai  $S < 0$ .

- 7) Mengulangi tahap empat sampai dengan enam, sehingga diperoleh *medoid* yang tidak berubah.

### 2.4.3 Metode Elbow.

Metode Elbow merupakan metode visual untuk mendapatkan jumlah *cluster* ( $k$ ) optimal yang akan dibentuk. Tahapannya dimulai dengan nilai kluster ( $k$ ) sama dengan 2 dan nilai  $k$  terus ditingkatkan disetiap tahapannya sebanyak satu  $k$ . Nilai  $k$  akan menjadi titik di mana total jumlah kuadrat dalam *cluster* turun drastis. Pada titik tersebut, grafik akan mencapai dataran tinggi ketika  $k$  meningkat lebih jauh sehingga membentuk sebuah siku [20].

### 2.5 Evaluation

Tahap *evaluation* dilakukan untuk mengetahui pencapaian tujuan yang telah ditetapkan pada tahap awal [18]. Evaluasi hasil *clustering* menggunakan nilai Davies-Bouldin Index (DBI), dapat dilihat dari nilai rasionya yang diperoleh dari pembagian *average within* (SSW) dengan *average between* (SSB). Hasil *clustering* yang paling baik diperoleh jika nilai DBI yang dihasilkan paling kecil [21].

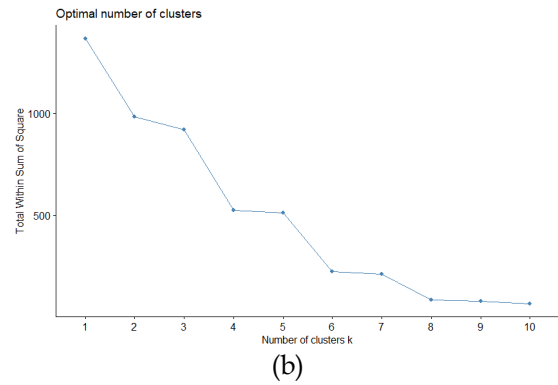
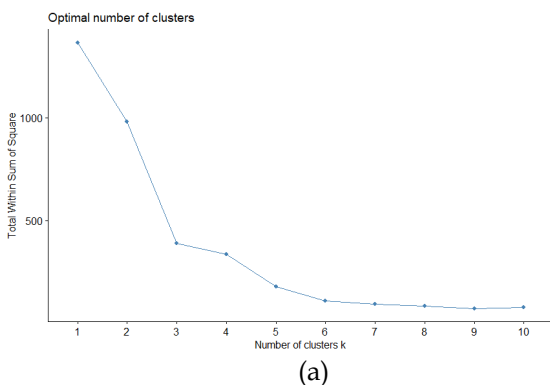
### 2.6 Deployment

*Deployment* adalah tahapan terakhir yaitu pengetahuan yang sudah diperoleh berupa segmentasi pelanggan disajikan dalam bentuk laporan agar PT XYZ dapat dengan mudah memahaminya. Laporan segmentasi pelanggan yang dihasilkan dapat digunakan sebagai masukan dalam membuat rekomendasi strategi pemasaran bagi perusahaan.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Jumlah Cluster Optimum

Metode Elbow digunakan dalam penentuan jumlah *cluster* optimum, dimana nilai *cluster* ( $k$ ) akan menjadi titik di mana total jumlah kuadrat dalam *cluster* turun drastis. Pada titik tersebut, grafik akan mencapai dataran tinggi ketika  $k$  meningkat lebih jauh sehingga membentuk sebuah siku. Gambar 2 menunjukkan hasil penentuan jumlah *cluster* optimum.

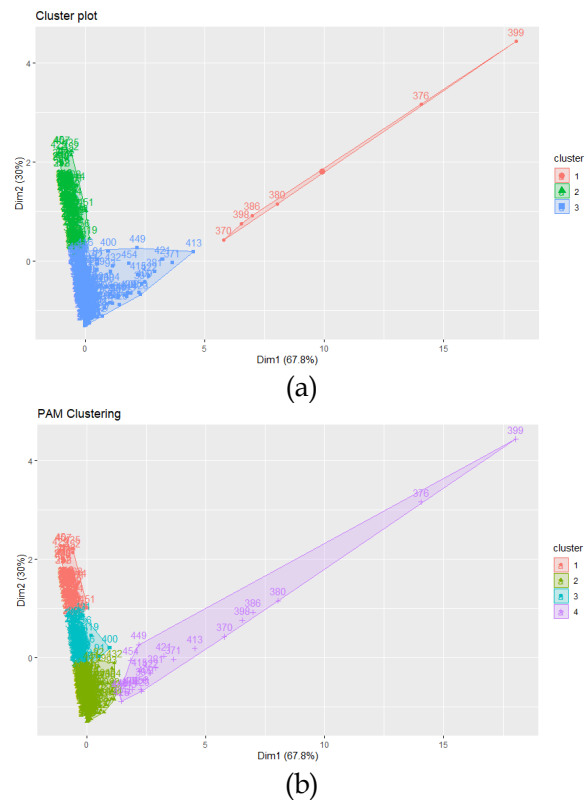


Gambar 2. Jumlah *Cluster* Optimum (a) Algoritma K-Means dan (b) Algoritma K-Medoids

Melalui penggunaan metode Elbow, jumlah *cluster* optimum pada algoritma K-Means diperoleh sebanyak 3 *cluster*. Sementara pada algoritma K-Medoids diperoleh jumlah *cluster* optimum sebanyak 4 *cluster*.

### 3.2 Hasil Clustering

Setelah mengetahui jumlah *cluster* optimum langkah selanjutnya adalah membentuk *cluster*. Gambar 3 menunjukkan hasil *cluster* algoritma K-Means dan K-Medoids dengan menggunakan aplikasi Rstudio.



Gambar 3. Hasil *Cluster* (a) Algoritma K-Means dan (b) Algoritma K-Medoids

Hasil *clustering* dengan menerapkan algoritma K-

Means diperoleh tiga buah cluster yang terdiri dari *cluster 1* terdapat 6 pelanggan, untuk *cluster 2* terdapat 170 pelanggan, dan untuk *cluster 3* terdapat 280 pelanggan. Pada algoritma K-Medoids diperoleh empat buah cluster yang terdiri dari *cluster 1* terdapat 112 pelanggan, untuk *cluster 2* terdapat 221 pelanggan, untuk *cluster 3* terdapat 98 pelanggan, dan untuk *cluster 4* terdapat 25 pelanggan.

### 3.3 Hasil Segmentasi Pelanggan

Untuk mengetahui segmentasi pelanggan yang dihasilkan dari proses *clustering*, diperoleh dengan membandingkan nilai rata-rata (*mean*) pada tiap atribut *recency*, *frequency* dan *monetary*. Cluster dengan nilai *recency* tinggi, *frequency* tinggi dan *monetary* tinggi termasuk dalam segmen *superstar* yaitu pelanggan yang sudah loyal pada perusahaan. Cluster dengan nilai *recency* sedang, *frequency* sedang dan *monetary* sedang termasuk dalam segmen *typical customer* yaitu pelanggan biasa yang berpotensi untuk menjadi pelanggan yang loyal. Cluster dengan nilai *recency* rendah, *frequency* rendah dan *monetary* rendah termasuk segmen *dormant customer* yaitu pelanggan yang sudah tidak aktif lagi. Cluster dengan nilai *recency* rendah, *frequency* sedang dan *monetary* rendah termasuk segmen *Customer Needing Attention* yaitu pelanggan yang membutuhkan perhatian karena dikhawatirkan akan menjadi *dormant customer*. Tabel 3 dan 4 menunjukkan hasil segmentasi pelanggan.

**Tabel 3.** Hasil Segmentasi Pelanggan Algoritma K-Means

Cluster	Mean (Nilai Rata-rata)			Jenis Pelanggan
	Recency	Frequency	Monetary	
Cluster 1	6,5	119	4.981.573.641	Superstar
Cluster 2	234,4	1,80	56.606.500	Dormant customer
Cluster 3	59,13	5,59	161.892.380	Typical customer

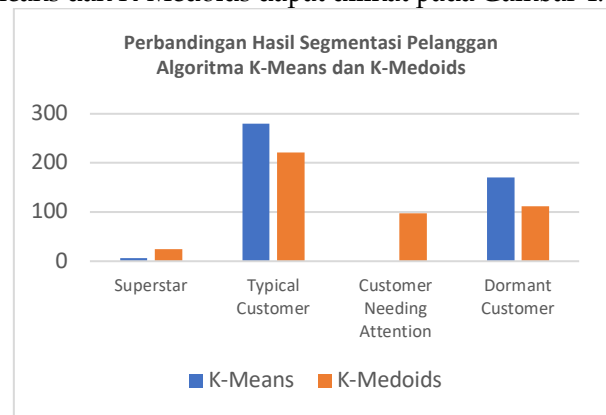
Segmentasi pelanggan pada algoritma K-Means menghasilkan tiga jenis pelanggan, diantaranya jenis pelanggan *superstar* yaitu pelanggan yang sudah loyal pada perusahaan, *dormant customer* yaitu pelanggan yang sudah tidak potensial lagi dan *typical customer* yaitu pelanggan biasa yang berpotensi untuk menjadi pelanggan yang loyal.

**Tabel 4.** Hasil Segmentasi Pelanggan Algoritma K-Medoids

Cluster	Mean (Nilai Rata-rata)			Jenis Pelanggan
	Recency	Frequency	Monetary	
Cluster 1	266	1,67	55.802.337	Dormant customer
Cluster 2	50,4	3,44	111.720.420	Typical customer
Cluster 3	153,4	2,31	71.343.238	Customer needing attention
Cluster 4	18,76	56,4	1.876.428.037	Superstar

Segmentasi pelanggan pada algoritma K-Medoids menghasilkan empat jenis pelanggan, diantaranya jenis pelanggan *dormant customer* yaitu pelanggan yang sudah tidak potensial lagi, *typical customer* yaitu pelanggan biasa yang berpotensi untuk menjadi pelanggan yang loyal, *customer needing attention* yaitu pelanggan yang membutuhkan perhatian karena dikhawatirkan akan menjadi *dormant customer*, dan *superstar* yaitu pelanggan yang sudah loyal pada perusahaan.

Dengan hasil segmentasi pelanggan pada Tabel 3 dan 4, perbandingan penerapan pada algoritma K-Means dan K-Medoids dapat dilihat pada Gambar 4.



**Gambar 4.** Perbandingan Hasil Segmentasi Pelanggan algoritma K-Means dan K-Medoids

Pada perbandingan hasil diatas, dapat diketahui bahwa terdapat perbedaan jenis pelanggan yang dihasilkan pada algoritma K-Means dan algoritma K-Medoids. Terdapat tiga jenis pelanggan yang dihasilkan pada algoritma K-Means, sedangkan algoritma K-Medoids terdapat empat jenis pelanggan. Pelanggan jenis *customer needing attention* hanya terdapat pada algoritma K-Medoids saja. Anggota cluster untuk jenis pelanggan *typical customer* dan *dormant customer* berjumlah lebih banyak pada algoritma K-Means dibandingkan jumlah pada algoritma K-Medoids. Sedangkan untuk pelanggan jenis *superstar* anggota cluster pada algoritma K-Medoids berjumlah lebih banyak daripada algoritma K-Means.

### 3.4 Evaluasi Clustering

Nilai DBI digunakan pada proses evaluasi hasil *clustering*. Nilai DBI menjadi acuan untuk memperoleh hasil klaster yang optimal, yaitu diperoleh jika nilai BDI paling kecil atau minimum. Tabel 5 berisi hasil nilai DBI.

**Tabel 5.** Nilai DBI

Algoritma	Nilai DBI
K-Means	0,696
K-Medoids	1,610

Berdasarkan nilai DBI, dapat diperoleh hasil bahwa nilai DBI pada algoritma K-Means jauh lebih kecil daripada nilai DBI algoritma K-Medoids. Hal ini menunjukkan bahwa algoritma K-Means lebih baik daripada algoritma K-Medoids dalam menentukan segmentasi pelanggan. Hasil penelitian ini selaras dengan yang dilakukan oleh Aryuni, dkk [12] dan Utomo [15] bahwa algoritma K-Means lebih baik daripada algoritma K-Medoids berdasarkan nilai DBI. Oleh karena itu, penulis merekomendasikan menggunakan algoritma K-Means dalam menentukan segmentasi pelanggan di PT XYZ. Hasil segmentasi pelanggan tersebut dapat menjadi masukan pada PT XYZ dalam membuat strategi pemasaran agar tepat sasaran.

#### 4. KESIMPULAN

Penelitian yang dilakukan di PT. XYZ dengan mengimplementasikan metode *clustering* partisi, yaitu algoritma K-Means dan algoritma K-Medoids diperoleh segmentasi pelanggan yang dapat menjadi masukan pada PT XYZ dalam membuat strategi pemasaran agar tepat sasaran. Segmentasi pelanggan menggunakan algoritma K-Means menghasilkan tiga jenis pelanggan yaitu *superstar*, *typical customer* dan *dormant customer*. Sedangkan segmentasi pelanggan menggunakan algoritma K-Medoids menghasilkan empat jenis pelanggan yaitu *superstar*, *typical customer*, *customer needing attention* dan *dormant customer*. Penelitian ini juga berupaya membandingkan kinerja penerapan algoritma K-Means dan algoritma K-Medoids, metode Davies Bouldin Index (DBI) digunakan dalam penentuan algoritma terbaik tersebut. Berdasarkan nilai DBI, diperoleh hasil nilai DBI K-Means jauh lebih kecil dari nilai DBI K-Medoids. Hasil ini menunjukkan bahwa pada metode *clustering* partisi, algoritma K-Means lebih baik daripada algoritma K-Medoids dalam melakukan segmentasi pelanggan pada PT XYZ.

#### Ucapan Terima Kasih

Dengan selesainya penelitian ini, penulis sampaikan rasa terima kasih yang sebesar-besarnya pada PT XYZ yang telah memberikan izin penelitian di perusahaannya, dan kepada Universitas Komputer Indonesia khususnya untuk Fakultas Pascasarjana Magister Sistem Informasi atas support yang besar baik material maupun non material selama penelitian dilakukan sehingga dapat berjalan dengan baik.

#### Daftar Pustaka

- [1] Shihab, S. H., Afroge, S., & Mishu, S. Z. (2019, February). RFM based market segmentation approach using advanced k-means and agglomerative clustering: a comparative study. In 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE) (pp. 1-4). IEEE
- [2] Adiana, B. E., Soesanti, I., & Permanasari, A. E. (2018). Analisis segmentasi pelanggan menggunakan kombinasi RFM model dan teknik clustering. *Jurnal Terapan Teknologi Informasi*, 2(1), 23-32
- [3] Sulistyawati, A. A. D., & Sadikin, M. (2021). Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan. *Sistemasi: Jurnal Sistem Informasi*, 10(3), 516-526.
- [4] Murpratiwi, S. I., Indrawan, I. G. A., & Aranta, A. (2021). Analisis Pemilihan Cluster Optimal Dalam Segmentasi Pelanggan Toko Retail. *Jurnal Pendidikan Teknologi dan Kejuruan*, 18(2), 152-163.
- [5] Shirole, R., Salokhe, L., & Jadhav, S. (2021). Customer Segmentation using RFM Model and K-Means Clustering. *More references*
- [6] Hung, P. D., Lien, N. T. T., & Ngoc, N. D. (2019, March). Customer segmentation using hierarchical agglomerative clustering. In *Proceedings of the 2019 2nd International Conference on Information Science and Systems* (pp. 33-37).
- [7] Sheshasaayee, A., & Logeshwari, L. (2018, May). Implementation of clustering technique based RFM analysis for customer behaviour in online transactions. In *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 1166-1170). IEEE.
- [8] Vercellis, C. (2009). *Business intelligence: data mining and optimization for decision making*. John Wiley & Sons.
- [9] Wei, J. T., Lin, S. Y., Yang, Y. Z., & Wu, H. H. (2020). Using a combination of RFM model and cluster analysis to analyze customers' values of a veterinary hospital. *IAENG International Journal of Computer Science*, 47(3), 1-7.
- [10] Mousavi, S., Boroujeni, F. Z., & Aryanmehr, S. (2020). Improving customer clustering by optimal selection of cluster centroids in K-means and K-medoids algorithms. *Journal of Theoretical and Applied Information Technology*, 98(18).
- [11] Muningsih, E. (2018). Komparasi Metode Clustering K-Means dan K-Medoids dengan Model Fuzzy RFM untuk Pengelompokan Pelanggan. *Evolusi: Jurnal Sains Dan Manajemen*, 6(2).
- [12] Aryuni, M., Madyatmadja, E. D., & Miranda, E. (2018, September). Customer segmentation in XYZ bank using K-means and K-medoids clustering. In *2018 International Conference on Information Management and Technology (ICIMTech)* (pp. 412-416). IEEE.
- [13] Monalisa, S., Nadya, P., & Novita, R. (2019). Analysis for customer lifetime value categorization with RFM model. *Procedia Computer Science*, 161, 834-840.
- [14] Parikh, Y., & Abdelfattah, E. (2020, October). Clustering algorithms and RFM analysis performed on retail transactions. In *2020 11th IEEE Annual*

- Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 0506-0511). IEEE.
- [15] Utomo, W. (2021). The comparison of k-means and k-medoids algorithms for clustering the spread of the covid-19 outbreak in Indonesia. *ILKOM Jurnal Ilmiah*, 13(1), 31-35
- [16] Nirmal, S. (2019). Comparative study between k-means and k-medoids clustering algorithms. *J. Classif*, 6, 839-844
- [17] Mariscal, G., Marban, O., & Fernandez, C. (2010). A survey of data mining and knowledge discovery process models and methodologies. *The Knowledge Engineering Review*, 25(2), 137-166.
- [18] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. SPSS inc.
- [19] Rahman, F., Ridho, I. I., Muflih, M., Pratama, S., Raharjo, M. R., & Windarto, A. P. (2020, April). Application of Data Mining Technique using K-Medoids in the case of Export of Crude Petroleum Materials to the Destination Country. In *IOP Conference Series: Materials Science and Engineering* (Vol. 835, No. 1, p. 012058). IOP Publishing
- [20] Hung, PD, Lien, NTT, & Ngoc, ND (2019, Maret). Segmentasi pelanggan menggunakan pengelompokan aglomerasi hierarkis. Dalam *Prosiding Konferensi Internasional ke-2 2019 tentang Ilmu dan Sistem Informasi* (hlm. 33-37).
- [21] Maulik, U., & Bandyopadhyay, S. (2002). Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on pattern analysis and machine intelligence*, 24(12), 1650-1654
- [22] Wardani, N. W., Dantes, G.R., & Indrawan, G. (2018). Prediksi customer churn dengan algoritma decision tree C4.5 berdasarkan segmentasi pelanggan untuk mempertahankan pelanggan pada perusahaan retail. *Jurnal RESISTOR (Rekayasa Sistem Komputer)*, 1(1), 16-24.