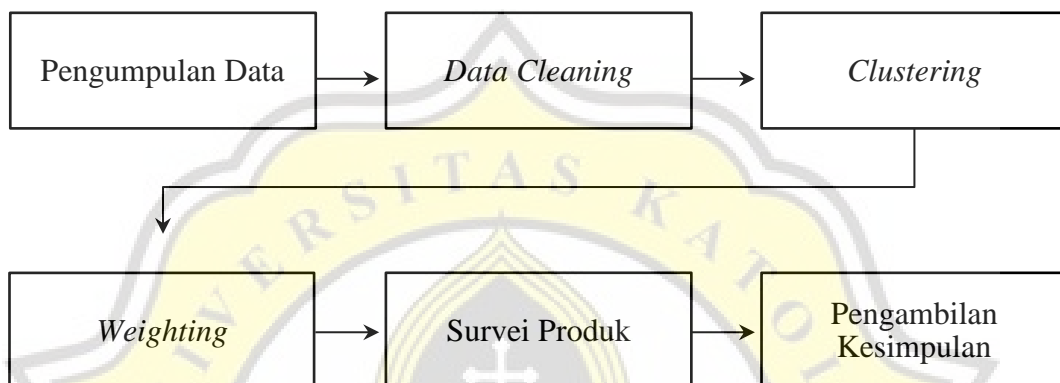


### 3. METODE PENELITIAN

Penelitian ini dilakukan menggunakan pendekatan penelitian survei yang diawali dengan tahap pengumpulan data, *data cleaning*, *clustering*, *weighting*, survei produk, dan pengambilan kesimpulan.



Gambar 1. Diagram Alir Penelitian

#### 3.1. Pengumpulan Data

##### 3.1.1. Media dan Waktu Penelitian

Media yang dipakai untuk melakukan penelitian ini adalah menggunakan *form* atau kuesioner yang disajikan melalui *google form* yang kemudian dibagikan kepada masyarakat berusia produktif berdasarkan kelompok usia melalui beberapa media sosial antara lain Instagram, Facebook, WhatsApp, dan LINE. Penelitian ini dilakukan secara *online* pada bulan Juli 2022 sampai Agustus 2022.

##### 3.1.2. Populasi Penelitian

Populasi dalam penelitian ini adalah masyarakat yang berusia produktif dengan semua jenis kelamin. Banyaknya jumlah populasi masyarakat yang berusia produktif tidak dibatasi oleh cakupan wilayah tertentu. Dalam penelitian ini tidak

dilakukan pembedaan berdasarkan jenis kelamin. Pembagian kelompok usia dapat dilihat pada Tabel 1.

**Tabel 1.** Populasi Kelompok Usia

Kelompok	Usia
A	15 sampai 24 tahun
B	25 sampai 34 tahun
C	35 sampai 44 tahun
D	45 sampai 54 tahun
E	54 sampai 64 tahun

### 3.1.3. Sampel Penelitian

Dalam penelitian ini, jumlah populasinya tidak diketahui. Menurut Daniel (1999), dalam penentuan jumlah sampel dengan jumlah populasi yang tidak diketahui dapat menggunakan rumus sebagai berikut:

$$n = \frac{Z^2 P(1 - P)}{d^2}$$

Di mana:

n = ukuran sampel

Z = statistik Z untuk tingkat kepercayaan,

P = prevalensi atau proporsi yang diharapkan (dalam proporsi satu; jika 20%, P = 0,2),

d = presisi (dalam proporsi satu; jika 5%, d = 0,05).

Hasil survei pendahuluan digunakan untuk menentukan jumlah responden yang akan digunakan dalam penelitian ini. Penentuan jumlah responden menggunakan rumus Daniel.

$$n = \frac{1,96^2 0,57(1-0,57)}{0,05^2} = 376,63 \text{ dibulatkan menjadi } 377 \text{ Responden}$$

Berdasarkan penghitungan jumlah responden menggunakan rumus Daniel, didapatkan hasil akhir yaitu 377 responden. Hasil dari penentuan jumlah responden tersebut ditambahkan nilai *error* sebesar 5% dari hasil perhitungan yaitu 19 (nilai pembulatan dari 18,85). Sehingga didapatkan hasil akhir penentuan jumlah responden yaitu sebanyak 396 orang.

### 3.1.4. Variabel Penelitian

Variabel dalam penelitian ini adalah antara lain *recency*, *frequency*, *monetary value*, kelompok usia, varian rasa, bentuk, dan warna wafer. *Recency* merupakan kapan terakhir kali konsumen membeli/mengonsumsi suatu produk. *Frequency* merupakan seberapa sering konsumen membeli/mengonsumsi suatu produk. *Monetary* merupakan seberapa banyak uang yang dikeluarkan konsumen untuk membeli/mengonsumsi suatu produk.

### 3.1.5. Penyusunan dan Pembagian Kuesioner

Kuesioner berisi tentang identitas diri responden meliputi nama, usia, jenis kelamin, domisili, varian wafer yang disukai dan sudah ada di pasaran, varian wafer yang belum ada di pasaran dan disukai, dan bentuk wafer yang disukai. Responden akan mengisi kuesioner secara mandiri secara *online*. Tabel kuesioner dapat dilihat pada Tabel 2.

**Tabel 2.** Kuesioner Survei

Pertanyaan	Pilihan Jawaban	Jenis Pertanyaan
Nama Lengkap	(tidak ada)	Jawaban Singkat
Jenis Kelamin	Laki-laki Perempuan	Pilihan Ganda
Usia	15 sampai 24 tahun	Pilihan Ganda
	25 sampai 34 tahun	
	35 sampai 44 tahun	
	45 sampai 54 tahun	
	55 sampai 64 tahun	

Pertanyaan	Pilihan Jawaban	Jenis Pertanyaan
Domisili	(tidak ada)	Jawaban Singkat
Kapan anda terakhir kali membeli wafer?	Kurang dari 7 hari yang lalu 7 sampai 14 hari yang lalu Lebih dari 14 hari yang lalu	Pilihan Ganda
Seberapa sering anda membeli wafer dalam sebulan?	Kurang dari 2 kali 2 sampai 5 kali Lebih dari 5 kali	Pilihan Ganda
Berapa uang yang rela anda keluarkan untuk membeli wafer dalam sekali pembelian?	Di bawah 10 ribu 10 ribu sampai 25 ribu Di atas 25 ribu	Pilihan Ganda
Untuk keperluan apa anda membeli wafer?	Dikonsumsi sendiri Dikonsumsi orang lain Dijual lagi	Pilihan Ganda
Rasa wafer yang disukai (boleh lebih dari 1)	(tidak ada)	Jawaban Panjang
Bentuk wafer yang disukai	Persegi Persegi Panjang Persegi dan Persegi Panjang	Pilihan Ganda
Warna krim yang disukai (boleh lebih dari 1)	(tidak ada)	Jawaban Panjang
Warna <i>sheet</i> yang disukai (boleh lebih dari 1)	(tidak ada)	Jawaban Panjang
Varian rasa wafer yang belum ada di pasaran dan diinginkan (boleh lebih dari 1)	(tidak ada)	Jawaban Panjang

### 3.1.6. Uji Kelayakan Kuesioner

Kuesioner yang telah dibuat dan diuji kelayakannya menggunakan uji validitas dan uji reliabilitas. Dalam uji ini, pertanyaan yang validitas dan reliabilitasnya rendah akan direduksi. Responden yang tidak mengisi kuesioner secara lengkap akan direduksi. Reduksi dilakukan sampai validitas dan reliabilitas dapat dipertanggungjawabkan dengan tingkat kepercayaan hingga 95%. Uji validitas diterapkan untuk mengetahui apakah kuesioner yang digunakan untuk mengukur variabel penelitian yang akan diperoleh dari masyarakat dengan usia produktif telah sesuai atau belum.

Dalam pengujian validitas, metode uji yang digunakan dalam penelitian ini adalah Korelasi Pearson. Korelasi Pearson merupakan alat analisis statistik yang digunakan untuk melihat keeratan hubungan linier antara 2 variabel yang skala datanya adalah interval atau rasio (Obilor *et al.*, 2018). Uji reliabilitas bertujuan untuk menemukan dan mengukur pemahaman responden tentang pertanyaan yang tercantum pada kuesioner. Tes reliabilitas dilakukan dengan memasukkan seluruh pertanyaan dari kuesioner. Dalam pengujian reliabilitas, metode uji yang digunakan dalam penelitian ini adalah *Cronbach Alpha*. Pengujian reliabilitas menggunakan uji *Cronbach Alpha* dilakukan untuk instrumen yang mempunyai jawaban benar lebih dari 1 (Adamson & Prion, 2013).

### **3.1.7. Survei Lanjutan**

Setelah hasil survei awal tersebut dapat dipertanggungjawabkan reliabilitas dan validitasnya, maka survei dilanjutkan dengan mencari sisa target respondennya. Kuesioner disebarluaskan kepada masyarakat berusia produktif menggunakan media *google form* melalui berbagai media sosial. Kuesioner disebarluaskan kepada masyarakat sebanyak-banyaknya agar mengantisipasi adanya reduksi pada data setelah dilakukan pengujian validitas dan reliabilitas.

## **3.2. Data Cleaning**

### **3.2.1. Tabulasi Data**

Tabulasi data adalah tahap data hasil kuesioner akan diubah menjadi tabel. Tahap ini diperlukan ketelitian yang tinggi agar tidak terjadi sebuah kesalahan. Tabel hasil tabulasi dapat berbentuk seperti tabel pemindahan, tabel biasa, atau tabel analisis (Hoq *et al.*, 2006).

### **3.2.2. Pembetulan Kesalahan Tipografi**

Dalam pengisian data, terkadang responden tidak mengisi data dengan format pengisian yang diharapkan. Maka dari itu, proses *data cleaning* membetulkan kesalahan tipografi perlu dilakukan. Tahap ini dilakukan untuk mempermudah dalam proses pengolahan data dari responden (Chapman, 2005).

### **3.2.3. Pengolahan Null Value**

*Null value* merupakan data suatu elemen pada tabel yang tidak memiliki nilai apapun. Dalam mengolah elemen ini, terdapat beberapa metode untuk memberi nilai pada *null value* yaitu dengan cara memberikan nilai baru yang diambil dari rata-rata atau modus dari nilai kolom yang bersangkutan (Thalheim & Schewe, 2010). Cara lain untuk mengolah *null value* adalah dengan cara menghilangkan baris yang memiliki nilai *null*.

### **3.2.4. Penghilangan Duplikat**

Sama halnya dengan proses sebelumnya, adanya duplikat data juga akan membuat data yang akan diolah menjadi tidak valid serta dapat menimbulkan kesalahan dalam proses pengolahan data. Menurut Marcus *et al* (2001) data duplikat harus ditandai pada tabulasi sehingga mereka dapat diidentifikasi dan dihapuskan dari analisis karena data duplikat dapat menimbulkan bias. Maka dari itu, duplikat data dari responden perlu untuk dihilangkan.

### **3.2.5. Pemberian Kode Data**

Pemberian kode data adalah sebuah proses pemberian nilai yang umumnya dilakukan pada elemen tabel yang memiliki nilai berupa data kualitatif yang berbentuk huruf, kalimat, maupun rentang. Tujuan dari pemberian kode data atau

*data coding* adalah untuk memberikan tingkatan dan nilai data yang jelas dalam mengolah data serta membentuk model *clustering* pada tahapan selanjutnya.

Pemberian kode data pada penelitian ini menggunakan *label encoding*. *Label encoding* merupakan sebuah metode *encoding* yang memberikan nilai berurutan pada suatu kolom (Zheng & Casari, 2018). *Label encoding* pada penelitian ini dilakukan pada kuesioner dengan pertanyaan yang merujuk pada indikator RFM (*Recency, Frequency, dan Monetary Value*) yaitu : “Kapan Anda terakhir kali membeli wafer ?”, “Seberapa sering Anda membeli wafer dalam sebulan ?”, dan “Berapa jumlah uang yang Anda rela keluarkan untuk membeli wafer ?”

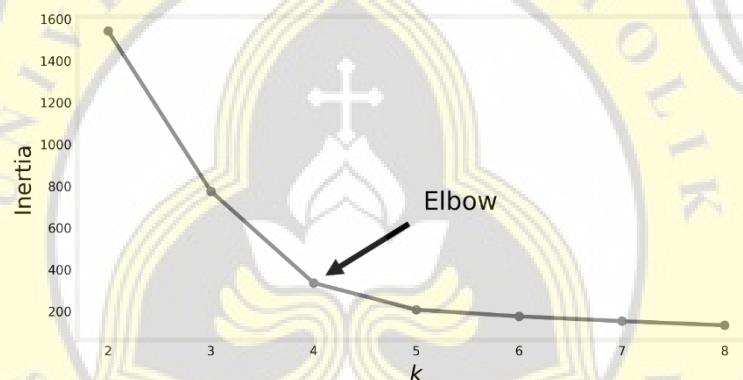
**Tabel 3.** Pemberian Kode Data

Indikator	Pilihan Kuesioner	Hasil Pemberian Data
Kapan anda terakhir kali membeli wafer?	Kurang dari 7 hari yang lalu	2
	7 sampai 14 hari yang lalu	1
	Lebih dari 14 hari yang lalu	0
Seberapa sering anda membeli wafer dalam sebulan?	Kurang dari 2 kali	2
	2 sampai 5 kali	1
	Lebih dari 5 kali	0
Berapa uang yang rela anda keluarkan untuk membeli wafer dalam sekali pembelian?	Di bawah Rp. 10.000	2
	Rp. 10.000 sampai Rp. 25.000	1
	Di atas Rp. 25.000	0

### 3.3. Clustering

#### 3.3.1. Elbow Method

Metode *Elbow* adalah metode visual untuk menguji konsistensi jumlah *cluster* terbaik dengan membandingkan *inertia* atau perbedaan jumlah dari kesalahan kuadrat (SSE) dari setiap jumlah *cluster*, perbedaan yang paling ekstrim yang membentuk sudut siku menunjukkan jumlah *cluster* terbaik (Lopez-Rubio *et al.*, 2018). Sebagai contoh, pada Gambar 2, jumlah *cluster* ( $k$ ) terbaik diperoleh dari pada jumlah cluster 4 karena jumlah *cluster* lainnya tidak memiliki perubahan yang signifikan.



**Gambar 2.** Contoh Grafik *Elbow Method*

#### 3.3.2. K-Means Clustering

*K-Means clustering* adalah algoritma yang diperlukan sebanyak input  $k$  yang membagi  $n$  objek menjadi  $k$  *cluster* sehingga tingkat kesamaan antara anggota dalam satu *cluster* tinggi sedangkan tingkat kemiripan dengan anggota di *cluster* lain sangat rendah (Dubey *et al.*, 2018). Tingkat kesamaan antara anggota dalam *cluster* diukur dengan kedekatan objek dengan nilai rata-rata pada *cluster* atau biasanya disebut sebagai *cluster centroid*. Metode *K-Means* merupakan metode pengelompokan yang paling sederhana dan paling yang umum digunakan (Sammour & Othman, 2016). *K-Means* sering digunakan karena memiliki kemampuan untuk mengelompokkan sejumlah besar data dengan waktu



perhitungan yang relatif cepat dan efisien. Tetapi kelemahan dari metode ini adalah kelemahan dalam menganalisis distribusi data dan tergantung pada inisialisasi sentroid. *K-Means* hanya melihat rentang data untuk setiap sentroid pada setiap *cluster*. Penempatan *cluster* responden dalam penelitian ini dipengaruhi oleh *recency*, *frequency*, dan *monetary* terhadap produk wafer dari responden itu sendiri.

### 3.3.3. *Weighting*

Metode *weighting* yang digunakan dalam penelitian ini adalah metode *Weighted Product*. *Weighted Product* merupakan salah satu metode yang dipakai dalam penyelesaian sistem pengambilan keputusan dengan pertimbangan kriteria dan bobot. Dalam penelitian ini digunakan metode *weighted product*, karena dalam pengambilan keputusan diperlukan beberapa pertimbangan yaitu antara lain kapan terakhir kali konsumen membeli produk tersebut (*recency*), seberapa sering konsumen membeli produk tersebut (*frequency*), dan berapa uang yang dikeluarkan konsumen tersebut untuk membeli produk tersebut (*monetary*). Pada metode ini, *frequency* memiliki bobot nilai yang lebih tinggi, diikuti oleh *monetary* dan kemudian *recency* (Monalisa *et al.*, 2019). Dalam penelitian ini, *recency*, *frequency*, dan *monetary* memiliki masing-masing 3 tingkatan yang terdiri dari tingkatan yang terendah sampai yang tertinggi.

## 3.4. Survei Produk

Survei produk wafer yang dijual di pasaran dilakukan di 2 *supermarket* (Transmart dan Farmers Market), 2 *minimarket* (Indomaret dan Alfamart), dan 2 pasar tradisional (Pasar Johar dan Pasar Peterongan) yang dilakukan di bulan Agustus 2022. Atribut varian wafer yang disurvei antara lain rasa, warna *sheet* dan krim, serta bentuk wafer. Survei produk ini bertujuan untuk mengetahui produk wafer apa saja yang dijual di pasaran yang kemudian hasilnya akan dihubungkan dengan hasil survei kuesioner oleh responden yang telah dilakukan.