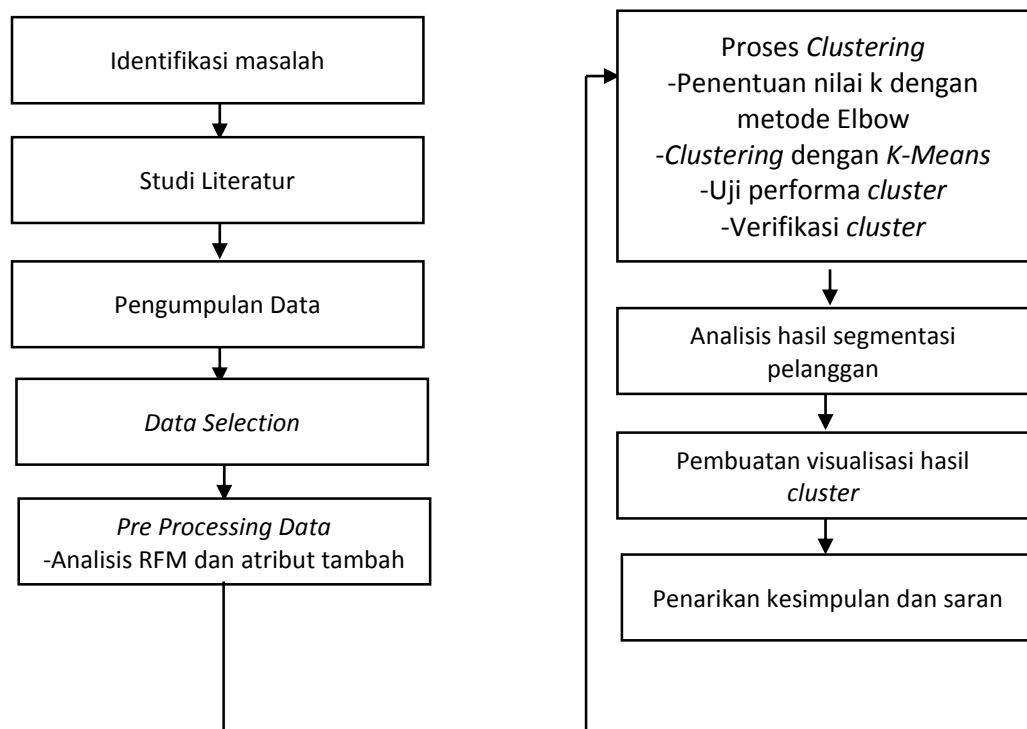


BAB 3 METODOLOGI

3.1 Metodologi Penelitian

Alur dari penelitian segmentasi pelanggan dengan metode *K-Means Clustering* berdasarkan model RFM pada Klinik Kecantikan *Belle Crown Malang* adalah sebagai berikut :



Gambar 3.1 Metodologi Penelitian

3.1.1 Identifikasi Masalah

Mengidentifikasi permasalahan terkait objek penelitian yakni pada Klinik Kecantikan Belle Crown Malang. Masalah tersebut terkait segmentasi pelanggan.

3.1.2 Studi Literatur

Adanya studi literatur adalah untuk mencari teori – teori yang relevan terhadap penelitian sejenis yang pernah dilakukan sebelumnya yang berkaitan dengan penemuan sebelumnya. Karena penelitian bertujuan untuk mendapatkan solusi dari permasalahan yang sedang dihadapi.

3.1.3 Pengumpulan Data

Dilakukan dengan teknik wawancara untuk menggali informasi dari pihak terkait. Selain wawancara juga dilakukan ekstraksi data dari *database* dengan format *file .xls*. Jumlah data yang didapatkan adalah 21.513 transaksi dan 4716 pelanggan. Hasil dari ekstraksi ini digunakan untuk mengidentifikasi atribut yang akan digunakan serta jenis data yang diperlukan dalam melakukan segmentasi pelanggan, yaitu :

1. Data profil pelanggan klinik kecantikan *Belle Crown* yang bersifat *anonymous* atau bukan data *credential* pasien seperti atribut ID dan nama (dikarenakan aturan kode etik).
2. Data transaksi pelanggan 6 bulan terakhir tahun 2017.
3. Data yang mendukung model RFM untuk proses *clustering*.
4. Data pendukung lainnya yang berkaitan.

3.1.4 Data Selection

Pengolahan data yang terlalu banyak menyebabkan proses *clustering* yang tidak optimal. Oleh karena itu perlu dilakukan pemilihan variabel data yang berkaitan dengan proses *clustering* dan analisis RFM. Kemudian data diekstraksi ke dalam variabel RFM sebagai *input* dalam proses *clustering*. Proses *data selection* pada data untuk memilih variabel preferensi RFM. Variabel *Recency* diperoleh dengan perhitungan selisih terakhir kali transaksi pelanggan dengan tanggal penelitian yakni 07 November 2017. Variabel *Frequency* dihitung dengan menjumlahkan transaksi pelanggan dalam satu periode. Sedangkan variabel *Monetary* dihitung dengan menjumlahkan besaran rupiah yang dikeluarkan dalam transaksi. Kemudian data diubah ke dalam format *.csv* untuk dilakukan *pre processing data*.

Atribut dasar terdiri dari *recency*, *frequency*, dan *monetary* (RFM). Hasil penelitian menunjukkan bahwa sejumlah atribut ini telah diterapkan oleh berbagai industri dan analisis dalam melakukan segmentasi pelanggan (Aggelis dkk., 2005; Chen dkk.,2005, Cheng dkk.,2009). Setelah data diekstraksi kemudian data diubah ke dalam format *.csv* untuk proses *pre processing data*.

3.1.5 Pre Processing Data

Pre processing data terdiri dari *data cleansing*, reduksi *outlier* dengan metode yang dilakukan Klodian Dhana, dan transformasi data yang termasuk di dalamnya menggunakan metode *Logarithmic* untuk menyeragamkan data ke dalam nilai yg lebih kecil, selanjutnya dilakukan normalisasi Min-Max, dimana nilai dari semua variabel diseragamkan menjadi 0 hingga 1 agar tidak terdapat data yang tidak normal. Selanjutnya data diintegrasikan ke dalam satu tabel. Seluruh rangkaian proses *pre processing* dilakukan menggunakan *software R*.

3.1.6 Proses Clustering

Data pelanggan yang telah melalui proses *pre processing data* kemudian dikelompokkan dengan salah satu metode *clustering*, yaitu metode *K-Means Clustering*. Pada proses *clustering* terdapat tiga tahapan yaitu menentukan nilai *k* dengan metode *Elbow*, proses *clustering* dan uji performa *cluster*. Masukan dari *clustering* ini adalah data dengan variabel RFM. Sedangkan keluarannya adalah *cluster* optimum. Seluruh rangkaian proses *clustering* dilakukan menggunakan *software R*.

3.1.6.1 Menentukan Nilai k dengan Metode Elbow

Penentuan jumlah *cluster* dilakukan dengan menggunakan metode *Elbow* seperti yang dilakukan pada penelitian terdahulu. Penentuan nilai *k* ini dilakukan secara *intuitive* dengan melihat penurunan garis pada grafik. Jika grafik terlihat mengalami penurunan secara drastis diikuti penurunan yang stabil, maka nilai tersebut menjadi acuan sebagai nilai *k*. Setelah itu dilakukan proses *clustering* menggunakan metode *K-Means*.

3.1.6.2 Proses Clustering Menggunakan K-Means

Proses *clustering* menggunakan metode *K-Means Clustering* dengan hasil *cluster* yang mencerminkan karakteristik pelanggannya.

3.1.6.3 Uji Performa Cluster

Uji performa *cluster* dilakukan untuk membandingkan hasil metode *Elbow* dengan nilai SSE yang diperoleh. Selain itu juga dilakukan uji performa menggunakan *package CValid()* yang di dalamnya memuat tiga metode, yakni *Dunn Index* yang menggambarkan bahwa suatu *cluster* semakin berbeda dengan *cluster* yang lain, *Silhouette Width* yang digunakan untuk

menentukan kebenaran penempatan suatu *cluster*, dan *Connectivity* yang digunakan untuk mengukur hubungan antar anggota dalam satu *cluster*.

3.1.7 Analisis Hasil *Cluster*

Selanjutnya data dilakukan denormalisasi guna mempermudah analisis *cluster* dengan membandingkan variable layanan. Fokus analisis adalah pada pencarian *cluster* optimal.

3.1.8 Pembuatan Visualisasi Hasil *Cluster*

Visualisasi *cluster* ditampilkan dalam bentuk *dashboard* menggunakan *software* Power BI. Masukan dalam proses ini yaitu *script* hasil eksekusi pada *software* R dengan data yang telah didenormalisasi untuk mempermudah pembuatan grafik. Keluarannya berupa *dashboard* yang merepresentasikan hasil segmentasi pelanggan yang telah diverifikasi oleh pemilik Belle Crown Malang melalui kuisisioner *System Usability Scale* (SUS). Hasil kuisisioner dilampirkan pada Lampiran.

3.1.9 Penarikan Kesimpulan dan Saran

Tahap terakhir yaitu penarikan kesimpulan berdasarkan rumusan masalah terkait segmentasi pelanggan. Kemudian pemberian saran untuk penelitian di masa mendatang.