

BAB III

METODE PENELITIAN

3.1 Subjek dan Objek Penelitian

3.1.1. Subjek Penelitian

Subjek penelitian ini merupakan pelanggan di Kedai Kenari, Kota Semarang, Jawa Tengah.

3.1.2. Objek Penelitian

Objek penelitian ini merupakan analisis perilaku pelanggan menggunakan nilai *Recency, Frequency, Monetary* (RFM).

3.2 Alat dan Bahan

Penelitian ini memerlukan pemanfaatan alat dan materi sebagai elemen pendukung untuk mencapai kesuksesan penelitian. Alat dan bahan yang dimaksud melibatkan:

3.2.1 Alat

Alat terbagi menjadi perangkat keras dan perangkat lunak. Rincian keduanya mencakup:

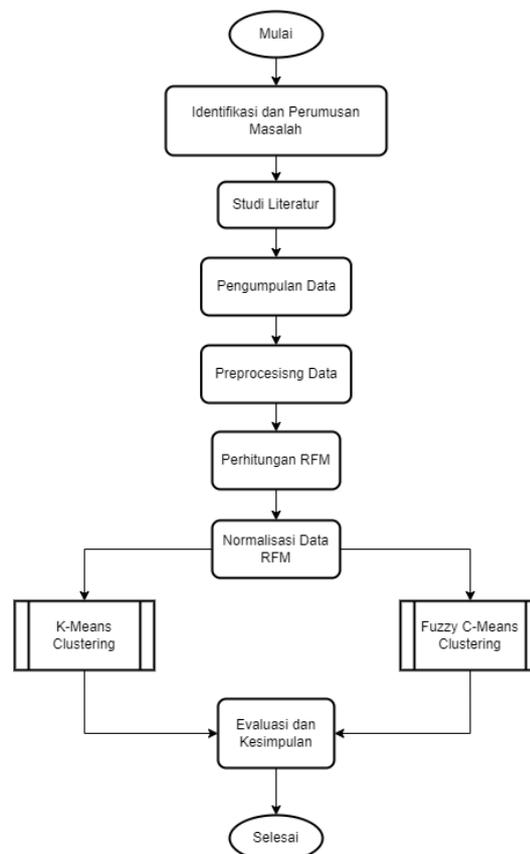
1. Perangkat Keras (*Hardware*)
 - a. Device : MSI Modern 14 B5M
 - b. Processor : Processor AMD Ryzen 7 5700U with Radeon Graphics 1.80 GHz
 - c. RAM : 8 GB
2. Perangkat Lunak (*Software*)
 - a. Sistem Operasi : *Windows 11 Home Single Language*
 - b. Bahasa Program : *Python*
 - c. Aplikasi : *Google Collabs* (Membuat Program),
Microsoft Word (Membuat Proposal),
Microsoft Excel (Mengumpulkan Data Mentah)

3.2.2 Bahan

Dataset yang dipakai dalam penelitian ini adalah *dataset* yang mencakup sejumlah atribut, termasuk *Customer ID*, *Outlet*, *Date*, *Time*, *Gross Sales*, *Discounts*, *Refunds*, *Net Sales*, *Gratuity*, *Tax*, *Total Collected*, *Total Amount*, *Other Note (Optional)*, *Receipt Number*, *Collected By*, *Served By*, *Customer*, *Customer Phone*, *Items*, *Payment Method*, *Event Type*, dan *Reason of Refund*. Data tersebut yang kemudian menjadi *dataset* pada penelitian ini.

3.3 Diagram Alir Penelitian

Dalam menyusun laporan penelitian, terdapat beberapa langkah yang diikuti. Berikut adalah diagram alur yang menggambarkan proses penelitian yang dilakukan dalam penyusunan laporan ini:



Gambar 3.1 Diagram Alir Penelitian

Berikut penjelasan mengenai Gambar 3.1, yaitu:

1. Identifikasi dan Perumusan Masalah

Langkah pertama dalam penelitian melibatkan pemahaman dan penjelasan terhadap permasalahan yang akan dibahas. Ini mencakup topik seperti penerapan nilai *Recency*, *Frequency*, *Monetary* (RFM) dan penilaian efektivitas dari perbandingan dua metode pengelompokan.

2. Studi Literatur

Setelah menentukan topik dan judul penelitian, langkah selanjutnya adalah melakukan studi literatur. Pada tahap ini, peneliti akan mencari informasi dari berbagai sumber seperti jurnal, buku, tesis, situs web, dan referensi lainnya yang terkait dengan permasalahan yang sedang diteliti.

3. Pengumpulan Data

Pengumpulan data dilakukan dengan melakukan survei secara langsung ke Kedai Kenari di Kota Semarang. Survei ini juga melibatkan wawancara singkat dengan Pimpinan Kedai Kenari yaitu Bapak Gabrielle D. Limas. Berdasarkan kesepakatan, data yang diambil adalah data transaksi penjualan yang berisi informasi mengenai pembelian konsumen di Kedai Kenari, Kota Semarang, pada periode Januari 2023 sampai dengan April 2024 dengan 7.445 transaksi. Data Transaksi Penjualan Kedai Kenari tersedia dalam format .csv dan mencakup berbagai informasi, antara lain:

Tabel 3.1 Tabel Data Transaksi Penjualan Kedai Kenari

	<i>Customer ID</i>	<i>Outlet</i>	<i>Date</i>	<i>Time</i>	<i>Disct</i>	...	<i>Items</i>	<i>Payment Method</i>	<i>Event Type</i>	<i>Reason of Refund</i>
0	CS75857	KEDAI KENARI	01-01-2023	13:09:18	0.0	...	Nasi Putih x 3, Kakap Cabe Garam, Cumi ...	BCA	<i>Payment</i>	NaN
1	CS61390	KEDAI KENARI	01-01-2023	16:02:17	0.0	...	Ayam Asam Manis, Nasi Putih, Paket Sapi (L...	BCA	<i>Payment</i>	NaN
2	CS23088	KEDAI KENARI	01-01-2023	16:32:45	0.0	...	Air Mineral, Teh Manis (Dingin), Kul...	<i>Cash</i>	<i>Payment</i>	NaN
3	CS81789	KEDAI KENARI	01-01-2023	18:48:30	0.0	...	Kakap Goreng Tepung, Es Jeruk Selas...	<i>Cash</i>	<i>Payment</i>	NaN
4	CS81789	KEDAI KENARI	01-01-2023	19:15:17	0.0	...	Nasi Goreng Merah, Teh Manis (Dingin), Air Mi...	BCA	<i>Payment</i>	NaN
...
7440	CS25007	KEDAI KENARI	30-04-2024	21:35:45	0.0	...	Sari Kacang Hijau x 2, SARI KACANG HIJAU, So...	<i>Bank Transfer</i>	<i>Payment</i>	NaN
7441	CS13400	KEDAI KENARI	30-04-2024	21:36:27	0.0	...	Sop Jagung (Ayam), Kakap Steam Saos Ke...	<i>Cash</i>	<i>Payment</i>	NaN

	<i>Customer ID</i>	<i>Outlet</i>	<i>Date</i>	<i>Time</i>	<i>Disct</i>	<i>...</i>	<i>Items</i>	<i>Payment Method</i>	<i>Event Type</i>	<i>Reason of Refund</i>
7442	CS26606	KEDAI KENARI	30-04-2024	21:37:33	0.0	...	Ayam Goreng Mentega, Es Campur, Ayam Goren...	<i>Bank Transfer</i>	<i>Payment</i>	NaN
7443	CS23233	KEDAI KENARI	30-04-2024	21:40:19	0.0	...	Jajan Pouch Sedang x 2	<i>Cash</i>	<i>Payment</i>	NaN
7444	CS87407	KEDAI KENARI	30-04-2024	21:56:40	0.0	...	Jagung Telur Asin, Ayam Gohyong, Bakpao Cokla...	<i>BCA QR</i>	<i>Payment</i>	NaN
<i>7445 rows x 22 columns</i>										

4. Preprocessing Data

Preprocessing data merupakan bagian penting dalam *data mining*. Hasil dari *preprocessing* data dapat menentukan hasil *cluster* yang akan dibentuk. Tujuan utama adalah melakukan tahapan *preprocessing* data untuk membersihkan, merapihkan, dan mengubah data agar dapat diolah dengan efisien menggunakan model algoritma dan menghasilkan hasil yang akurat. Beberapa langkah dalam proses *preprocessing* data meliputi:

- a *Data Cleaning*: melakukan proses penghapusan *noise*, nilai yang hilang, dan data yang tidak relevan sebagai langkah untuk menjaga kualitas *dataset*. Sebelum melakukan *data cleaning*, berikut penjelasan data yang akan digunakan dalam penelitian ini pada tabel 3.2.

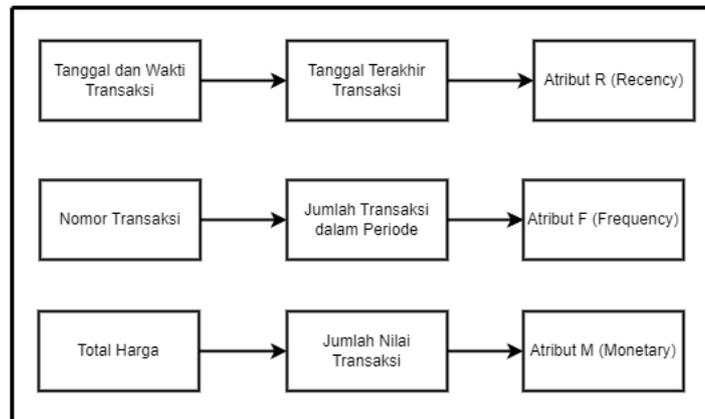
Tabel 3.2 Penjelasan Data

Atribut	Tipe Data	Keterangan
<i>Customer ID</i>	Kategorikal	Kombinasi dari huruf CS dan angka 4–5-digit secara unik diberikan untuk setiap pelanggan.
<i>Outlet</i>	Kategorikal	Menunjukkan Kedai Kenari tempat transaksi dilakukan.
<i>Date</i>	Kategorikal	Tanggal Transaksi
<i>Time</i>	Kategorikal	Waktu Transaksi (WIB).
<i>Gross Sales</i>	Numerik	Harga total barang yang dijual sebelum dikurangi <i>diskon</i> dan <i>refund</i> (IDR).
<i>Discounts</i>	Numerik	Jumlah potongan harga yang diberikan kepada pelanggan.
<i>Net Sales</i>	Numerik	Harga total barang yang dijual setelah dikurangi <i>diskon</i> dan <i>refund</i> (IDR).
<i>Gratuity</i>	Numerik	Jumlah uang tambahan yang diberikan pelanggan sebagai penghargaan kepada pelayanan (IDR).
<i>Tax</i>	Numerikal	Jumlah pajak yang dikenakan atas penjualan barang (IDR).

Atribut	Tipe Data	Keterangan
<i>Total Collected</i>	Numerik	Jumlah uang yang diterima dari pelanggan, termasuk <i>gratuity</i> dan <i>tax</i> (IDR).
<i>Total Amount</i>	Numerik	Jumlah keseluruhan yang harus dibayar pelanggan (IDR).
<i>Other Note (Optional)</i>	Kategorikal	Catatan tambahan mengenai transaksi.
<i>Receipt Number</i>	Kategorikal	Nomor unik yang tertera pada struk belanja.
<i>Collected By</i>	Kategorikal	Nama kasir atau kedai yang menerima pembayaran pelanggan.
<i>Served By</i>	Kategorikal	Nama <i>staff</i> yang melayani pelanggan.
<i>Customer</i>	Kategorikal	Nama <i>customer</i> yang transaksi
<i>Customer Phone</i>	Kategorikal	Nomor telepon customer
<i>Items</i>	Kategorikal	List apa saja menu yang dipesan oleh <i>customer</i>
<i>Payment Method</i>	Kategorikal	Cara pembayaran yang digunakan pelanggan.
<i>Event Type</i>	Kategorikal	Berisi informasi sudah dibayar atau belum
<i>Reason of Refund</i>	Kategorikal	Berisi alasan mengapa barang dikembalikan

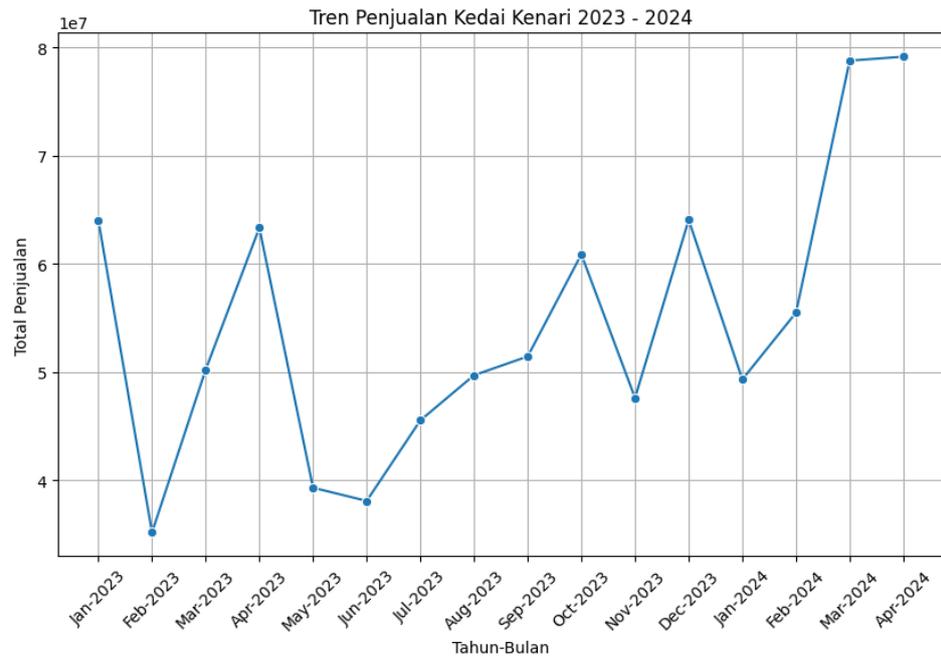
Setelah itu, langkah selanjutnya adalah melihat karakteristik data yang mendalam. Dalam tahap data cleaning, perlu berfokus pada upaya untuk memastikan bahwa tidak ada nilai yang hilang atau kosong dalam *dataset*. Adanya data yang kosong dapat mengganggu kinerja algoritma dan menghambat pencapaian hasil yang optimal. Selain itu, konsistensi data juga harus diperhatikan dengan cermat. Data yang tidak konsisten dapat menghasilkan ketidakjelasan atau ambigu dalam interpretasi data. Oleh karena itu, penting untuk memverifikasi kembali bahwa setiap atribut dalam *dataset* memiliki konsistensi baik dalam hal tipe data maupun nilai yang terkandung dalam atribut tersebut.

- b. *Data Reduction*; suatu proses dalam analisis data yang bertujuan untuk mengurangi jumlah data yang harus dianalisis, tanpa mengorbankan informasi yang penting. Tujuan dari *data reduction* adalah untuk mengurangi *volume* data dengan menghilangkan catatan dan atribut kolom yang tidak diperlukan. Pengurangan kolom dilakukan untuk mengurangi ukuran data agar lebih mudah diolah, khususnya saat melakukan teknik analisis seperti *clustering*. Atribut-atribut yang akan tetap digunakan dalam analisis mencakup *Customer ID*, *Date* dan *Time*, *Receipt Number*, dan *Total Amount*. Hal ini disebabkan karena atribut-atribut tersebut sangat relevan dalam perhitungan dan analisis nilai *Recency*, *Frequency*, *Monetary* (RFM). Atribut tersebut digunakan sebagai berikut:
- *Customer ID*: Digunakan untuk mengidentifikasi pelanggan secara individual dalam analisis RFM.
 - *Date* dan *Time*: Digunakan untuk menghitung nilai *Recency* (seberapa baru pelanggan melakukan pembelian).
 - *Receipt Number*: Digunakan untuk menghitung *Frequency* (seberapa sering pelanggan melakukan pembelian).
 - *Total Amount*: Digunakan untuk menghitung *Monetary* (nilai total pembelian pelanggan).
- c. *Data Transformation*; proses mengubah format, tipe, atau struktur data dari satu bentuk ke bentuk lain. Tujuan dari *data transformation* adalah untuk membuat data lebih sesuai untuk analisis, pemrosesan, atau penggunaan tertentu. Penjelasan lebih lanjut bisa dilihat pada Gambar 3.2.



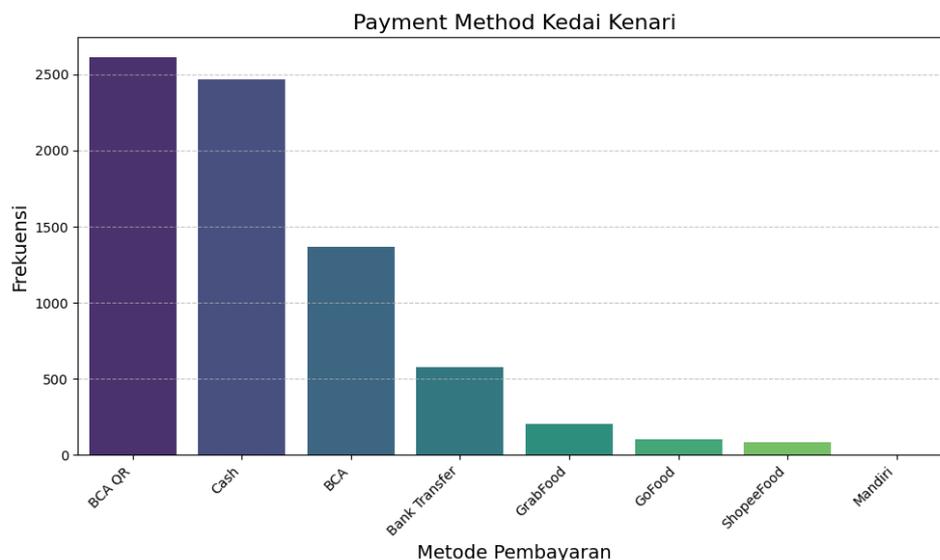
Gambar 3.2 *Data Transformation*

- d. *Exploratory Data Analysis (EDA)*; proses analisis data awal yang digunakan untuk memahami karakteristik, pola, dan hubungan dalam *dataset*. Tujuan utama EDA adalah untuk mendapatkan wawasan awal tentang data dan membantu peneliti atau analis dalam merencanakan analisis selanjutnya. EDA membantu mengidentifikasi pertanyaan yang relevan, menentukan strategi analisis yang tepat, dan memastikan kualitas data sebelum digunakan untuk pengambilan keputusan atau pemodelan. Berikut merupakan hasil EDA dari *dataset* transaksi penjualan Kedai Kenari berikut:



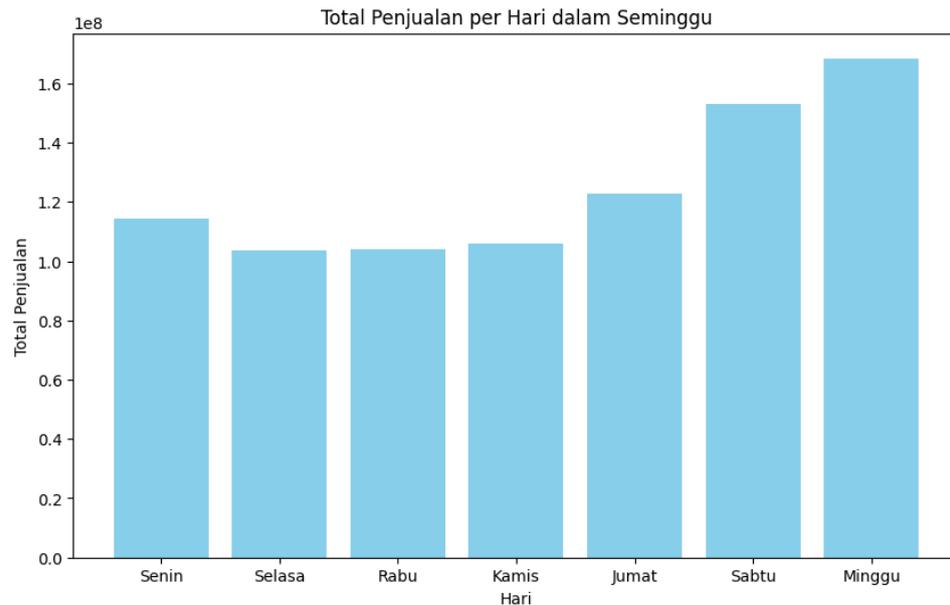
Gambar 3.3 Trend Penjualan Kedai Kenari 2023 - 2024

Pada Gambar 3.3, menunjukkan bahwa penjualan Kedai Kenari mengalami peningkatan yang signifikan di tahun 2024 khususnya pada bulan Maret dibandingkan tahun 2023. Terjadi kenaikan penjualan tahunan sebesar 23,73%, dengan total penjualan di tahun 2024 pada bulan April mencapai Rp79.134.630,00, dibandingkan dengan Rp64.065.350,00 di tahun 2023 pada bulan Desember. Hal ini menunjukkan potensi pasar yang besar di bulan April untuk Kedai Kenari.



Gambar 3.4 Metode Pembayaran Kedai Kenari

Pada Gambar 3.4 menunjukkan bahwa *BCA QR* dan *Cash* menjadi dua metode pembayaran terfavorit di Kedai Kenari, dengan jumlah pengguna yang hampir setara. *BCA QR* sedikit lebih unggul dengan 2.613 transaksi, dibandingkan 2.468 transaksi untuk *Cash*. *Bank Transfer* dan *Kredit BCA* juga digunakan oleh beberapa pelanggan, namun jumlahnya jauh lebih sedikit. Popularitas *BCA QR* dan *Cash* kemungkinan besar didorong oleh kemudahan penggunaan dan kecepatan transaksi.



Gambar 3.5 Total Penjualan Per Hari dalam Seminggu

Pada Gambar 3.5, terdapat fakta pada periode 01 Januari 2023 sampai 30 April 2024 bahwa *weekend* mendominasi dengan Sabtu dan Minggu sebagai hari terlaris, diikuti Jumat dan Senin di *weekdays*. Tren *weekend* yang kuat ini menunjukkan minat pelanggan yang lebih tinggi untuk berbelanja di Kedai Kenari pada akhir pekan. Meskipun *weekdays* memiliki penjualan yang lebih rendah, Jumat dan Senin menunjukkan potensi pasar yang cukup besar.

5. Perhitungan RFM

Dalam analisis RFM menggunakan Bahasa Pemrograman Python terdiri dari tiga kriteria utama yaitu *Recency*, *Frequency*, dan *Monetary*. *Recency*, *Frequency*, dan *Monetary* dapat digunakan untuk mengekstraksi atribut-atribut yang dibutuhkan dalam mengukur loyalitas pelanggan. Data yang telah melalui tahap pengumpulan dan *preprocessing* akan digunakan dalam perhitungan bobot untuk masing-masing kriteria RFM dan data awal ini akan dimanfaatkan dalam analisis RFM pada Tabel 3.3.

Tabel 3.3 Data Awal RFM

	<i>Customer ID</i>	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
0	CS22250	2023-12-31 19:45:06	3BO8QS	458700
1	CS40694	2023-12-31 17:05:20	3BO8QR	93600
2	CS8400	2023-12-31 16:11:45	3BO8QQ	96600
3	CS81117	2023-12-31 15:32:49	3BO8QP	307400
4	CS79899	2023-12-31 14:58:11	3BO8QO	498000

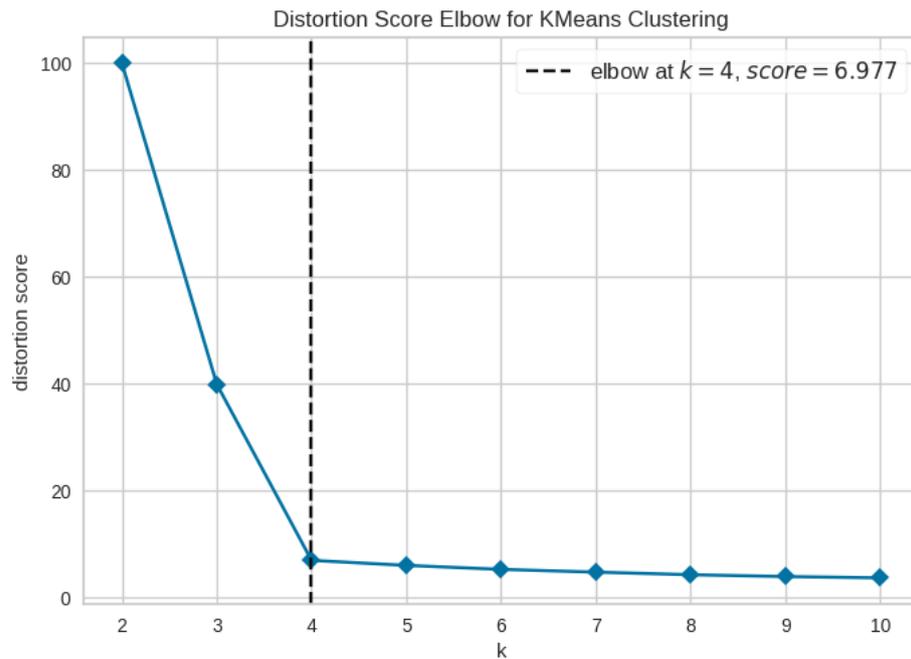
6. Normalisasi Data RFM

Setelah data RFM (*Recency*, *Frequency*, *Monetary*) dikumpulkan, langkah selanjutnya adalah melakukan normalisasi data menggunakan teknik *Min-Max Scaling*. Proses ini bertujuan untuk mentransformasi nilai data ke dalam skala yang sama, sehingga mempermudah perbandingan data antar pelanggan. Data hasil normalisasi ini kemudian digunakan sebagai input untuk dua algoritma *clustering*, yaitu *K-Means* dan *Fuzzy C-Means*.

7. *K-Means*

Setelah mengumpulkan data, memproses data, dan menghitung nilai RFM, langkah berikutnya adalah menerapkan algoritma *K-Means* untuk mengelompokkan pelanggan berdasarkan skor *Recency*, *Frequency*, *Monetary* (RFM) pelanggan. Langkah-langkah perancangan *K-Means* dilakukan melalui serangkaian tahapan sebagai berikut:

- a. Menentukan jumlah *cluster* (*k*) dengan menggunakan *elbow method*. Pada Gambar 3.7, terlihat bahwa jumlah kluster terbaik yang digunakan adalah 4 kluster berdasarkan metode *elbow*.



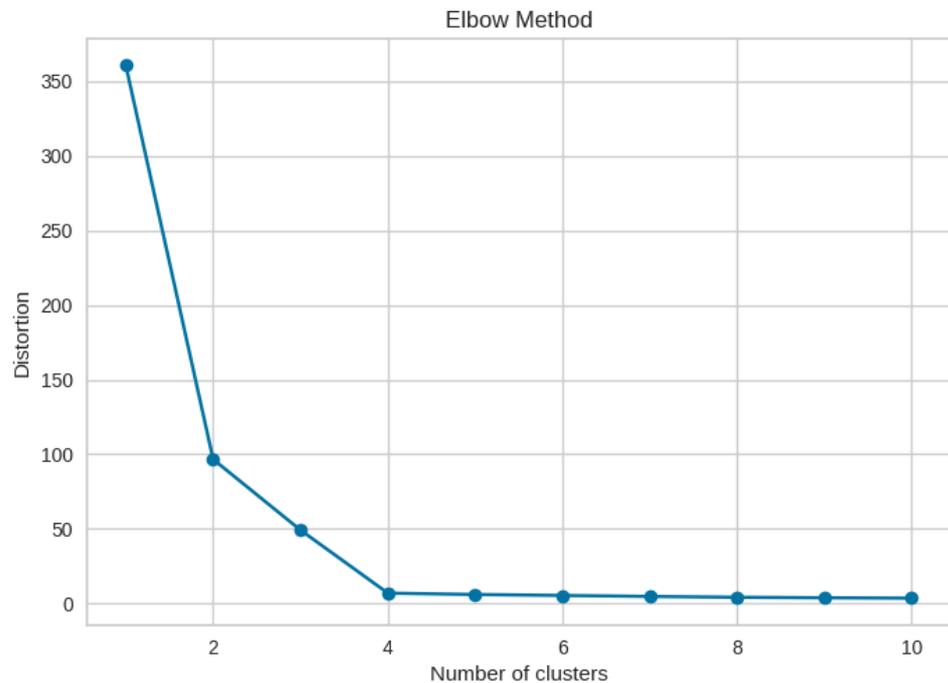
Gambar 3.6 *Elbow Method K-Means*

- b. Inisialisasi pusat cluster secara acak dengan *random state* = 42.
- c. Hitung jarak *centroid* berdasarkan nilai RFM yang sudah di transformasi.
- d. Pengelompokan berdasarkan objek *minimum*.
- e. Selanjutnya, dilakukan perhitungan ulang nilai *centroid* untuk *cluster* yang baru terbentuk dan diperiksa apakah terjadi perubahan atau tidak. Proses ini diulangi hingga tidak ada lagi perubahan pada nilai *centroid*.
- f. Hasil *clustering* akan dievaluasi dengan *Davies-Bouldin Index* (DBI) dan *Silhouette Coefficient*.

8. *Fuzzy C-Means*

Setelah mengumpulkan data, memproses data, dan menghitung nilai RFM, langkah berikutnya adalah menerapkan *Fuzzy C-Means* (FCM) untuk mengelompokkan pelanggan berdasarkan skor *Recency*, *Frequency*, *Monetary* (RFM) mereka. Tahap-tahap perancangan *Fuzzy C-Means* (FCM) dilakukan melalui serangkaian tahapan sebagai berikut:

- a. Jumlah *cluster* yang digunakan adalah 4 *cluster* berdasarkan hasil dari metode *elbow*.



Gambar 3.7 *Elbow Method FCM*

- b. Menentukan n dan T serta *maxiter* (maksimal iterasi).
- c. Inisialisasi matriks *pseido partition* dengan nilai acak.
- d. Hitung *centroid* setiap *cluster*.
- e. Hitung nilai derajat keanggotaan.
- f. Hitung nilai objektif.
- g. Lakukan perhitungan nilai *centroid* dengan mempertimbangkan derajat keanggotaan data selama ketentuan-ketentuan berikut masih terpenuhi, seperti:
- Perubahan dalam nilai fungsi tujuan masih di atas nilai batas yang ditetapkan.
 - Perubahan nilai pusat *cluster* masih melampaui nilai batas yang telah ditentukan.
 - Batas iterasi maksimum belum terpenuhi.
- g. Evaluasi hasil *cluster* akan dilakukan dengan menggunakan *Davies-Bouldin Index (DBI)* dan *Silhouette Coefficient*.

9. Evaluasi dan Kesimpulan

Setelah melalui proses *K-Means* dan *Fuzzy C-Means* (FCM) serta telah menyelesaikan proses evaluasi, analisis hasil pengelompokan dilakukan untuk memahami karakteristik masing-masing kelompok pelanggan. Analisis ini bertujuan untuk mengidentifikasi pola dan perilaku pelanggan dalam setiap kelompok yang kemudian dapat digunakan untuk mendukung pengambilan keputusan bisnis lebih lanjut. Dalam kerangka penelitian ini, evaluasi dilakukan dengan maksud membandingkan kinerja dua algoritma pengelompokan, yakni *K-Means* dan *Fuzzy C-Means* berdasarkan nilai *Recency, Frequency, Monetary* (RFM). Tujuan utama dari evaluasi ini adalah untuk memilih algoritma yang paling cocok dengan persyaratan bisnis, didasarkan pada analisis RFM.