



TUGAS AKHIR - KS141501

**IMPLEMENTASI FUZZY C-MEANS DAN MODEL
RFM UNTUK SEGMENTASI PELANGGAN (STUDI
KASUS : PT. XYZ)**

***IMPLEMENTATION OF FUZZY C-MEANS AND
RFM MODEL FOR CUSTOMER SEGMENTATION
(CASE STUDY : PT. XYZ)***

Denny Bintang Saputra
NRP 5214100135

Dosen Pembimbing :
Edwin Riksakomara, S.Kom, M.T.

DEPARTEMEN SISTEM INFORMASI
Fakultas Teknologi Informasi dan Komunikasi
Institut Teknologi Sepuluh Nopember
Surabaya 2018



TUGAS AKHIR - KS141501

IMPLEMENTASI FUZZY C-MEANS DAN MODEL RFM UNTUK SEGMENTASI PELANGGAN (STUDI KASUS : PT. XYZ)

Denny Bintang Saputra
NRP 5214100135

Dosen Pembimbing :
Edwin Riksakomara, S.Kom, M.T.

DEPARTEMEN SISTEM INFORMASI
Fakultas Teknologi Informasi dan Komunikasi
Institut Teknologi Sepuluh Nopember
Surabaya 2018





FINAL PROJECT - KS141501

IMPLEMENTATION OF FUZZY C-MEANS AND RFM MODEL FOR CUSTOMER SEGMENTATION (CASE STUDY : PT. XYZ)

Denny Bintang Saputra
NRP 5214100135

Supervisor :
Edwin Riksakomara, S.Kom, M.T.

DEPARTMENT OF INFORMATION SYSTEMS
Faculty of Information and Communication Technology
Sepuluh Nopember Institute of Technology
Surabaya 2018

LEMBAR PENGESAHAN

IMPLEMENTASI FUZZY C-MEANS DAN MODEL RFM UNTUK SEGMENTASI PELANGGAN (STUDI KASUS : PT. XYZ)

TUGAS AKHIR

Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada
Departemen Sistem Informasi
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember

Oleh :

DENNY BINTANG SAPUTRA
5214 100 135

Disetujui Tim Pengaji : Tanggal Ujian : 16 Januari 2018
Periode Wisuda : Maret 2018

Edwin Riksakomara, S.Kom, M.T.

(Pembimbing I)

Wiwik Anggraeni, S.Si, M.Kom.
(Pengaji I)

Ahmad Mukhlason, S.Kom, M.Sc, Ph.D. (Pengaji II)

LEMBAR PENGESAHAN

IMPLEMENTASI FUZZY C-MEANS DAN MODEL RFM UNTUK SEGMENTASI PELANGGAN (STUDI KASUS : PT. XYZ)

TUGAS AKHIR

Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada
Departemen Sistem Informasi
Fakultas Teknologi Informasi dan Komunikasi
Institut Teknologi Sepuluh Nopember

Oleh :

DENNY BINTANG SAPUTRA
5214 100 135

Surabaya, Januari 2018

**PLH Kepala
Departemen Sistem Informasi**



Edwin Riksakomara, S.Kom, M.T.
NIP 19690725 20031210 01

IMPLEMENTASI FUZZY C-MEANS DAN MODEL RFM UNTUK SEGMENTASI PELANGGAN (STUDI KASUS : PT. XYZ)

Nama Mahasiswa : Denny Bintang Saputra
NRP : 5214100135
Departemen : Sistem Informasi
Pembimbing I : Edwin Riksakomara, S.Kom, M.T.

ABSTRAK

Era globalisasi yang disertai dengan berlakunya MEA menyebabkan persaingan dalam berbagai sektor bisnis, termasuk sektor industri. Hal ini memaksa perusahaan-perusahaan yang bergerak di sektor industri di Indonesia untuk bersaing dengan negara-negara tetangga di ASEAN, termasuk PT. XYZ. Banyaknya perusahaan kompetitor yang menawarkan harga rendah membuat pelanggan PT. XYZ beralih pada kompetitor. Terkait dengan hal tersebut, PT. XYZ telah melakukan berbagai kegiatan promosi dalam rangka mengelola hubungan dengan pelanggan. Namun kegiatan promosi selama ini diimplementasikan kepada semua pelanggan tanpa adanya segmentasi sehingga strategi promosi menjadi tidak efektif dan efisien.

Untuk menyelesaikan masalah pelanggan tersebut, salah satu solusi yang dapat dilakukan adalah menentukan segmentasi pelanggan. Segmentasi yang dilakukan menggunakan metode clustering Fuzzy C-Means. Sedangkan untuk menentukan jumlah K, metode yang digunakan adalah metode Elbow. Dalam proses clustering, diterapkan variabel RFM yang menggambarkan perilaku pelanggan dalam melakukan transaksi. Setelah hasil clustering muncul, cluster tersebut akan

divisualisasikan sehingga memudahkan untuk menganalisa setiap cluster.

Hasil dari clustering dengan Fuzzy C-Means dan model RFM adalah tiga segmen pelanggan. Segmen 1 adalah segmen terburuk dengan rata-rata nilai CLV 0.0538. Segmen 2 adalah segmen menengah dengan rata-rata nilai CLV 0.1674. Segmen 3 adalah segmen terbaik dengan rata-rata nilai CLV 0.3516. Segmen 1 memiliki 42 anggota, segmen 2 memiliki 64 anggota, dan segmen 3 memiliki 17 anggota. Penelitian ini menghasilkan visualisasi berbasis web yang menyajikan grafik hasil clustering.

Kata kunci : Clustering, Fuzzy C-Means, Model RFM, Segmentasi Pelanggan

IMPLEMENTATION OF FUZZY C-MEANS AND RFM MODEL FOR CUSTOMER SEGMENTATION (CASE STUDY : PT. XYZ)

Student Name	: Denny Bintang Saputra
NRP	: 5214100135
Department	: Sistem Informasi
Supervisor I	: Edwin Riksakomara, S.Kom, M.T.

ABSTRACT

Globalization era and Asean Economic Community caused competitions in many business sectors, including industry sector. This condition forced many companies in industry sector in Indonesia to compete with other ASEAN countries, including PT. XYZ. Several company competitors that offer low price of product make the customers of PT. XYZ move to the competitors. Actually, PT XYZ had applied some marketing strategies to manage relationship with customers. But, these marketing strategies were implemented to all customers without any segmentations, so the marketing strategies are not effective and efficient.

One of the solution to solve that customer problem is customer segmentation analysis. This segmentation is constructed using Fuzzy C-Means clustering method. Meanwhile, Elbow method is used to determine the K value of clustering. During the clustering process, the variables that used refer to RFM model describing behavior of customers when transactions occurred. After clustering result constructed, the clusters will be visualized to make cluster analysis easier.

The clustering process with Fuzzy C-Means and RFM model resulted three customer segments. Segment 1 is the worst segment which has 0.0538 CLV average. Segment 2 is the medium segment which has 0.1674 CLV average. Segment 3 is the best segment which has 0.0538 CLV average. Segment 1 has 42 members, segment 2 has 64 members, and segment 3 has 17 members. This research resulted web-based visualization that presents graphics of clustering results.

Keywords : Clustering, Fuzzy C-Means, RFM Model, Customer Segmentation

KATA PENGANTAR

Alhamdulillah penulis panjatkan kepada Allah SWT atas rahmat, nikmat, barakah, dan kemudahan yang telah diberikan sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul : **IMPLEMENTASI FUZZY C-MEANS DAN MODEL RFM UNTUK SEGMENTASI PELANGGAN (STUDI KASUS : PT. XYZ)**, sebagai salah satu syarat kelulusan di Departemen Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi, Institut Teknologi Sepuluh Nopember Surabaya.

Penyelesaian Tugas Akhir ini tidak akan tercapai tanpa bantuan dan dukungan dari banyak pihak yang sudah mendedikasikan waktu, tenaga, pikiran, dan materi untuk penulis selama pengerjaan Tugas Akhir. Oleh karena itu, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada berbagai pihak, yaitu :

1. Allah SWT yang telah menganugerahkan rahmat dan barakahnya sehingga Tugas Akhir ini dapat selesai.
2. Kedua orang tua, kakak, adik, dan seluruh keluarga yang senantiasa memberikan dukungan materi, semangat dan doa untuk penulis.
3. Pihak PT. XYZ Surabaya, yang telah memberikan kesempatan bagi penulis untuk melaksanakan penelitian Tugas Akhir.
4. Bapak Dr. Ir. Aris Tjahyanto, M.Kom, selaku Kepala Departemen Sistem Informasi FTIK ITS, yang telah menyediakan dukungan untuk penelitian mahasiswa.
5. Bapak Edwin Riksakomara, S.Kom., M.T. selaku dosen pembimbing yang telah meluangkan waktunya untuk

memberi masukan, arahan dan bimbingan untuk penelitian Tugas Akhir ini.

6. Bapak Nisfu Asrul Sani selaku dosen wali yang telah memberikan dorongan, perwalian dan semangat sejak awal perkuliahan hingga penggerjaan Tugas Akhir.
7. Seluruh Bapak dan Ibu dosen pengajar beserta karyawan Departemen Sistem Informasi, FTIK ITS Surabaya yang telah memberikan ilmu, pengalaman berharga dan bantuan kepada penulis selama masa perkuliahan.
8. Mas Ricky selaku admin laboratoriun RDIB yang membantu penulis dalam hal administrasi penyelesaian Tugas Akhir.
9. Rekan-rekan pengurus HMSI, BEM FTIf, BIMITS, dan organisasi lain yang telah menjadi tempat penulis dalam mengembangkan *softskill*.
10. Rekan-rekan pejuang wisuda 117 laboratorium RDIB yang memberikan bantuan dan semangat.
11. Teman-teman OSIRIS yang menjadi keluarga tempat berbagi susah dan senang selama penulis menempuh perkuliahan.
12. Berbagai pihak yang telah turut serta membantu yang tidak dapat disebut satu persatu.

Penulis sangat menyadari bahwa penulisan Tugas Akhir ini masih belum sempurna serta memiliki banyak kekurangan Untuk itu penulis mengharapkan saran, masukan, dan kritik yang membangun untuk perbaikan selanjutnya. Penulis juga berharap Tugas Akhir ini bermanfaat bagi berbagai pihak.

Surabaya, Januari 2018

Penulis

DAFTAR ISI

ABSTRAK	i
ABSTRACT	iii
KATA PENGANTAR	v
DAFTAR ISI	vii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xiii
DAFTAR SCRIPT	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Tugas Akhir	5
1.5 Manfaat Tugas Akhir	5
1.6 Relevansi Tugas Akhir	5
1.7 Sistematika Penulisan Tugas Akhir.....	6
BAB II TINJAUAN PUSTAKA.....	9
2.1 Studi Sebelumnya.....	9
2.2 Dasar Teori	11
2.2.1 PT. XYZ Cabang Surabaya	12
2.2.2 Customer Relationship Management	13
2.2.3 Segmentasi Pelanggan.....	14
2.2.4 Penggalian Data.....	16
2.2.5 Clustering	17
2.2.6 Fuzzy C-Means (FCM)	19
2.2.7 Model RFM.....	21
2.2.8 Normalisasi <i>Min-Max</i>	22
2.2.9 Metode <i>Elbow</i>	23
2.2.10 Analytical Hierarchy Process	23
2.2.11 Customer Lifetime Value (CLV)	26
2.2.12 Sum Square Error (SSE)	27

BAB III METODOLOGI PENELITIAN	29
3.1 Metodologi	29
3.1.1 Identifikasi Masalah	30
3.1.2 Studi Literatur.....	31
3.1.3 Pengumpulan Data.....	31
3.1.4 Praproses Data	32
3.1.5 Pembobotan RFM.....	33
3.1.6 Perhitungan CLV	33
3.1.7 Clustering Data.....	34
3.1.8 Pengujian Kualitas <i>Cluster</i>	34
3.1.9 Visualisasi <i>Cluster</i>	35
3.1.10 Analisis Segmentasi.....	35
3.1.11 Penyusunan Laporan Tugas Akhir.....	35
BAB IV PERANCANGAN.....	37
4.1 Pengumpulan data	37
4.2 Praproses data.....	38
4.3 Pembobotan RFM.....	39
4.4 Penentuan jumlah <i>cluster</i>	40
4.5 Rancangan visualisasi.....	40
4.5.1 Use Case Diagram	40
4.5.2 <i>Use Case Description</i>	41
4.5.3 Desain Antarmuka	42
BAB V IMPLEMENTASI	45
5.1 Pra Proses Data.....	45
5.1.1 Data Cleaning	45
5.1.2 Pemodelan RFM	47
5.1.3 Normalisasi.....	48
5.2 Pembobotan RFM.....	50
5.3 Perhitungan CLV	51
5.4 Proses Clustering	51
5.4.1 Penentuan Jumlah K	51
5.4.2 Clustering dengan Fuzzy C-Means.....	52
5.5 Pengujian Kualitas <i>Clustering</i>	53

5.6	Pembuatan Visualisasi.....	54
5.6.1	Halaman Utama.....	54
5.6.2	Halaman Visualisasi Segmen	56
5.6.3	Halaman Tabel	63
BAB VI	HASIL PEMBAHASAN	67
6.1	Pembobotan RFM dengan AHP	67
6.2	Perhitungan CLV.....	70
6.3	Proses <i>Clustering</i>	70
6.3.1	Penentuan Nilai K	71
6.3.2	<i>Clustering Fuzzy C-Means</i>	72
6.4	Pengujian Kualitas <i>Clustering</i>	76
6.4.1	Validasi Derajat Keanggotaan.....	76
6.4.2	Pengujian SSE.....	77
6.5	Analisis Hasil Clustering.....	77
6.5.1	Pemeringkatan Segmen	78
6.5.2	Analisis Karakteristik Segmen 1	79
6.5.3	Analisis Karakteristik Segmen 2	80
6.5.4	Analisis Karakteristik Segmen 3	81
6.5.5	Perbandingan Karakteristik Segmen	82
6.6	Analisis Visualisasi Clustering.....	82
6.6.1	Grafik <i>Bar Chart</i> dan <i>Pie Chart</i>	83
6.6.2	Grafik <i>Scatter Plot 3D</i>	84
6.6.3	Grafik <i>Box Plot</i>	87
6.7	Kesimpulan Analisis	92
BAB VII	KESIMPULAN DAN SARAN	95
7.1	Kesimpulan	95
7.2	Saran.....	96
DAFTAR PUSTAKA	97
BIODATA PENULIS	102
LAMPIRAN A	A-1
LAMPIRAN B	B-1
LAMPIRAN C	C-1
LAMPIRAN D	D-1

LAMPIRAN E	E-1
LAMPIRAN F	F-1
LAMPIRAN G	G-1
LAMPIRAN H	H-1

DAFTAR TABEL

Tabel 2.1 Penelitian Sebelumnya Judul Paper 1	9
Tabel 2.2 Penelitian Sebelumnya Judul Paper 2	10
Tabel 2.3 Penelitian Sebelumnya Judul Paper 3	10
Tabel 2.4 Penelitian Sebelumnya Judul Paper 4	11
Tabel 2.5 Skala Preferensi Penilaian Pairwise Comparison...	24
Tabel 2.6 Nilai Random Indeks.....	26
Tabel 4.1 Data Mentah Transaksi Penjualan.....	37
Tabel 4.2 Data Mentah Tabel Pelanggan	38
Tabel 4.3 Use case melihat grafik bar chart dan pie chart.....	42
Tabel 5.1 Atribut Terpilih	45
Tabel 5.2 Baris Data Tidak Valid.....	46
Tabel 5.3 Sebagian Hasil Pemodelan RFM.....	48
Tabel 5.4 Rentang Nilai Atribut	49
Tabel 6.1 Hasil Kuesioner Responden 1	67
Tabel 6.2 Hasil Kuesioner Responden 2	67
Tabel 6.3 Hasil Kuesioner Responden 3	67
Tabel 6.4 Hasil Rata-Rata Penilaian Kuesioner	68
Tabel 6.5 Normal Comparisons Matrix.....	68
Tabel 6.6 Overall Score Bobot	68
Tabel 6.7 Sebagian Hasil Perhitungan CLV Pelanggan.....	70
Tabel 6.8 Selisih Penurunan Nilai Error	71
Tabel 6.9 Percobaan Terhadap Perubahan Nilai Parameter ...	73
Tabel 6.10 Sebagian Hasil Validasi Derajat Keanggotaan....	77
Tabel 6.11 Peringkat Segmen Berdasarkan CLV	78
Tabel 6.12 Karakteristik Pelanggan pada Segmen 1	79
Tabel 6.13 Karakteristik Pelanggan pada Segmen 2	80
Tabel 6.14 Karakteristik Pelanggan pada Segmen 3	81
Tabel 6.15 Perbandingan Karakteristik 3 Segmen	82
Tabel B.1 Use Case Melihat Grafik Scatter Plot 3D	B-1
Tabel B.2 Use Case Melihat Grafik Box-Plot.....	B-2

Tabel B.3 Use Case Melihat Rentang Nilai Segmen.....	B-3
Tabel B.4 Use Case Melihat Pelanggan Teratas dan Terbawah	B-4
Tabel B.5 Use Case Melihat Rataan Variabel RFM.....	B-5
Tabel B.6 Use Case Melihat Peringkat Pelanggan Setiap Cluster	B-6
Tabel B.7 Use Case Melihat Tabel Data	B-7
Tabel C.1 Pemodelan RFM Keseluruhan Pelanggan	C-1
Tabel D.1 Hasil Keseluruhan Normalisasi Variabel RFM ..	D-1
Tabel E.1 Hasil Keseluruhan Perhitungan CLV dan Clustering	E-1
Tabel F.1 Keseluruhan Hasil Percobaan Perubahan Nilai Parameter.....	F-1
Tabel G.1 Keseluruhan Validasi Derajat Keanggotaan Pelanggan	G-1

DAFTAR GAMBAR

Gambar 3.1 Alur Metodologi Penelitian	29
Gambar 4.1 Use Case Diagram	41
Gambar 4.2 Halaman Utama Dashboard.....	43
Gambar 4.3 Halaman Visualisasi Segmen	44
Gambar 4.4 Halaman Tabel Data.....	44
Gambar 5.1 Hasil Eksekusi Visualisasi Rata-Rata Variabel RFM	58
Gambar 5.2 Hasil Eksekusi Visualisasi Tabel Peringkat Pelanggan	61
Gambar 5.3 Hasil Eksekusi Visualisasi Tabel Data	65
Gambar 6.1 Plot Penentuan Nilai K dengan Metode Elbow ..	72
Gambar 6.2 Hasil clustering Fuzzy C-Means untuk 3 cluster	74
Gambar 6.3 Hasil Pengujian SSE.....	78
Gambar 6.4 Grafik Pie Chart 3 Segmen.....	83
Gambar 6.5 Grafik Bar Chart 3 Segmen	83
Gambar 6.6 Grafik Scatter Plot 3D Keseluruhan Segmen	84
Gambar 6.7 Grafik Scatter Plot 3D Segmen 1	85
Gambar 6.8 Grafik Scatter Plot 3D Segmen 2	86
Gambar 6.9 Grafik Scatter Plot 3D Segmen 3	87
Gambar 6.10 Grafik Box Plot CLV Semua Segmen	88
Gambar 6.11 Grafik Box Plot Recency Semua Segmen	89
Gambar 6.12 Grafik Box Plot Frequency Semua Segmen	90
Gambar 6.13 Grafik Box Plot Monetary Semua Segmen	92

DAFTAR SCRIPT

Script 5.1 Query Ekstraksi Data	48
Script 5.2 Normalisasi Data.....	49
Script 5.3 Penentuan Jumlah K.....	52
Script 5.4 Proses Clustering	53
Script 5.5 Perhitungan SSE	53
Script 5.6 Visualisasi Bar Chart	55
Script 5.7 Visualisasi Pie Chart.....	56
Script 5.8 Visualisasi rata-rata variabel RFM	57
Script 5.9 Visualisasi Pelanggan Teratas dan Terbawah.....	58
Script 5.10 Visualisasi Grafik Scatter Plot 3D	59
Script 5.11 Visualisasi Tabel Peringkat Pelanggan	60
Script 5.12 Visualisasi Grafik Box-Plot (1).....	62
Script 5.13 Visualisasi Grafik Box-Plot (2).....	63
Script 5.14 Visualisasi Tabel Data	64

BAB I

PENDAHULUAN

Pada bab ini, akan dijelaskan tentang Latar Belakang, Rumusan Masalah, Batasan Masalah, Tujuan Tugas Akhir, Manfaat, Relevansi, dan Sistematika Penulisan Tugas Akhir.

1.1 Latar Belakang

Era globalisasi yang beberapa tahun lalu dimulai, menyebabkan perubahan besar bagi banyak negara. Perubahan besar terjadi pada berbagai sektor menyebabkan persaingan antar negara semakin ketat. Terlebih setelah disepakatinya MEA pada awal tahun 2016, menyebabkan dampak buruk pada berbagai sektor, termasuk sektor industri [1]. Hal ini memaksa perusahaan-perusahaan yang bergerak di sektor industri di Indonesia untuk bersaing dengan negara-negara tetangga di ASEAN. Salah satu industri yang saat ini tengah bersaing adalah industri pertambangan dan minyak bumi. Kekayaan tambang yang dimiliki Indonesia tak serta merta membuat persaingan lebih mudah. Industri pertambangan dan penggalian mencatat penurunan pertumbuhan di tengah mengilapnya perekonomian nasional [2]. Badan Pusat Statistik menyebut, sektor pertambangan dan penggalian mengalami kontraksi negatif 0,49 persen bahkan di saat pertumbuhan ekonomi yang relatif membaik pada kuartal I tahun 2017. Banyak faktor yang menyebabkan melemahnya industri pertambangan, salah satunya adalah turunnya harga mineral dunia yang rendah.

PT. XYZ merupakan cabang perusahaan yang bergerak di industri tambang dan minyak bumi. PT. XYZ menangani penjualan dan pemasaran produk tambang di wilayah Jawa Timur, Bali, dan Nusa Tenggara. Berbagai produk olahan

tambang yang dipasarkan diantaranya adalah grup produk jenis bitumen (aspal) dan *special chemical*. Seperti halnya perumbuhan industri tambang yang melemah di kuartal I tahun 2017, PT. XYZ juga mengalami hal serupa. Selama bulan Januari hingga Juni 2017, angka penjualan produk tersebut tidak mencapai target yang ditetapkan setiap bulannya [3]. Adapun tidak tercapainya target penjualan disebabkan berbagai faktor. Faktor pertama adalah rendahnya permintaan produk di berbagai wilayah dibanding periode yang sama tahun lalu. Selain itu, harga jual produk lebih rendah yang ditawarkan oleh kompetitor menyebabkan sebagian pelanggan beralih kepada kompetitor. Kompetitor yang bersaing tidak hanya dari dalam negeri tapi juga luar negeri.

Terkait dengan berpindahnya pelanggan ke pihak kompetitor, PT. XYZ telah melakukan berbagai kegiatan promosi untuk menarik kembali pelanggan [3]. Kegiatan promosi tersebut antara lain, program diskon berjenjang pada pembelian produk. Program diskon ini diberlakukan pada setiap pelanggan yang membeli produk dengan jumlah tertentu. Selain itu, dilakukan *door-to-door marketing* dengan cara melakukan pendekatan ke pelanggan di lingkungan masing-masing. Pendekatan kegiatan promosi tersebut ditargetkan kepada semua pelanggan tanpa kriteria tertentu. PT. XYZ menerapkan strategi marketing kepada pelanggan tanpa adanya segmentasi pelanggan. Oleh karena itu, PT. XYZ menerapkan strategi yang sama untuk semua pelanggan. Penerapan strategi yang sama ini dapat menimbulkan beberapa dampak, yaitu strategi promosi yang tidak tepat sasaran sehingga strategi tersebut hanya bermanfaat pada pelanggan tertentu saja. Selain itu, waktu dan dana yang digunakan untuk promosi menjadi kurang efektif dan efisien.

Untuk menyelesaikan masalah pelanggan tersebut, salah satu solusi yang dapat dilakukan adalah menentukan segmentasi pelanggan. Segmentasi pelanggan adalah proses menghasilkan segmen-semen pelanggan berdasarkan karakteristik tertentu sebagai dasar pengelompokan. Setiap pelanggan yang masuk ke dalam segmen tertentu, memiliki kemiripan karakteristik dengan pelanggan dalam satu segmen serta memiliki karakteristik yang berbeda dengan pelanggan pada segmen lainnya. Dari setiap segmen yang dihasilkan dapat diketahui perilaku pelanggan yang dominan karakteristiknya. Dengan hasil segmentasi ini, pihak manajemen dapat menganalisis mengenai strategi marketing yang tepat terhadap setiap segmen berdasarkan kecenderungan perilaku pelanggan di dalamnya.

Banyak metode yang dapat digunakan dalam melakukan segmentasi pelanggan, salah satunya adalah penggalian data terhadap riwayat transaksi pembelian pelanggan pada periode tertentu. Data transaksi tersebut kemudian ditentukan variabel-variabel yang berpengaruh dan dapat dijadikan kriteria untuk menentukan segmen pelanggan. Beberapa variabel yang umumnya digunakan antara lain :

- a. Selisih jarak dari waktu transaksi terakhir kali dilakukan dengan waktu saat ini
- b. Total jumlah transaksi yang dilakukan selama periode tertentu
- c. Jumlah nilai uang dari transaksi yang dilakukan

Ketiga variabel tersebut lazim disebut sebagai model RFM (*Recency*, *Frequency*, dan *Monetary*). Dalam model RFM, *Recency* mewakili waktu pembelian terakhir, *Frequency* mewakili jumlah transaksi, dan *Monetary* mewakili jumlah pemasukan yang diterima. Ketiga variabel tersebut kemudian diproses dengan menggunakan algoritma penggalian data *clustering*. *Clustering* adalah mengelompokkan kumpulan

objek tertentu berdasarkan karakteristiknya, mengumpulkan berdasarkan kesamaan tertentu [15].

Algoritma *clustering* yang digunakan adalah *Fuzzy C-Means*. *Fuzzy C-Means* merupakan salah satu teknik untuk menentukan *cluster* optimal pada suatu ruang vektor yang didasarkan pada bentuk normal *Euclidean* untuk jarak antar vektor. *Fuzzy C-Means* menggunakan derajat keanggotaan sebagai dasar penentuan suatu vektor masuk ke dalam *cluster* tertentu. Untuk mengecek validitas hasil *clustering*, digunakan *Sum of Square Error (SSE)* sebagai metode uji. Semakin kecil nilai SSE, maka semakin bagus hasil *clustering* yang dilakukan.

Dengan hasil penelitian ini, diharapkan dapat memberikan pendukung dalam menentukan strategi pemasaran terhadap pelanggan, dengan bantuan visualisasi berbasis *web* sehingga mempermudah analisis.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, berikut adalah rumusan masalah yang menjadi permasalahan utama dalam tugas akhir ini :

1. Bagaimana implementasi metode *Fuzzy C-Means* pada segmentasi pelanggan berdasarkan model RFM?
2. Bagaimana visualisasi dapat mempermudah analisis segmentasi pelanggan?

1.3 Batasan Masalah

Dalam tugas akhir ini, terdapat beberapa batasan masalah penelitian, yaitu :

1. Data yang akan diolah yaitu data transaksi penjualan produk PT. XYZ pada rentang waktu 1 Januari 2017 hingga 30 Juni 2017.

2. Variabel yang digunakan sebagai parameter adalah jarak waktu pembelian terakhir, frekuensi pembelian, dan jumlah biaya pembelian yang mengacu pada metode RFM.
3. Perangkat lunak yang digunakan sebagai alat bantu *clustering* adalah R Studio
4. Visualisasi *cluster* yang digunakan berbasis *web*

1.4 Tujuan Tugas Akhir

Dalam tugas akhir ini, tujuan yang ingin dicapai adalah :

1. Menghasilkan segmentasi pelanggan dengan metode *clustering Fuzzy C-Means*
2. Membuat visualisasi *clustering* sehingga mempermudah analisis segmentasi

1.5 Manfaat Tugas Akhir

Manfaat penelitian yang dilakukan adalah membantu PT. XYZ dalam menentukan strategi pemasaran dalam manajemen hubungan pelanggan yang tepat berdasarkan segmentasi yang telah dibuat.

1.6 Relevansi Tugas Akhir

Relevansi tugas akhir ini terhadap ruang lingkup Sistem Informasi berada pada Laboratorium Rekayasa Data dan Intelijensi Bisnis dengan topik penggalian data dan *clustering*. Penelitian ini merupakan implementasi dari berbagai mata kuliah yaitu : Sistem Cerdas, Penggalian Data dan Analitika Bisnis, dan Manajemen Rantai Pasok dan Hubungan Pelanggan.

1.7 Sistematika Penulisan Tugas Akhir

Pada penulisan tugas akhir, terdapat sistematika penulisan yang digunakan. Sistematika ini terbagi menjadi tujuh bab yang akan dijabarkan sebagai berikut :

BAB I PENDAHULUAN

Pada bab pendahuluan ini akan dijelaskan mengenai latar belakang, masalah yang akan diselesaikan, rumusan permasalahan, batasan tugas akhir, tujuan dan manfaat yang ingin dicapai, relevansi dan sistematika penyusunan dari topik yang diambil.

BAB II TINJAUAN PUSTAKA

Pada bab tinjauan pustaka ini akan dijelaskan mengenai referensi-referensi yang terkait dan penjelasan teori-teori yang mengenai *clustering*, algoritma *Fuzzy C-Means*, model RFM, penggalian data, manajemen hubungan pelanggan, Metode *Elbow*, Normalisasi *Min-Max*, SSE, AHP, dan perhitungan CLV

BAB III METODOLOGI PENELITIAN

Pada bab metodologi penelitian ini akan dijelaskan mengenai tahapan-tahapan yang dilakukan pada tugas akhir ini, meliputi yaitu identifikasi masalah, studi literatur, pengumpulan data, praproses data, pembobotan RFM, perhitungan CLV, *clustering* data, pengujian kualitas *clustering*, pembuatan visualisasi, analisa hasil segmentasi, dan penyusunan laporan. Semua tahapan akan dideskripsikan dan digambarkan dalam jadwal penggerjaan tugas akhir pada tabel waktu.

BAB IV PERANCANGAN

Pada bab perancangan ini akan dijelaskan mengenai rancangan dari tugas akhir yang terdiri dari proses pengumpulan data, praproses data, pembobotan RFM, penentuan jumlah *cluster* dan rancangan visualisasi.

BAB V IMPLEMENTASI

Pada bab implementasi ini akan dijelaskan mengenai implementasi berdasarkan dari rancangan yang telah dibuat yaitu pra proses data, pembobotan RFM, perhitungan CLV, proses clustering, pengujian kualitas *cluster*, dan proses pembuatan visualisasi.

BAB VI HASIL DAN PEMBAHASAN

Pada bab hasil dan pembahasan ini akan dijelaskan mengenai hasil dan pembahasan dari pembobotan RFM, perhitungan CLV, proses *clustering*, pengujian kualitas *cluster*, analisis hasil *clustering*, analisis visualisasi *cluster*, dan penarikan kesimpulan.

BAB VII PENUTUP

Pada bab penutup ini akan dijelaskan mengenai kesimpulan yang didapat dari penelitian tugas akhir dan saran untuk penelitian selanjutnya.

Halaman ini sengaja dikosongkan

BAB II

TINJAUAN PUSTAKA

Pada bab ini, akan dijelaskan tentang referensi yang terkait dengan penelitian tugas akhir dan teori yang digunakan dalam tugas akhir ini.

2.1 Studi Sebelumnya

Penelitian-penelitian sebelumnya dapat menjadi referensi dan dasar penelitian tugas akhir ini. Berikut adalah beberapa penelitian terdahulu yang digunakan sebagai dasar penelitian pada tabel 2.1 hingga 2.4 :

Tabel 2.1 Penelitian Sebelumnya Judul Paper 1

Judul Penelitian 1	<i>Online Auction Customer Segmentation Using A Neural Network Model [4]</i>
Penulis; Tahun	Chu-Chai Henry Chan; 2005
Deskripsi Umum	Segmentasi pelanggan lelang daring yang bertujuan untuk menggambarkan perilaku pelanggan yang melakukan lelang daring dengan studi kasus laman eBay Taiwan. Segmentasi yang dilakukan menggunakan model SOM <i>Neural Network</i> dengan 3 segmen hasil. Tiga segmen tersebut adalah <i>Impulsive Deal</i> , <i>Patient Deal</i> dan <i>Analytic Deal</i> . Selain itu, atibut yang digunakan untuk mendeteksi perilaku pelanggan mengacu pada model RFM. Hasilnya, 39,3% pelanggan termasuk <i>impulsive deals</i> , 27,8% pelanggan termasuk <i>analytic deals</i> , dan 32,2% pelanggan termasuk <i>patient deals</i> ,
Keterkaitan dengan Tugas Akhir	Referensi penggunaan metode <i>cluster</i> dan model RFM untuk segmentasi pelanggan.

Tabel 2.2 Penelitian Sebelumnya Judul Paper 2

Judul Penelitian 2	<i>Customer Lifetime Value (CLV) Measurement Based on RFM Model [28]</i>
Penulis; Tahun	Babak Sohrabi, Amir Khanlari; 2007
Deskripsi Umum	Penelitian ini menekankan pada pengukuran <i>Customer Lifetime Value (CLV)</i> terhadap pelanggan pada sebuah cabang bank Iranian dengan tujuan mengetahui berapa keuntungan yang diberikan pelanggan pada perusahaan. Data yang digunakan adalah hasil segmentasi 214 data dengan metode <i>K-Means</i> dengan dasar model RFM. Selain itu, penelitian ini menggunakan <i>discriminant analysis</i> untuk menentukan bagaimana <i>cluster</i> dapat digunakan untuk membedakan sampel pelanggan. Hasilnya, 99.5% kasus grup dapat diklasifikasikan dengan benar.
Keterkaitan dengan Tugas Akhir	Referensi dalam perhitungan CLV menggunakan model RFM terhadap hasil <i>clustering</i> data untuk mengetahui keuntungan yang diperoleh dalam sklus hidup pelanggan.

Tabel 2.3 Penelitian Sebelumnya Judul Paper 3

Judul Penelitian 3	<i>Fuzzy C-Means Clustering With Prior Biological Knowledge [29]</i>
Penulis; Tahun	Luis Tari, Chitta Baral, Seungchan Kim; 2008
Deskripsi Umum	Implementasi metode <i>Fuzzy C-Means</i> terhadap data ekspresi gen dengan tujuan untuk mengelompokkan gen dengan fungsi yang berkaitan. Metode <i>clustering</i> yang digunakan adalah <i>Fuzzy C-Means</i> yang dikombinasikan dengan pengetahuan biologis, yaitu <i>Gen Ontology Annotation (GO)</i> sebagai domain <i>cluster</i> yang akan dibentuk. Oleh karena itu, metode <i>clustering</i> ini disebut sebagai <i>semi-</i>

	<i>supervised clustering.</i> Hasilnya, modifikasi metode <i>Fuzzy C-Means</i> diklaim mampu membentuk <i>cluster</i> yang konsisten dengan menentukan <i>cluster</i> awal dengan <i>prior knowledge</i> .
Keterkaitan dengan Tugas Akhir	Referensi penggunaan metode <i>Fuzzy C-Means</i> dimodifikasi untuk <i>clustering</i> .

Tabel 2.4 Penelitian Sebelumnya Judul Paper 4

Judul Penelitian 4	<i>Medical Image Segmentation Using Fuzzy C-Means for MRI Images [34]</i>
Penulis; Tahun	Sonakshi Gupta, Parminder Kaur; 2015
Deskripsi Umum	Penggunaan metode <i>Fuzzy C-Means</i> untuk segmentasi gambar medis yang bertujuan untuk mendekripsi tumor. Dengan penggunaan metode <i>Fuzzy C-Means</i> ini, waktu komputasi dan segmentasi menjadi lebih efisien. Segmentasi dilakukan berdasarkan <i>Fuzzy Code Division</i> . Hasilnya, waktu pemrosesan gambar menjadi lebih cepat jika menggunakan dua kombinasi metode tersebut.
Keterkaitan dengan Tugas Akhir	Referensi penggunaan <i>Fuzzy C-Means</i> sebagai metode untuk segmentasi.

2.2 Dasar Teori

Dasar teori akan memberikan gambaran secara umum dari landasan penjabaran tugas akhir ini. Pada bagian ini menjelaskan mengenai teori-teori yang berkaitan dengan penelitian tugas akhir ini.

2.2.1 PT. XYZ Cabang Surabaya

PT. XYZ merupakan perusahaan yang bergerak pada sektor industri pertambangan dan pengolahan minyak bumi dan gas yang berdiri sejak 1950. PT. XYZ sebagai Badan Usaha Milik Negara bertanggung jawab dalam pengelolaan tambang-tambang migas dari mulai proses hulu hingga ke hilir. PT. XYZ melaksanakan pengusahaan migas mulai dari mengelola dan menghasilkan migas dari ladang-ladang minyak di seluruh wilayah Indonesia, mengolahnya menjadi berbagai produk dan menyediakan serta melayani kebutuhan bahan bakar minyak dan gas di seluruh Indonesia.

PT. XYZ mendirikan cabang di Surabaya yang bertugas untuk memasarkan produk dan melayani pendistribusian produk di wilayah Jawa Timur, Bali, dan Nusa Tenggara. Salah satu divisi di PT. XYZ cabang Surabaya yang memasarkan produk olahan tambang adalah divisi *Petrochemical Trading*. Divisi *Petrochemical Trading* adalah divisi yang bergerak di bidang pemasaran dan pendistribusian produk *petrochemical* yaitu bitumen (aspal) dan *special chemical*. Dari kedua produk ini akan dibagi menjadi beberapa material yang dijual yakni:

- a. Bitumen (Aspal), meliputi ;
 1. Aspal Curah
 2. Aspal Drum
- b. Special Chemical, meliputi;
 1. Green Coke
 2. Minarex (A,H,I)
 3. Paraffinic (60,95)
 4. Slack Wax (LMO, DAO, SPO, MMO)
 5. Parafin Wax (HSWW,YBW,Palm Wax)
 6. Smooth Fluid (02,05)
 7. Tenac Sticker
 8. TB 192

9. Solphy
10. Solvent dan Kondensat
11. Lube Base Oil

2.2.2 Customer Relationship Management

Customer Relationship Management (CRM) adalah sebuah strategi perusahaan untuk memahami dan mempengaruhi perilaku pelanggan melalui komunikasi intensif dengan tujuan untuk meningkatkan pendapatan, penyimpanan, loyalitas, dan profitabilitas dari pelanggan [20]. Proses CRM menyediakan struktur untuk menciptakan hubungan yang sukses dengan pelanggan.

CRM merupakan komitmen dari perusahaan untuk menempatkan pengalaman pelanggan pada pusat dari prioritas perusahaan dan untuk memastikan sistem bahwa yang insentif, proses, dan sumber daya informasi dapat mempengaruhi hubungan dengan menambah pengalaman [21]. Dalam konteks teknologi, CRM merupakan desain komunikasi dan penggunaan informasi untuk memastikan bahwa pelanggan tumbuh lebih yakin, percaya, dan rasa nilai pribadi dalam hubungannya dengan perusahaan.

Fokus utama dalam strategi CRM adalah untuk membuat organisasi menciptakan dan mempertahankan keuntungan pelanggan [20]. Sebagian besar strategi dikembangkan dari tiga aspek, yaitu *customer profitability*, *customer acquisition*, dan *customer retention*, berdasarkan pengurangan biaya mempertahankan pelanggan daripada memperoleh pelanggan baru [22].

Membangun hubungan dengan pelanggan membutuhkan data dari pelanggan. Jika data akan digunakan, data tersebut harus bersih dan tepat pada waktunya, dan kesan diperoleh bahwa

perusahaan memiliki data ekstensif tentang pelanggan [23]. Dalam penggunaan teknologi, sejumlah aplikasi teknologi dapat diidentifikasi yang digunakan pada pengembangan strategi CRM. Tiga komponen utama komponen sistem CRM adalah [24]:

a. *Operational CRM*

Operational CRM meliputi aplikasi yang diakses pelanggan, seperti otomatisasi penjualan, otomatisasi pemasaran perusahaan, dan dukungan dan layanan pelanggan. Pusat bantuan pelanggan juga termasuk komponen dari *operational CRM*, dan telah diidentifikasi sebagai aspek dominan pada sistem CRM.

b. *Analytical CRM*

Analytical CRM menganalisa data yang telah dibentuk melalui *operational CRM* untuk membangun gambaran pelanggan. *Analytical CRM* meliputi menangkap, menyimpan, mengekstrasi, memproses, menginterpretasi dan melaporkan data pelanggan yang disimpan di *data warehouse*.

c. *Collaborative CRM*

Collaborative CRM menggunakan teknologi komunikasi baru dan tradisional untuk memungkinkan pelanggan berinteraksi dengan organisasi. *Collaborative CRM* menyediakan tingkatan yang lebih baik dari respon ke kebutuhan pelanggan dengan mengembangkan semua anggota dari rantai pasok seperti pemasok atau yang lain.

2.2.3 Segmentasi Pelanggan

Segmentasi adalah proses dimana data dikelompokkan menjadi komponen yang tidak tumpang tindih serta memiliki nilai [36]. Sedangkan dalam konteks pelanggan, segmentasi pelanggan adalah proses membagi keseluruhan pasar menjadi kelompok

yang memiliki nilai dengan kemiripan kebutuhan dan keinginan produk yang sama [25]. Segmentasi dapat dipandang sebagai kompleksitas yang berhubungan dengan pelanggan secara individual, masing-masing dengan kebutuhan dan nilai potensial yang berbeda [26]. Perusahaan mengimplementasikan segmentasi pelanggan berdasarkan fakta bahwa setiap pelanggan berbeda dan usaha pemasaran akan terlayani dengan baik jika terdapat target yang spesifik, kelompok kecil dengan maksud bahwa konsumen tersebut akan relevan dan mendorong mereka untuk membeli sesuatu. Perusahaan juga berharap untuk memperoleh pemahaman lebih mendalam dari preferensi pelanggan mereka dan menggunakan strategi untuk masing-masing segmen.

Segmentasi pelanggan menghasilkan segmen-segmen berdasarkan karakteristik tertentu. Karakteristik yang dapat menjadi acuan antara lain [27]:

- a. Demographics : umur, jenis kelamin, besarnya keluarga, besarnya tempat tinggal, siklus kehidupan keluarga, pemasukan, pekerjaan, pendidikan kepemilikan rumah, status sosial ekonomi, agama, kewarganegaraan.
- b. Psychographics : kepribadian, gaya hidup, nilai-nilai, sikap
- c. Behavior : manfaat yang dicari, status pembelian, tingkat penggunaan produk, frekuensi pembelian
- d. Geographics : negara, provinsi, kota, kode pos, iklim

Dengan menerapkan segmentasi pelanggan untuk target pelanggan spesifik, model segmentasi memperkenankan alokasi efektif untuk sumber daya pemasaran dan memaksimalkan *cross-selling* dan *up-selling*. Segmentasi

pelanggan dapat juga meningkatkan layanan pelanggan dan membantu loyalitas dan pendapatan. Keuntungan lain dari segmentasi pelanggan adalah menjadikan perusahaan selangkah lebih maju dari kompetitor pada bagian yang spesifik dari pasar dan mengidentifikasi produk baru yang ada atau pelanggan potensial dapat tertarik atau meningkatkan kualitas produk untuk mencapai keinginan pelanggan.

2.2.4 Penggalian Data

Penggalian data adalah teknologi yang memadukan metode analisis data tradisional dengan algoritma canggih untuk memproses volume data yang besar [17]. Penggalian data didefinisikan sebagai proses menemukan pola-pola dalam data. Proses ini dapat bersifat otomatis atau semi otomatis. Pola yang ditemukan harus memiliki manfaat, semisal manfaat ekonomis. Penggalian data sering disebut juga *knowledge discovery in database* (KDD), yaitu kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam himpunan data berukuran besar.

Teknik dalam penggalian data terbagi atas dua kategori utama, yaitu prediktif dan deskriptif. Kategori prediktif bertujuan untuk memprediksi nilai dari atribut tertentu berdasarkan nilai pada atribut lain. Sedangkan kategori deskriptif bertujuan untuk menurunkan pola-pola (korelasi, tren, *cluster*, trayektori, dan anomali). Beberapa tujuan dari penggalian data adalah [18] :

1. *Explanatory*

Untuk menjelaskan beberapa kondisi penelitian, seperti mengapa penjualan truk pick-up di Colorado meningkat.

2. *Confirmatory*

Untuk mempertegas hipotesa, contohnya hipotesa mengenai dua kali pendapatan keluarga lebih suka

dipakai untuk membeli peralatan keluarga dibandingkan dengan satu kali pendapatan keluarga.

3. *Exploratory*

Untuk menganalisa data yang memiliki hubungan yang baru. Misalnya pola apa yang cocok untuk kasus penggelapan aktru kredit.

Tugas dalam penggalian data terbagi menjadi dua kategori, yaitu langsung (*directed*) dan tidak langsung (*undirected*) [19]. Penggalian data langsung bertujuan untuk menggunakan data yang tersedia untuk membangun sebuah model yang menggambarkan salah satu variabel tertentu yang menarik dari data tersebut. Adapun tugas yang termasuk penggalian data langsung yaitu *classification*, *estimation*, dan *prediction*.

Penggalian data tidak langsung berarti tidak terdapat variabel terpilih sebagai target karena tujuannya untuk membangun beberapa hubungan pada semua variabel. Adapun tugas yang termasuk penggalian data tidak langsung adalah *clustering*, *affinity group* atau *association rules*, *description*, dan *visualization*. Dalam penelitian ini, tugas yang digunakan adalah *clustering* karena tidak variabel terpilih sebagai target dan bertujuan untuk membentuk pola suatu kelompok.

2.2.5 *Clustering*

Clustering dalam konteks penggalian data adalah mengelompokkan kumpulan objek tertentu berdasarkan karakteristiknya, mengumpulkan berdasarkan kesamaan tertentu [15]. *Clustering* atau segmentasi ini adalah satu metode penggalian data yang *unsupervised* atau tidak terawasi, karena tidak ada atribut yang digunakan sebagai panduan dan tidak adanya label pada data dalam proses pembelajaran.

Algoritma *clustering* membangun sebuah model dengan melakukan serangkaian pengulangan dan berhenti ketika model tersebut sudah terpusat dan batasan segmentasi lebih stabil. Hasil *clustering* yang baik ditentukan oleh ukuran kesamaan dan metode yang digunakan. Pendekatan dalam *clustering* dibagi menjadi dua kelompok utama yaitu [16]:

1. Metode Hirarki, yaitu metode yang membentuk *cluster* yang membagi partisi secara berulang-ulang dari atas ke bawah atau sebaliknya. Hasil dari metode hirarki berupa dendrogram yang mewakili kelompok objek dan tingkat kesamaan di mana terdapat perubahan pengelompokkan.
2. Metode Partisi, yaitu metode yang membuat inisial partisi untuk membentuk. Kemudian secara iteratif menggunakan teknik relokasi dengan mencoba berulang-ulang memindahkan objek dari satu kelompok ke kelompok lain untuk memperoleh partisi optimal.

Terdapat beberapa algoritma *clustering* yang dapat diaplikasikan pada himpunan data untuk mempartisi informasi. Beberapa algoritma tersebut berdasarkan karakteristik himpunan data, antara lain [15]:

1. Centroid-based
Pada tipe ini, setiap *cluster* mereferensi sebuah nilai vektor. Masing-masing objek merupakan bagian dari kluster yang memiliki jarak terpendek disbanding *cluster* lain. Jumlah *cluster* harus ditentukan sebelumnya.
2. Distributed-based
Tipe *cluster* ini berhubungan dengan model statistik, metodologi terdistribusi mengkombinasikan objek yang bernilai sama dengan distribusi.
3. Connectivity-based
Pada tipe ini, setiap objek dihubungkan pada tetangganya, tergantung derajat keanggotaan pada jarak

diantara keduanya. Dengan asumsi ini, cluster terbentuk dengan objek berdekatan dan dapat dideskripsikan sebagai batas jarak maksimum.

4. Density-based

Tipe *cluster* ini membentuk *cluster* berdasarkan kepadatan anggota pada himpunan data pada lokasi tertentu. Algoritma ini mengumpulkan perkiraan jarak pada standar level kepadatan anggota grup pada *cluster*.

Pada penelitian ini, metode yang digunakan adalah metode partisi guna mengelompokkan objek hanya pada satu *cluster* saja.

2.2.6 Fuzzy C-Means (FCM)

Fuzzy C-Means (FCM) pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981. FCM merupakan salah satu teknik untuk menentukan *cluster* optimal pada suatu ruang vektor yang didasarkan pada bentuk normal *Euclidean* untuk jarak antar vektor. FCM akan mengcluster data dimana posisi tiap-tiap titik data dalam suatu *cluster* ditentukan oleh derajat keanggotaan.

Fase pertama dalam FCM adalah menentukan pusat *cluster*, yang akan menandai posisi rata-rata untuk tiap-tiap *cluster*. Pada awalnya, posisi pusat *cluster* ini belum sepenuhnya akurat. Setiap titik data memiliki derajat keanggotaan untuk setiap *cluster*, sehingga tergabung pada *cluster* tertentu. Secara berulang, pusat *cluster* akan dihitung ulang, sehingga posisi pusat *cluster* secara bertahap akan berulang. Begitu juga dengan derajat keanggotaan setiap titik akan bergerak menuju posisi yang tepat.

a. Input Data

Input data yang akan *dicluster* X, berupa matriks berukuran $n \times m$, dengan $n =$ jumlah sampel data, $m =$

atribut setiap data). X_{ij} = data sampel ke- i ($i = 1,2,3,\dots,n$), atribut j ($j = 1,2,3,\dots,m$).

b. Menentukan nilai variabel

- Jumlah *cluster* $= c;$
- Pangkat/pembobot $= w;$
- Maksimum Iterasi $= MaxIter;$
- Error terkecil yang diharapkan $= \xi;$
- Fungsi objektif awal $= P_0 = 0;$
- Iterasi awal $= t = 1;$

c. Membangkitkan Nilai Random

Bangkitkan nilai random μ_{ik} , $i = 1,2,3,\dots,n$; $k = 1,2,3,\dots,c$ sebagai elemen – elemen matriks partisi awal μ_{ik} . μ_{ik} adalah derajat keanggotaan yang merujuk pada seberapa besar kemungkinan suatu data bisa menjadi anggota ke dalam suatu *cluster*. Hitung jumlah setiap kolom (atribut) :

$$Q_j = \sum_{k=1}^c \mu_{ik}, j = 1,2,3, \dots, m \quad (1)$$

Hitung :

$$\mu_{ik} = \frac{\mu_{ik}}{Q_j} \quad (2)$$

d. Menghitung Pusat *Cluster* ke- k

Hitung pusat cluste ke- k : V_{kj} dengan $k = 1,2,3,\dots,c$; $j = 1,2,3,\dots,m$.

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (3)$$

e. Menghitung Fungsi Objektif

Hitung fungsi objektif pada iteasi ke- t , P_t :

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (u_{ik})^w \right) \quad (4)$$

f. Menghitung Perubahan Matriks

Hitung perubahan matriks partisi :

$$u_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}; i = 1, 2, 3, \dots, n; k = 1, 2, 3, \dots, c \quad (5)$$

- g. Mengecek Kondisi Berhenti
- Jika ($|P_t - P_{t-1}| < \xi$) atau ($t > MaxIter$) maka berhenti;
 - Jika tidak, $t = t + 1$, ulangi langkah d.

2.2.7 Model RFM

Untuk mengidentifikasi kebiasaan pelanggan, metode yang umum digunakan adalah *recency, frequency, and monetary* (RFM) *model* yang digunakan untuk merepresentasikan karakteristik kebiasaan pelanggan [4]. Model RFM pertama kali diperkenalkan oleh Hughes dan saat ini banyak digunakan oleh industri termasuk manufaktur, *retailer*, dan industri jasa [5].

Model RFM melibatkan 3 variabel yaitu *Recency of the last purchases*, *Frequency of the purchases*, dan *Monetary value of the purchases* [6]. Adapun penjelasan ketiga variabel tersebut adalah :

- a. *Recency (R)*
Recency adalah jarak dari waktu transaksi terakhir kali dilakukan dengan waktu saat ini. Semakin kecil nilai jarak waktu, maka semakin besar nilai R.
- b. *Frequency (F)*
Frequency adalah total jumlah transaksi yang dilakukan selama periode tertentu. Semakin besar jumlah transaksi, maka semakin besar nilai F.
- c. *Monetary (M)*
Monetary adalah total nilai produk dalam bentuk uang dalam periode tertentu. Semakin besar nilai produk, maka semakin besar nilai M.

Representasi nilai R dan F yang semakin besar, maka pelanggan kemungkinan akan melakukan transaksi kembali. Di samping itu, jika nilai M semakin besar, maka pelanggan cenderung memberikan respon kepada produk dan layanan perusahaan [7]. Dalam implemtasinya, variabel R, F dan M diberikan bobot sesuai dengan kebutuhan.

Model RFM memiliki beberapa kelebihan, antara lain tidak memerlukan biaya besar dalam analisis dan mengukur kebiasaan pelanggan. Selain itu, RFM sangat mudah dalam memprediksi respon dan dapat meningkatkan keuntungan perusahaan dalam waktu singkat. Kelebihan lain adalah RFM sangat efektif dalam memodelkan variabel dengan jumlah variabel yang sangat sedikit. Terakhir, RFM adalah metode jangka panjang untuk mengukur tingkat kekuatan hubungan pelanggan dan menentukan pelanggan yang berharga.

Akan tetapi model RFM juga memiliki beberapa kekurangan antara lain RFM hanya berfokus dalam mengidentifikasi pelanggan terbaik saja. Yang kedua, RFM hanya dapat digunakan untuk jumlah variabel yang terbatas. Selain itu, variabel RFM bersifat independen. Terakhir, RFM berfokus pada pelanggan yang sudah terdaftar dan tidak dapat diaplikasikan pada pelanggan baru [8].

2.2.8 Normalisasi *Min-Max*

Normalisasi *Min-Max* merupakan salah satu metode normalisasi dengan melakukan perubahan linier terhadap data asli [9]. Normalisasi *Min-Max* memetakan nilai d dari P ke d' dalam rentang $[new_min(p), new_max(p)]$. Normalisasi *Min-Max* dihitung dengan formula berikut :

$$d' = \frac{[d - \min(p)] * [new_{\max}(p) - new_{\min}(p)]}{[\max(p) - \min(p)]} + new_{\min}(p) \quad (6)$$

Keterangan :

d'	= nilai yang telah dinormalisasi
d	= nilai asli
$\min(p)$	= nilai terkecil atribut
$\max(p)$	= nilai terbesar atribut
$\text{new_min}(p)$	= rentang minimal
$\text{new_max}(p)$	= rentang maksimal

2.2.9 Metode *Elbow*

Metode *Elbow* adalah suatu metode untuk menghasilkan informasi dalam menentukan jumlah *cluster* terbaik dengan cara melihat prosentase hasil perbandingan antara jumlah *cluster* yang akan membentuk sudut siku pada suatu titik [32]. Hasil prosesntase dari setiap perhitungan ditunjukkan dalam grafik dan dibandingkan. Jika nilai suatu *cluster* dengan nilai *cluster* kedua mengalami penurunan paling dalam atau nilainya mengalami penurunan paling besar maka nilai *cluster* tersebut terbaik. Langkah-langkah dalam perhitungan metode *Elbow* adalah sebagai berikut [33]:

1. Mulai
2. Inisialisasi awal nilai $K = 1$
3. Naikkan nilai K
4. Hitung hasil SSE dari tiap nilai K
5. Melihat hasil SSE nilai K yang turun secara drastis
6. Tetapkan nilai K yang membentuk siku
7. Selesai.

2.2.10 Analytical Hierarchy Process

Analytical Hierarchy Process (AHP) merupakan salah satu alat pendukung keputusan yang pertama kali diperkenalkan oleh Saaty pada 1977. AHP digunakan untuk mengembangkan nilai numerik untuk peringkat setiap alternatif keputusan, berdasarkan seberapa baik setiap alternatif memenuhi kriteria pembuat keputusan. Metode peringkat alternatif keputusan ini

akan memilih yang terbaik ketika membuat keputusan memiliki beberapa tujuan, atau kriteria yang menjadi dasar pada keputusan [13]. Langkah yang harus dilakukan untuk memperoleh bobot dengan metode AHP adalah [14] :

1. Membuat *pairwise comparison*, dengan melakukan survei terlebih dahulu melalui kuesioner untuk memberikan nilai setiap kriteria yang dipasangkan mengacu pada skala preferensi tabel berikut :

Tabel 2.5 Skala Preferensi Penilaian Pairwise Comparison

Skala Nilai	Tingkat Preferensi	Penjelasan
1	Sama Pentingnya	Kedua kriteria memiliki pengaruh yang sama
3	Sedikit lebih penting	Penilaian dan pengalaman sedikit memihak pada salah satu kriteria tertentu dibandingkan kriteria pasangannya
5	Lebih penting	Penilaian dan pengalaman memihak pada salah satu kriteria tertentu dibandingkan kriteria pasangannya
7	Jelas lebih penting	Salah satu kriteria lebih diprioritaskan dan relatif lebih penting dibandingkan kriteria pasangannya.
9	Mutlak sangat penting	Salah satu kriteria sangat jelas lebih penting dibandingkan kriteria pasangannya.
2,4,6,8		Diberikan bila ada keraguan penilaian diantara dua tingkat kepentingan yang berdekatan.

2. Setelah menentukan prioritas masing-masing kriteria menggunakan matriks *pairwise comparison*, kemudian

melakukan normalisasi matriks. Pertama-tama menambahkan jumlah setiap kolom, kemudian membagi setiap elemen dengan hasil penjumlahan setiap kolomnya. Matriks baru yang terbentuk disebut *normal comparison*.

3. Menghitung bobot setiap kriteria dengan cara menghitung rata-rata angka setiap baris dalam matriks *normal comparison*.
4. Melakukan uji rasio konsistensi, yaitu instrument yang menentukan konsistensi dan menunjukkan tingkat kepercayaan prioritas yang diperoleh dari perbandingan. Terdapat dua tahapan dalam uji rasio konsistensi, yaitu :
 - a. Setiap elemen pada tabel yang berisi matriks *pairwise comparison* dikalikan dengan bobot dari setiap kriteria yang didapatkan pada langkah ke 3, lalu menghitung total skor dengan menjumlahkan setiap barisnya.
 - b. Mencari nilai x sebagai masukan untuk menghitung CI dengan cara mengalikan total skor dengan bobot kemudian hasilnya dirata-rata.
 - c. Menghitung indeks konsistensi dengan persamaan berikut :
 - d. Menghitung tingkat ketidakkonsistensian yang dapat diterima dengan persamaan berikut :

$$CR = \frac{CI}{RI} \quad (7)$$

Keterangan :

CR	= <i>Consistency Ratio</i>
CI	= indeks konsistensi
RI	= random indeks

Tabel 2.6 Nilai Random Indeks

n	2	3	4	5	6	7	8	9	10
RI	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.51

Tingkat ketidakkonsistensian diterima apabila nilai yang didapatkan dari perhitungan *CR* lebih kecil dari 0.1. Apabila CR bernilai lebih dari 0.1 atau tidak konsisten, maka dapat digunakan algoritma sebagai berikut [35] :

1. Tentukan vektor w_n menggunakan metode AHP dan cek indeks konsistensi *CR*.
2. Jika $CR < 0.1$, hentikan perhitungan, jika sebaliknya, lanjutkan ke tahap selanjutnya.
3. Tentukan matriks *pairwise comparison* baru, W , dari formula :

$$W_s = W_n \left(\frac{1}{w_n} \right)^T \quad (8)$$

4. Pada basis matriks W_s dan tabel skala preferensi, tentukan usulan baru untuk matriks perbandingan.
5. Meminta pertimbangan pada pengambil keputusan untuk menyetujui usulan matriks perbandingan atau mengevaluasi *pairwise comparison* dari matriks W .
6. Jika pengambil keputusan menyetujui matriks perbandingan baru, maka hentikan perhitungan. Jika tidak, maka ulangi langkah 3.

2.2.11 Customer Lifetime Value (CLV)

Customer Lifetime Value (CLV) adalah sebuah metric pemasaran yang menggambarkan nilai dari seorang pelanggan sepanjang hubungan pelanggan dengan perusahaan berlangsung [11]. CLV menentukan nilai pelanggan ke perusahaan selama siklus hidup pelanggan. CLV berusaha

memaksimalkan keuntungan dengan menganalisis perilaku pelanggan dan siklus bisnis untuk mengidentifikasi dan menargetkan pelanggan dengan potensi nilai terbesar dari waktu ke waktu.

Dalam penelitian tugas akhir ini, penggunaan CLV mengacu pada indeks nilai CLV. Indeks CLV ditentukan melalui bobot variabel RFM yang sebelumnya telah diperoleh dan nilai normalisasi masing-masing variabel RFM. Adapun formula untuk menghitung nilai CLV adalah sebagai berikut :

$$CLV = NRxWR + NFxWF + NMxWM$$

Keterangan :

CLV	= nilai CLV yang dicari
NR	= nilai normalisasi variabel <i>recency</i>
WR	= nilai bobot variabel <i>recency</i>
NF	= nilai normalisasi variabel <i>frequency</i>
WF	= nilai bobot variabel <i>frequency</i>
NM	= nilai normalisasi variabel <i>monetary</i>
WM	= nilai bobot variabel <i>monetary</i>

2.2.12 Sum Square Error (SSE)

Sum Square Error (SSE) adalah salah satu formula sederhana yang digunakan untuk mengevaluasi hasil *cluster*. Dalam perhitungan SSE, setiap objek pada *cluster* akan dihitung tingkat kesalahan terhadap *centroid* terdekat. Semakin kecil nilai SSE, maka semakin berkualitas hasil *clustering* yang telah dilakukan. Formula untuk menghitung nilai SSE adalah [31]:

$$SSE = \sum_{i=1}^K \sum_{x \in C_i} dist(m_i, x)^2 \quad (9)$$

Keterangan :

$$K = \text{jumlah cluster}$$

x = objek pada *cluster C*

m_i = *centroid cluster i*

$dist$ = jarak objek ke centroid terdekat pada masing-masing
cluster i

C_i = *cluster ke-i*

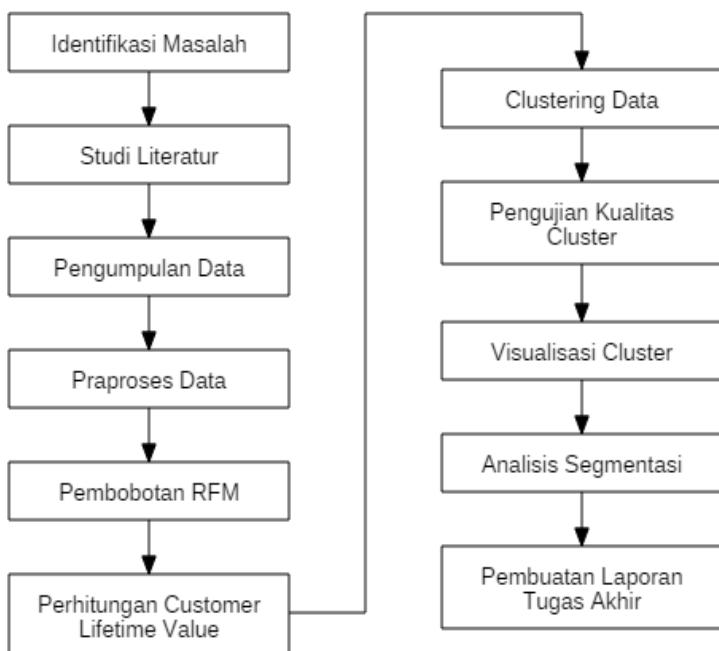
BAB III

METODOLOGI PENELITIAN

Pada bab ini akan menjelaskan tentang metodologi yang akan digunakan dengan deskripsi setiap tahapannya

3.1 Metodologi

Metodologi digunakan sebagai panduan dalam penyusunan tugas akhir agar terstruktur dan sistematis. Adapun urutan dari pengerjaan tugas akhir pada Gambar 3.1 berikut ini :



Gambar 3.1 Alur Metodologi Penelitian

Penelitian dimulai dari identifikasi masalah yang menghasilkan rumusan masalah, batasan masalah, tujuan dan manfaat. Tahap selanjutnya adalah studi literatur untuk memperoleh referensi terkait penelitian ini. Kemudian dilakukan pengumpulan data untuk mendapatkan data transaksi penjualan. Data transaksi ini kemudian diproses pada tahap *data cleaning* untuk memperoleh data bersih. Setelah itu, data bersih ditransformasikan menjadi tiga variabel, yaitu *recency*, *frequency*, dan *monetary* menghasilkan data RFM. Nilai pada data RFM selanjutnya akan dinormalisasi untuk menyetarakan skala nilai dan menghasilkan data RFM ternormalisasi. Kemudian, dilakukan pembobotan pada setiap variabel RFM dengan metode *Analytical Hierarchy Process* sehingga menghasilkan bobot untuk masing-masing variabel. CLV merupakan indeks pada setiap pelanggan yang menunjukkan seberapa bernilai pelanggan bagi perusahaan. CLV digunakan untuk menilai segmen mana yang menjadi prioritas. Selanjutnya, pada tahap perhitungan *Customer Lifetime Value*, data RFM ternormalisasi dikalikan dengan bobot masing-masing dan menghasilkan data RFM bobot untuk masukan pada tahap *clustering*. Setelah itu, ditentukan terlebih dahulu nilai K dengan metode Elbow dan dilakukan proses *clustering* terhadap data RFM bobot. Hasil *clustering* kemudian divisualisasi dan dianalisis karakteristik setiap segmennya.

3.1.1 Identifikasi Masalah

Tahap ini merupakan tahap awal dalam pengerjaan tugas akhir ini. Pada tahap ini dilakukan analisis permasalahan yang akan diselesaikan pada tugas akhir ini. Studi kasus yang diambil adalah PT. XYZ, dimana permasalahan yang diangkat adalah segmentasi pelanggan produk. Setelah permasalahan telah diidentifikasi, selanjutnya adalah penetapan rumusan masalah, batasan masalah, tujuan, dan manfaat.

3.1.2 Studi Literatur

Tahap selanjutnya adalah studi literatur. Pada tahap ini dilakukan pencarian dan pengumpulan berbagai referensi penelitian terdahulu. Literatur yang dibahas adalah terkait dengan penggalian data, *clustering*, *Fuzzy C-Means*, metode RFM, dan segmentasi pelanggan yang berhubungan dengan tugas akhir. Studi literatur dilakukan berdasarkan narasumber dengan teknik wawancara, buku, penelitian sebelumnya dan dokumen lainnya.

3.1.3 Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data-data yang diperlukan untuk penelitian tugas akhir, terutama yang terkait dengan proses *clustering*. Adapun data-data yang dikumpulkan meliputi :

1. Data pelanggan yang pernah melakukan transaksi
2. Data transaksi penjualan produk ke pelanggan dari Januari 2017 hingga Juni 2017

Data transaksi penjualan yang diperoleh memiliki atribut seperti tanggal transaksi, nama pelanggan, jenis produk yang dibeli, kuantitas produk, dan pendapatan yang diperoleh perusahaan. Dari data transaksi yang diperoleh, nilai variabel sudah cukup variatif sehingga bisa untuk dilakukan *clustering*. Selain pengumpulan data, dilakukan juga wawancara kepada narasumber terkait kondisi perusahaan yang berhubungan dengan segmentasi sebagai pendukung dalam penelitian tugas akhir. Luaran dari tahap ini adalah data transaksi penjualan yang dilakukan selama 6 bulan dari Januari hingga Juni 2017.

3.1.4 Praproses Data

Setelah pengumpulan data, tahap selanjutnya adalah melakukan pra proses data sebelum masuk ke proses *clustering*. Praproses data dibutuhkan agar data mentah yang telah dikumpulkan dapat diolah sesuai dengan kriteria yang diinginkan. Masukan tahap ini adalah data transaksi yang sebelumnya diperoleh. Tahap praproses data ini meliputi *data cleaning*, pemodelan RFM dan normalisasi.

3.1.4.1 Data Cleaning

Data cleaning merupakan subproses untuk melakukan pembersihan data mentah dari nilai yang kosong, negatif, dan tidak sesuai kriteria. Pembersihan data dilakukan bertujuan untuk menghasilkan data yang lebih bersih sehingga hasil *clustering* lebih akurat.

3.1.4.2 Pemodelan RFM

Pemodelan RFM merupakan subproses untuk membentuk variabel *recency*, *frequency*, dan *monetary* dari data transaksi yang masih mentah. Adapun pemodelan variabel-variable tersebut antara lain :

- a. *Recency* : variabel *recency* diperoleh dari selisih waktu transaksi yang terakhir kali dilakukan oleh masing-masing pelanggan per tanggal 1 Oktober 2017. Atribut yang berkaitan dengan variabel ini adalah tanggal transaksi.
- b. *Frequency* : variabel *frequency* diperoleh dari jumlah transaksi yang dilakukan oleh masing-masing pelanggan dengan mengagregasikan semua data berdasarkan pelanggan.
- c. *Monetary* : variabel *monetary* diperoleh dari jumlah total pendapatan yang diperoleh perusahaan berdasarkan

masing-masing pelanggan. Atribut yang berkaitan dengan variabel ini adalah pendapatan yang dijumlah berdasarkan tiap pelanggan.

3.1.4.3 Normalisasi

Normalisasi merupakan subproses untuk menormalkan nilai suatu variabel agar seimbang dengan variabel lainnya. Adapun variabel yang dinormalisasi adalah variabel *monetary* karena memiliki nilai dengan rentang lebih jauh dibanding dengan variabel *recency* dan *frequency*.

Setelah tiga subproses tersebut dilakukan, akan menghasilkan data transaksi yang teragregasi berdasarkan tiap pelanggan, dengan tiga atribut, yaitu *recency*, *frequency*, dan *monetary* yang siap untuk *dicluster*.

3.1.5 Pembobotan RFM

Pada tahapan ini dilakukan pembobotan terhadap variabel RFM dengan metode AHP. Penilaian bobot diperoleh dengan memberikan kuesioner kepada tiga karyawan pada divisi *Petrochemical Trading*, yaitu *Sales Executive Petrochemical*, *Senior Supervisor Sales Bitumen*, dan *Senior Supervisor Sales Special Chemical*. Dari perhitungan AHP tersebut, diperoleh nilai bobot untuk variabel *recency*, *frequency*, dan *monetary*. Setelah itu dilakukan uji konsistensi terhadap nilai bobot yang telah dihitung.

3.1.6 Perhitungan CLV

Tahapan selanjutnya yaitu menghitung indeks CLV terhadap pelanggan. Masukan dari tahap ini adalah nilai atribut RFM dari masing-masing pelanggan dan bobot masing-masing variabel. Kemudian dilakukan perkalian nilai variabel RFM dengan bobot masing-masing. Indeks CLV didapatkan dari

penjumlahan semua nilai variabel RFM. Tingkat loyalitas pelanggan dapat diketahui dari besar kecilnya indeks CLV. Luaran dari tahap ini adalah nilai CLV dari masing-masing pelanggan.

3.1.7 Clustering Data

Tahap selanjutnya adalah *clustering* data, dimana data yang sebelumnya telah siap diolah kemudian melalui proses *clustering*. Masukan dari tahap ini adalah nilai CLV dari tahap sebelumnya. Sebelum *clustering* dilakukan, sebelumnya ditentukan terlebih dahulu nilai K sebagai jumlah *cluster* yang akan dibentuk. Metode yang digunakan untuk menentukan jumlah K adalah *Elbow*. Metode *clustering* yang digunakan adalah *Fuzzy C-Means*. *Clustering* bertujuan untuk mengelompokkan data ke dalam beberapa *cluster* yang memiliki kemiripan sifat atau atribut. *Clustering* dilakukan berdasarkan nilai ketiga variabel yang diplot ke dalam ruang koordinat tiga dimensi. Setelah proses *clustering* selesai, luarannya adalah *cluster* yang berisikan data pelanggan yang memiliki perilaku pelanggan tertentu. Perilaku pelanggan bisa berupa frekuensi pembelian yang tinggi dan rendah atau pelanggan dengan nilai pembelian yang banyak dan sedikit.

3.1.8 Pengujian Kualitas Cluster

Setelah proses *clustering* selesai, selanjutnya adalah tahap pengujian kualitas *cluster*. Masukan dari tahap ini adalah data pelanggan serta pada *cluster* apa pelanggan tersebut berada. Pengujian hasil *cluster* dilakukan dengan menguji validitas berdasarkan SSE. Nilai SSE diperoleh dari rata-rata jarak semua objek terhadap titik pusat *cluster* pada masing-masing *cluster*. Nilai SSE yang semakin kecil menunjukkan kualitas *cluster* yang semakin baik.

3.1.9 Visualisasi *Cluster*

Pada tahap visualisasi *cluster*, hasil *cluster* akan divisualisaikan dengan media dashboard berbasis web dengan alat bantu RShiny. Visualisasi ini meliputi grafik *scattered plot 3D*, *pie chart*, dan *Box and Whisker*. Visualisasi ini bertujuan untuk memudahkan dalam pembacaan hasil *clustering* untuk menganalisis segmentasi pelanggan pada setiap *cluster*-nya.

3.1.10 Analisis Segmentasi

Luaran hasil *clustering* memberikan gambaran bagaimana pelanggan-pelanggan dikelompokkan dalam beberapa *cluster*. Analisis segmentasi dilakukan untuk menjelaskan ciri atau kemiripan anggota dalam satu segmen, serta perbandingan dengan segmen lainnya. Analisis ini juga bertujuan untuk mengidentifikasi segmen mana yang menjadi prioritas dibanding segmen yang lain. Prioritas ditentukan berdasarkan tinggi rendahnya nilai ketiga variabel pada keseluruhan anggota cluster. Selain itu juga dilakukan penjelasan mengenai hasil visualisasi grafik sebelumnya.

3.1.11 Penyusunan Laporan Tugas Akhir

Tahap terakhir adalah penyusunan laporan tugas akhir sebagai dokumentasi terhadap keseluruhan tugas akhir. Seluruh pelaksanaan atau penggerjaan tugas akhir di dokumentasikan dalam sebuah buku Tugas Akhir (TA) dengan mengikuti format yang telah ditetapkan oleh laboratorium Rekayasa Data dan Intelelegensi Bisnis (RDIB) serta yang berlaku di Departemen Sistem Informasi ITS.

Halaman ini sengaja dikosongkan

BAB IV

PERANCANGAN

Pada bab ini akan dijelaskan tentang rancangan dari tugas akhir yang terdiri dari proses pengumpulan data, praproses data, pemberian bobot RFM, penentuan jumlah *cluster* dan rancangan visualisasi.

4.1 Pengumpulan data

Pengumpulan data dilakukan secara langsung pada divisi Petrochemical Trading, PT. XYZ Surabaya. Data diperoleh dari salah satu karyawan yang bertanggungjawab dalam pelaporan penjualan bulanan. Data yang dibutuhkan adalah riwayat transaksi penjualan dari tanggal 1 Januari 2017 sampai 30 Juni 2017 dan daftar pelanggan yang pernah melakukan transaksi. Selain itu, wawancara dilakukan untuk menggali permasalahan yang dihadapi terkait dengan strategi pemasaran yang dilakukan. Data transaksi penjualan dan pelanggan didapatkan dalam bentuk dokumen *excel* dengan variabel dijelaskan pada Tabel 4.1 dan 4.2.

Tabel 4.1 Data Mentah Transaksi Penjualan

Variabel	Tipe Data	Keterangan
Date	Date	Tanggal transaksi dilakukan
Calendar Year/Month	Text	Keterangan bulan dan tahun transaksi dilakukan
Sales Organization	Text	Divisi yang menjual produk
Material	Text	Kode produk yang dijual
Material Name	Text	Nama produk yang dijual
Region	Text	Area transaksi
Sold-to-party	Number	Kode pelanggan
Sold-to-party	Text	Nama pelanggan

Unit of measure	Text	Satuan produk yang dijual
Billing Quantity SU	Number	Jumlah produk yang dijual
Net Value Invoice in RP	Number	Nilai uang yang diterima
Material Code	Text	Kode produk
Product Group	Text	Grup produk dari produk yang dijual

Data transaksi yang diperoleh berjumlah 1355 baris data, terhitung mulai Januari hingga Juni 2017.

Tabel 4.2 Data Mentah Tabel Pelanggan

Variabel	Tipe Data	Keterangan
Sold-to-party	Number	ID pelanggan
Customer	Text	Nama pelanggan
Area	Text	Area transaksi
Address	Text	Alamat pelanggan

Data pelanggan yang diperoleh berjumlah 123 baris data, terhitung melakukan transaksi dari Januari hingga Juni 2017.

4.2 Praproses data

Sebelum dapat diolah, data mentah memasuki tahap praproses data, sehingga data mentah yang diperoleh dapat diolah dan menghasilkan hasil yang diharapkan. Tahap praproses yang dilakukan antara lain :

1. Pemilihan data untuk menentukan atribut yang akan digunakan sebagai variabel *cluster*. Atribut yang dibutuhkan mengacu pada model RFM, yaitu *recency*, *frequency*, dan *monetary*.
2. Pembersihan data untuk menghilangkan data yang kosong atau nilai atribut yang dimiliki tidak valid. Pembersihan data dilakukan agar hasil *clustering* lebih baik dengan tingkat kesalahan lebih kecil.

3. Pemodelan variabel RFM untuk mendapatkan nilai variabel RFM yang dibutuhkan berdasarkan atribut yang tersedia.
 - a. Variabel *recency* diperoleh dari selisih waktu transaksi yang terakhir kali dilakukan oleh masing-masing pelanggan per tanggal 1 Oktober 2017. Atribut yang berkaitan dengan variabel ini adalah tanggal transaksi.
 - b. Variabel *frequency* diperoleh dari jumlah transaksi yang dilakukan oleh masing-masing pelanggan dengan mengagregasikan semua data berdasarkan setiap pelanggan.
 - c. Variabel *monetary* diperoleh dari jumlah total pendapatan yang diperoleh perusahaan berdasarkan masing-masing pelanggan. Atribut yang berkaitan dengan variabel ini adalah pendapatan yang diakumulasi berdasarkan tiap pelanggan.
4. Transformasi data untuk mengubah format *file .xlsx* menjadi format *file .csv* agar bisa diinputkan ke dalam aplikasi RStudio.
5. Normalisasi data untuk menyetarakan skala nilai variabel *monetary* yang cukup jauh dengan nilai variabel *recency* dan *frequency*. Metode normalisasi yang digunakan adalah metode *min-max*.

4.3 Pembobotan RFM

Tahap selanjutnya dalam analisis RFM adalah menentukan bobot untuk setiap variabel RFM. Bobot yang diberikan pada setiap variabel menentukan tingkat kepentingan setiap variabel dibanding dengan variabel lainnya. Bobot variabel ini diperoleh melalui proses pengisian kuesioner oleh pihak terkait di fungsi *Petrochemical PT. XYZ Surabaya*. Responden yang mengisi kuesioner ada tiga orang, yaitu *senior sales executive*, *supervisor bitumen*, dan *supervisor special chemical*. Kuesioner yang diberikan berisikan data responden, petunjuk

pengisian kuesioner, penjelasan variabel RFM, skala preferensi, contoh pengisian, dan daftar pertanyaan. Rancangan kuesioner dapat dilihat pada Lampiran A.

4.4 Penentuan jumlah *cluster*

Sebelum proses *clustering* dilakukan, tahap sebelumnya adalah penentuan nilai k atau jumlah *cluster* yang akan dibentuk. Dalam penentuan nilai k ini, metode yang digunakan adalah metode *Elbow*. Hasil dari metode *Elbow* adalah grafik yang menunjukkan nilai SSE pada setiap nilai k . Nilai k dipilih ketika ada penurunan nilai SSE yang drastis pada suatu nilai k , diikuti dengan nilai yang stabil pada nilai k selanjutnya. Aplikasi yang digunakan untuk menghitung nilai k adalah RStudio.

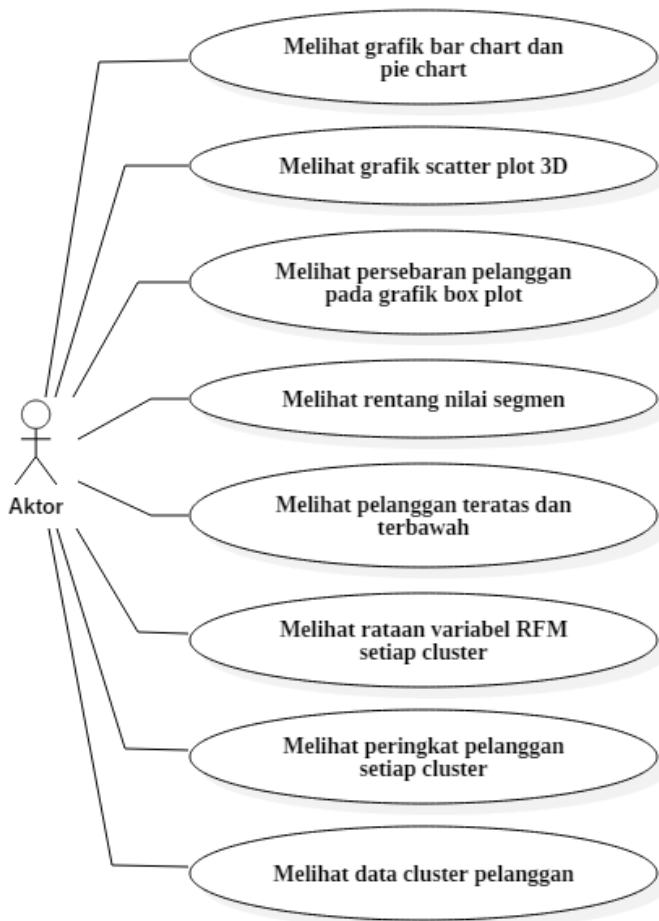
4.5 Rancangan visualisasi

Rancangan visualisasi hasil *clustering* meliputi penyusunan *use case diagram*, *use case description* dan desain antarmuka yang menjelaskan mengenai alur penggunaan aplikasi visualisasi.

4.5.1 Use Case Diagram

Visualisasi ini dibuat untuk menampilkan setiap grafik yang menjelaskan informasi karakteristik dari setiap pelanggan hasil dari *clustering*. *Use case diagram* yang tersaji pada Gambar 4.3, terbagi dalam 8 *use case* utama, yaitu :

1. Melihat grafik *bar chart* dan *pie chart*
2. Melihat grafik *scatter plot* 3D
3. Melihat persebaran pada grafik *box-plot*
4. Melihat rentang nilai segmen
5. Melihat pelanggan teratas dan terbawah
6. Melihat rataan variabel RFM setiap *cluster*
7. Melihat peringkat pelanggan setiap *cluster*
8. Melihat tabel data *cluster* pelanggan



Gambar 4.1 Use Case Diagram

4.5.2 Use Case Description

Use case melihat grafik *bar chart* dan *pie chart* dideskripsikan menjadi *use case description*, ditampilkan pada Tabel 4.3.

Untuk tujuh *use case description* yang lain dapat dilihat pada Lampiran B.

Tabel 4.3 Use case melihat grafik bar chart dan pie chart

UC-1 Melihat grafik bar chart dan pie chart		
Purpose	Mengetahui jumlah dan prosentase pelanggan pada masing-masing <i>cluster</i> dalam grafik <i>bar chart</i> dan <i>pie chart</i>	
Overview	Dimulai ketika pengguna mengakses aplikasi	
Actors	Pengguna aplikasi	
Pre Condition	Data hasil <i>clustering</i> telah dimasukkan pada aplikasi	
Post Condition	Grafik <i>bar chart</i> dan <i>pie chart</i> ditampilkan	
Typical Course Event	Actor 1. Mengklik ikon  3. Mengklik menu “Halaman Utama”	System 2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan grafik “Jumlah Anggota Segmen” dan “Prosentase Anggota Segmen”
Alternate Flow Event	-	
Exceptional Flow Event	Aplikasi tertutup karena <i>error</i>	

4.5.3 Desain Antarmuka

Desain antarmuka menjelaskan mengenai tampilan antarmuka dari aplikasi yang diakses oleh pengguna. Aplikasi visualisasi menggunakan web ini didesain dengan 3 halaman yang dapat diakses, yaitu :

a. Halaman utama dashboard

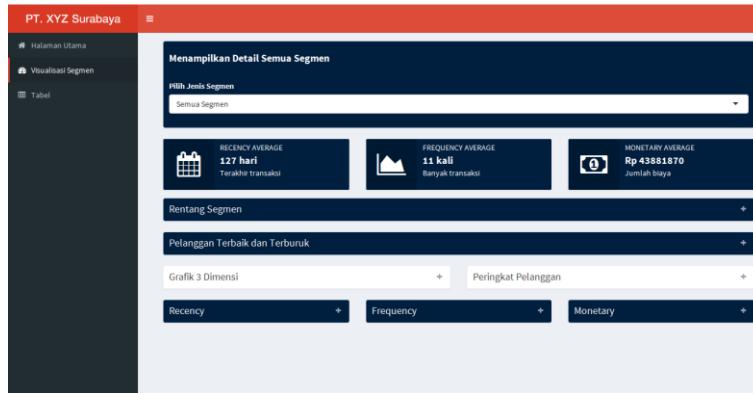
Halaman utama dashboard menampilkan informasi jumlah pelanggan, jumlah pemasukan, jumlah transaksi, grafik jumlah pelanggan setiap segmen dalam bentuk *bar chart*, dan prosentase jumlah pelanggan setiap segmen dalam bentuk *pie chart*. Rancangan halaman utama dashboard disajikan pada Gambar 4.2.



Gambar 4.2 Halaman Utama Dashboard

b. Halaman visualisasi segmen

Halaman visualisasi segmen menampilkan informasi rata-rata variabel RFM, rentang nilai variabel RFM, pelanggan teratas dan terbawah, grafik *scater plot 3D*, tabel peringkat pelanggan, dan grafik *box plot* variabel RFM setiap segmen yang dipilih. Rancangan visualisasi grafik cluster disajikan pada Gambar 4.3.



Gambar 4.3 Halaman Visualisasi Segmen

c. Halaman tabel data

Halaman tabel data menampilkan informasi tabel detail peringkat seluruh pelanggan. Rancangan halaman tabel data disajikan pada Gambar 4.4.

Code	Recency	Frequency	Monetary	CLV
2053	0.0818163005350907	0.350470327214513	0.0617809732456539	0.494067600995258
2138	0.0898745720446953	0.0598017503753412	0.00309050163414517	0.146966882783448
2166	0.081301732807197	0.0534615753780701	0.234287816274624	0.369051124219629
700032	0.081301732807197	0.15444550975887	0.00664492310783205	0.24239320669918
700189	0.02675753225047	0	0.0001137751199346374	0.0268713074498163
700286	0.080272597514097	0.0358410502252046	0.00689136204747512	0.122805093094089
700385	0.0463965210777906	0.0176205251128023	0.00178095714587574	0.0690800033383687
700493	0.000081740756736189	0	0.0000818427059704637	0.0013895027243236
700503	0.069466070265431	0.0594017503753412	0.00826627626439832	0.13713469905383
700529	0.081301732807197	0.089302635680117	0.00196741415575437	0.172371772325963

Gambar 4.4 Halaman Tabel Data

BAB V

IMPLEMENTASI

Pada bab ini menjelaskan proses implementasi berdasarkan dari rancangan yang dibuat yaitu pra proses data, pembobotan RFM, perhitungan CLV, proses *clustering*, uji performa *clustering* dan proses pembuatan visualisasi.

5.1 Pra Proses Data

Praproses data dibutuhkan agar data mentah yang telah dikumpulkan dapat diolah sesuai dengan kriteria yang diinginkan. Masukan tahap ini adalah data transaksi yang sebelumnya diperoleh. Tahap praproses data ini meliputi *data cleaning*, pemodelan RFM dan normalisasi.

5.1.1 Data Cleaning

Pembersihan data dilakukan untuk menghilangkan data yang tidak valid atau redundan. Sebelum itu, dipilih terlebih dahulu atribut yang penting berkaitan dengan model RFM. Dari 13 atribut yang ada, atribut yang dipilih adalah date, sold-to-party, dan net value invoice in RP. Penjelasan mengenai atribut yang dipilih tersaji pada Tabel 5.1.

Tabel 5.1 Atribut Terpilih

Atribut	Keterangan
date	Atribut date digunakan untuk memperoleh variabel <i>recency</i> , dengan mencari selisih hari.
sold-to-party	Atribut date digunakan untuk memperoleh variabel <i>frequency</i> , dengan menghitung jumlah transaksi per pelanggan.

net value invoice in RP	Atribut net value invoice in RP digunakan untuk memperoleh variabel <i>monetary</i> , dengan menjumlah net value per pelanggan
-------------------------	--

Setelah itu, dilakukan penghapusan data yang tidak valid pada atribut-atribut tersebut. Saat dilakukan validasi, ditemukan baris yang bernilai 0 dan negatif pada atribut net value in RP. Pencarian baris yang bernilai 0 dan negatif dilakukan dengan fitur *sort* pada *excel*. Setelah diurutkan, didapatkan 12 baris data yang bernilai negatif dan 3 data yang bernilai 0. 15 baris data yang tidak valid tersebut kemudian dihapus dan tersisa 1340 baris data. Adapun data yang dihapus ditampilkan ada Tabel 5.2.

Tabel 5.2 Baris Data Tidak Valid

Date	Sold-to-party	Net Value Invoice in RP
06/01/2017	2053	-5747335.916
15/01/2017	821868	-360221.926
27/01/2017	728260	-195.072
15/02/2017	830033	-230595.75
19/02/2017	821868	-101542.072
22/02/2017	821868	-50152.32
04/04/2017	704904	-4560
05/04/2017	706383	-678587.092
06/04/2017	700385	-75243.827
24/04/2017	722171	0
24/04/2017	821868	-3072.278
02/06/2017	719930	-526292.484
18/06/2017	892303	0
21/06/2017	883122	0
25/06/2017	704402	-177.51

5.1.2 Pemodelan RFM

Pemodelan RFM dilakukan untuk membentuk nilai variabel *recency*, *frequency*, dan *monetary* sebagai masukan proses *clustering*. Nilai RFM didapat dengan mengekstrak nilai atribut yang sebelumnya dipilih. Pemodelan RFM dilakukan dengan *query* pada basis data *MySQL*.

Nilai variabel *recency* atau R adalah selisih hari sejak transaksi terakhir kali dilakukan oleh setiap pelanggan dengan hari pengerajan, yaitu tanggal 1 Oktober 2017. Atribut yang dibutuhkan untuk memperoleh nilai R adalah sold-to-party dan date. Nilai R didapatkan dari tanggal transaksi yang paling terakhir menggunakan fungsi MAX berdasarkan tiap pelanggan dengan fungsi GROUP BY, sedangkan untuk mencari selisihnya menggunakan fungsi DATEDIFF.

Nilai variabel *frequency* atau F adalah jumlah transaksi yang dilakukan oleh setiap pelanggan. Atribut yang diperlukan untuk memperoleh nilai F adalah sold-to-party dan date. Nilai F diperoleh dengan menghitung banyaknya date dengan fungsi COUNT dan GROUP BY berdasarkan sold-to-party.

Nilai variabel *monetary* atau M adalah jumlah pemasukan yang diterima perusahaan dari tiap pelanggan. Atribut yang dibutuhkan untuk memperoleh nilai F adalah sold-to-party dan net value invoice in RP. Nilai M didapatkan dengan menjumlah atribut net value invoice in RP dengan fungsi COUNT dan fungsi GROUP BY berdasarkan sold-to-party.

Setelah dihitung masing-masing variabel, atribut yang diperoleh adalah customer_id, R, F, dan M yang diurutkan berdasarkan customer_id. Untuk *query* yang dilakukan dapat dilihat pada Script 5.1.

```

SELECT
id_stp as STP,
DATEDIFF('2017-10-01',MAX(date)) as R,
COUNT(date) as F,
SUM(net_value) as M
FROM `sales` GROUP BY id_stp
ORDER BY customer_id

```

Script 5.1 Query Ekstraksi Data

Setelah *query* dieksekusi, diperoleh data sejumlah 123 pelanggan yang melakukan transaksi sejak Januari 2017 hingga Juni 2017. Sebagian hasil pemodelan RFM dapat dilihat pada Tabel 5.3 dan lebih lengkapnya dapat dilihat pada Lampiran C.

Tabel 5.3 Sebagian Hasil Pemodelan RFM

STP	R	F	M
2053	97	60	94075983
2138	93	11	4731484
2166	98	10	356683172
700032	98	27	10145412
700189	204	1	200000
700286	100	7	10517522
700385	160	4	2750130
700493	250	1	151085
700503	121	11	12610553
700529	98	16	3021795

5.1.3 Normalisasi

Normalisasi data dilakukan untuk menyetarakan skala variabel *monetary* yang terlalu jauh dengan variabel *recency* dan *frequency*. Normalisasi dilakukan dengan aplikasi RStudio, oleh karena itu format file harus diubah terlebih dulu. Format

file data mentah berupa file *Excel Worksheet* (.xlsx) harus diubah menjadi file *Comma Separated Value* (.csv) agar bisa diproses. Perubahan format file dilakukan melalui fitur *Save As* pada *Microsoft Excel* dengan format file .csv.

Setelah itu, normalisasi dilakukan dengan metode *Min-Max* dari skala 0 sampai 1. Normalisasi dilakukan pada semua nilai variabel R, F, dan M dengan memperhatikan nilai maksimum dan minimum setiap variabel. Penjelasan lebih detail mengenai nilai maksimum dan minimum setiap variabel dapat dilihat pada Tabel 5.4.

Tabel 5.4 Rentang Nilai Atribut

Variabel	Maksimum	Minimum
<i>Recency</i>	256	93
<i>Frequency</i>	60	1
<i>Monetary</i>	861123629	26800

Perhitungan normalisasi *Min-Max* dilakukan berdasarkan rumus yang telah dijelaskan pada Bab II. Setiap nilai dirangi dengan nilai minimal variabelnya, kemudian dibagi dengan selisih antara nilai maksimum dan nilai minimum setiap variabel. Perhitungan lebih detail dapat dilihat pada Script 5.2.

```
> R_Norm <- (DataRFM$R-min(DataRFM$R))/(max(DataRFM$R)-min(DataRFM$R))
> F_Norm <- (DataRFM$F-min(DataRFM$F))/(max(DataRFM$F)-min(DataRFM$F))
> M_Norm <- (DataRFM$M-min(DataRFM$M))/(max(DataRFM$M)-min(DataRFM$M))
> hasil <- data.frame(DataRFM,R_Norm,F_Norm,M_Norm)
> write.csv(hasil,"D:/TA/Data/RFM.csv")
```

Script 5.2 Normalisasi Data

Nilai variabel R yang dihasilkan memiliki makna berbeda dengan variabel F dan M. Nilai R yang semakin kecil menunjukkan nilai yang paling baik, sedangkan nilai F dan M yang semakin kecil menunjukkan nilai paling buruk. Untuk itu, variabel R diubah dengan cara mengurangkan nilai R terhadap 1 pada kolom baru bernama R_Rev. Hasil lengkap nilai RFM yang ternormalisasi dapat dilihat pada Lampiran D.

5.2 Pembobotan RFM

Dari proses pengisian kuesioner, didapatkan nilai preferensi terhadap tiga kriteria RFM dari tiga responden. Hasil kuesioner tersebut digunakan sebagai masukan pada tahap pembobotan nilai RFM. Data kuesioner diolah dengan cara merata-rata semua pengisian responden terhadap perbandingan tiga kriteria RFM.

Tahap selanjutnya yaitu melakukan normalisasi matriks. Normalisasi dilakukan dengan cara membagi setiap nilai perbandingan dengan jumlah kolom pada masing-masing kolom. Kemudian, bobot setiap kriteria diperoleh dengan merata-rata kriteria setiap baris. Tabel baru yang dihasilkan tersebut disebut sebagai *normal comparisons matrix*.

Setelah bobot masing-masing kriteria diperoleh, perlu dilakukan uji konsistensi. Uji konsistensi dilakukan untuk mengetahui apakah hasil kuesioner bersifat konsisten atau tidak. Uji konsistensi dilakukan dengan cara mengalikan setiap nilai perbandingan pada *normal comparisons matrix* dengan bobot masing-masing kriteria. Lalu setiap baris kriteria dijumlahkan menjadi *overall score*. Selanjutnya, dicari nilai x dengan merata-rata hasil pembagian *overall score* dengan nilai bobot masing-masing kriteria.

5.3 Perhitungan CLV

Perhitungan CLV dilakukan untuk mengetahui nilai pelanggan terhadap perusahaan di masa mendatang. Selain itu, nilai CLV juga digunakan sebagai masukan untuk proses *clustering*, akan tetapi hanya nilai RFM-nya saja. Data yang dibutuhkan adalah data RFM tiap pelanggan dan bobot untuk setiap variabel RFM. Setiap nilai RFM masing-masing pelanggan akan dikalikan dengan bobot tiap variabel. Hasil perkalian inilah yang digunakan sebagai masukan untuk proses clustering. Kemudian, hasil perkalian tersebut dijumlahkan untuk memperoleh nilai CLV, dan digunakan sebagai salah satu bahan analisis segmentasi.

5.4 Proses Clustering

Proses *clustering* dilakukan untuk menghasilkan *cluster* yang beranggotakan pelanggan-pelanggan dengan karakteristik yang sama. Masukan dari proses ini adalah data nilai variabel RFM semua pelanggan yang telah dikalikan dengan bobot setiap variabel. Dalam proses ini terdapat dua subproses, yaitu penentuan jumlah K dan *clustering* dengan *Fuzzy C-Means*. Keseluruhan proses *clustering* dilakukan dengan aplikasi RStudio.

5.4.1 Penentuan Jumlah K

Data dari proses sebelumnya, yaitu nilai variabel RFM semua pelanggan yang telah dikalikan dengan bobot setiap variabel, digunakan sebagai dataset untuk menentukan jumlah K atau jumlah *cluster* yang dihasilkan. Metode yang diimplementasikan adalah metode Elbow, dimana pada metode ini akan dihasilkan sebuah grafik dua dimensi. Pada sumbu X grafik tersebut menggambarkan jumlah K, sedangkan sumbu Y merepresentasikan nilai SSE yang diperoleh pada setiap K.

Umumnya, secara bertahap nilai SSE akan menurun seiring dengan naiknya nilai K.

Dalam penentuan nilai K ini, rentang K yang digunakan adalah 2 hingga 10. Kemudian ditampilkan sebuah grafik yang merepresentasikan nilai SSE pada setiap nilai K. Selanjutnya adalah menentukan nilai K terbaik dengan memperhatikan gambaran grafik yang dihasilkan. Kode yang digunakan untuk proses ini ditampilkan pada Script 5.3.

```
input <- data.frame(HasilRFMBobot$R_Bobot,Hasil
RFMBobot$F_Bobot,HasilRFMBobot$M_Bobot)
k.max <- 10
elbow <- sapply(2:k.max, function(k)
  {cmeans(input, k, 100,verbose = TRUE,
method = "cmeans",
weights = 0.5)$withinerror})
plot(2:k.max, elbow,type="b", pch = 19,
xlab="Jumlah K",
ylab="Sum of Squares")
```

Script 5.3 Penentuan Jumlah K

5.4.2 Clustering dengan Fuzzy C-Means

Setelah memperoleh nilai K terbaik, tahap selanjutnya adalah *clustering*. Dataset yang digunakan sama seperti tahap sebelumnya, yaitu dataset nilai variabel RFM semua pelanggan yang telah dikalikan dengan bobot setiap variabel. Metode *clustering* yang dijalankan *Fuzzy C-Means*. Setelah dataset dimasukkan, proses *clustering* dijalankan dengan menggunakan nilai K yang dipilih. Metode *Fuzzy C-Means* akan menghasilkan derajat keanggotaan tiap pelanggan terhadap setiap *cluster*. Hasil *clustering* kemudian disimpan dalam format file *csv*. Kode yang digunakan untuk proses ini ditampilkan pada Script 5.4.

```

> input <- data.frame(HasilRFMBobot$R_Bobot, Ha
silRFMBobot$F_Bobot, HasilRFMBobot$M_Bobot)
> cl<-cmeans(input, 3, 20,verbose = TRUE, rate.
par = 0.5, method = "cmeans", weights = 0.5)
> output <- data.frame(HasilRFMBobot$Customer_I
D,cl$membership,cl$cluster)

> write.csv(output,"D:/TA/Data/Hasilklastter.csv
")

```

Script 5.4 Proses Clustering

5.5 Pengujian Kualitas *Clustering*

Setelah *clustering* dilakukan, dihasilkan *cluster* atau segmen yang membagi semua pelanggan kedalamnya. Untuk mengetahui apakah jumlah segmen yang ditentukan sudah optimal, diperlukan uji kualitas *clustering*. Pengujian ini mengimplementasikan metode SSE. Data yang diperlukan adalah data nilai variabel RFM semua pelanggan yang telah dikalikan dengan bobot setiap variabel. Algoritma SSE menghitung kuadrat selisih setiap objek data pelanggan dengan titik pusat *clusternya*, kemudian menjumlahkan total keseluruhan. Perhitungan SSE ditampilkan dengan grafik yang menunjukkan nilai SSE untuk setiap jumlah *cluster*. Proses perhitungan SSE menggunakan Script 5.5.

```

input <- data.frame(HasilRFMBobot$R_Bobot,
HasilRFMBobot$F_Bobot,
HasilRFMBobot$M_Bobot)

cl<-cmeans(input, 3, 20, method = "cmeans",
verbose = TRUE, rate.par = 0.5, weights = 0.5)

print(cl$withinerror)

```

Script 5.5 Perhitungan SSE

5.6 Pembuatan Visualisasi

Proses visualisasi dilakukan menggunakan aplikasi Rshiny. Package yang perlu dijalankan dalam aplikasi yaitu package shiny dan untuk desainnya mengguakan package shinydashboard. Penjelasan lebih rinci mengenai isi dari setiap halaman akan dijelaskan pada setiap aktivitas berikut.

5.6.1 Halaman Utama

Halaman utama menampilkan dua grafik, yaitu *bar chart* dan *pie chart*. *Bar chart* menggambarkan jumlah pelanggan setiap segmen sedangkan *pie chart* menggambarkan prosesntase pelanggan setiap segmen.

5.6.1.1 Tampilan Bar Chart

Untuk menampilkan grafik *bar chart*, dibutuhkan data masukan hasil *clustering*. Data tersebut kemudian akan diagregasikan berdasarkan kelompok *cluster*. *Bar chart* digambarkan dengan sumbu X berupa jenis *cluster*, sedangkan sumbu Y berupa jumlah pelanggan dalam *cluster* tersebut. *Package* yang dibutuhkan adalah package *plotly*. Proses visualisasi *bar chart* dapat dilihat pada Script 5.6.

```
#Data Masukan
segmen <- c("Segmen 1","Segmen 2",
          "Segmen 3")
value <- c(sum(data$c1.cluster == "1"),
           sum(data$c1.cluster == "2"),
           sum(data$c1.cluster == "3"))
#Bar Chart
output$barchart <- renderPlotly ({
  plot_ly(x = segmen,
           y = value,
           name = "Jenis Segmen",
           type = "bar",
           hoverinfo = 'text',
           text = paste(value,
                         ' Pelanggan'),
           marker = list(color= warna)
  )})
#Kode Warna
warna <- c("#00FFFF", "#8A2BE2", "#E9967A")
```

Script 5.6 Visualisasi Bar Chart

5.6.1.2 Tampilan Pie Chart

Untuk menampilkan grafik *bar chart*, dibutuhkan data masukan hasil *clustering*. Data tersebut kemudian akan diagregasikan berdasarkan kelompok *cluster*. *Bar chart* digambarkan dengan sumbu X berupa jenis *cluster*, sedangkan sumbu Y berupa jumlah pelanggan dalam *cluster* tersebut. *Package* yang dibutuhkan adalah package *plotly*. Proses visualisasi *bar chart* dapat dilihat pada Script 5.7.

```

#Data Masukan
segmen <- c("Segmen 1","Segmen 2",
          "Segmen 3")
value <- c(sum(data$c1.cluster == "1"),
           sum(data$c1.cluster == "2"),
           sum(data$c1.cluster == "3"))

#Pie Chart
output$piechart <- renderPlotly ({
  plot_ly(labels = segmen,
          values = value,
          type = 'pie',
          textposition = 'inside',
          textinfo = 'label+percent',
          insidetextfont = list(color =
            '#FFFFFF'),
          hoverinfo = 'text',
          text = paste(value,
                      ' Pelanggan'),
          marker = list(colors = warna,
                        line = list(color =
                          '#FFFFFF',
                          width = 1))
)
})

#Kode Warna
warna <- c("#00FFFF", "#8A2BE2", "#E9967A")

```

Script 5.7 Visualisasi Pie Chart

5.6.2 Halaman Visualisasi Segmen

Halaman visualisasi segmen menampilkan informasi rata-rata variabel RFM, rentang nilai variabel RFM, pelanggan teratas dan terbawah, grafik *scater plot 3D*, tabel peringkat pelanggan, dan grafik *box plot* variabel RFM setiap segmen yang dipilih. Masing-masing tampilan akan dijelaskan lebih detail pada poin berikut.

5.6.2.1 Tampilan rata-rata variabel RFM

Rata-rata variabel RFM ditampilkan dalam bentuk kotak-kotak informasi terpisah. Data yang diperlukan untuk menghitung rata-rata variabel RFM adalah nilai RFM masing-masing pelanggan yang sudah dikalikan dengan bobot variabel. Proses visualisasi rata-rata variabel RFM dapat dilihat pada Script 5.8.

```
#Rata-rata Recency
output$recency <- renderInfoBox({
  data <- pilihdata()infoBox(
    title = "Recency Average",
    value = paste(round(mean(data$R)), "hari"),
    subtitle = "Terakhir transaksi",
    icon = icon("calendar"),
    color = "navy",
    fill = TRUE) })

#Rata-rata Frequency
output$frequency <- renderInfoBox({
  data <- pilihdata()infoBox(
    title = "Frequency Average",
    value = paste(round(mean(data$F)), "kali"),
    subtitle = "Banyak transaksi",
    icon = icon("area-chart"),
    color = "navy",
    fill = TRUE) })

#Rata-rata Monetary
output$monetary <- renderInfoBox({
  data <- pilihdata()infoBox(
    title = "Monetary Average",
    value = paste("Rp", round(mean(data$M))),
    subtitle = "Jumlah biaya",
    icon = icon("money"),
    color = "navy",
    fill = TRUE) })
```

Script 5.8 Visualisasi rata-rata variabel RFM

Hasil eksekusi Script 5.8 ditampilkan pada Gambar 5.1.



Gambar 5.1 Hasil Eksekusi Visualisasi Rata-Rata Variabel RFM

5.6.2.2 Tampilan pelanggan teratas dan terbawah

Informasi pelanggan teratas dan terbawah ditampilkan dalam bentuk kotak informasi. Data yang dibutuhkan adalah data nilai RFM dikali bobot dan CLV masing-masing pelanggan. Pelanggan teratas diperoleh dari nilai CLV tertinggi pada segmen, sedangkan pelanggan terbawah diperoleh dari nilai CLV terendah pada segmen. Proses visualisasi rata-rata variabel RFM dapat dilihat pada Script 5.9.

```
#Pelanggan Teratas
output$baik<- renderValueBox({
  inputan <- pilihdata()
  cari <- sqldf("Select Customer_ID From
    (select Customer_ID,max(CLV)
     From inputan)")
  valueBox(value=tags$h5("PELANGGAN TERATAS"),
    subtitle = tags$h4(tags$strong (cari)),
    icon = icon("thumbs-o-up"),
    color = "navy"))
}

#Pelanggan Terbawah
output$buruk<- renderValueBox({
  inputan <- pilihdata()
  cari <- sqldf("Select Customer_ID From (sel
  ect Customer_ID,min(clv) From inputan)")
  valueBox(value=tags$h5("PELANGGAN TERBAWAH"),
    subtitle = tags$h4(tags$strong (cari)),
    icon = icon("thumbs-o-down"),
    color = "navy")})
```

Script 5.9 Visualisasi Pelanggan Teratas dan Terbawah

5.6.2.3 Tampilan grafik scatter plot 3 dimensi

Grafik *scatter plot* ditampilkan dengan bentuk kubus 3 dimensi yang mewakili tiga variabel. Sumbu X, sumbu Y, dan sumbu Z masing-masing mewakili variabel *recency*, *frequency*, dan *monetary*. Setiap titik dalam kubus menggambarkan setiap pelanggan yang dipetakan berdasarkan nilai RFM yang dimiliki. Anggota segmen memiliki warna titik yang berbeda dengan anggota segmen lain. Kubus dapat diputar secara tiga dimensi untuk melihat dari sudut pandang yang lain. Data yang diperlukan untuk visualisasi *scatter plot* ini adalah data nilai RFM dikali bobot masing-masing pelanggan dan cluster pelanggan berada. Proses visualisasi grafik *scatter plot* 3D dapat dilihat pada Script 5.10.

```
#Scatter Plot 3 Dimensi
output$scatter <- renderRglwidget ({
  splot <- pilihdata()
  open3d()
  plot3d(x=splot$R, y=splot$F, z=splot$M,
          col = warna[splot$cl.cluster],
          type = 's', size = 1,
          xlab = "recency",
          ylab = "frequency",
          zlab = "monetary",
          colkey = list(length = 0.5,
                        width = 0.5,
                        cex.lab = 0.75)
  )
  scene1 <- scene3d()
  rgl.close()
  save <- options(rgl.inShiny = TRUE)
  on.exit(options(save))
  rglwidget(scene1) })
```

Script 5.10 Visualisasi Grafik Scatter Plot 3D

5.6.2.4 Tampilan tabel peringkat pelanggan

Tabel peringkat pelanggan ditampilkan dalam bentuk tabel yang telah terurut berdasarkan nilai CLV pelanggan dari terbesar hingga terendah pada setiap segmennya. Data yang dibutuhkan adalah data pelanggan dan nilai CLV masing-masing. Proses visualisasi tabel peringkat pelanggan dapat dilihat pada Script 5.11.

```
#Tabel peringkat
output$clv <- renderDataTable({
  tabel <- pilihdata()
  tampil <- data.frame(sqlpdf(
    "Select Customer, R, F, M, CLV
     From tabel
     order by CLV DESC")
  )

  colnames(tampil) <- c("Customer",
                        "Recency",
                        "Frequency",
                        "Monetary",
                        "CLV")
  },
  list(lengthMenu = c(5, 10, 15),
       pageLength = 8))
```

Script 5.11 Visualisasi Tabel Peringkat Pelanggan

Hasil eksekusi Script 5.11 ditampilkan pada Gambar 5.2.

Peringkat Pelanggan					
	Show 5 entries	Search:			
Customer	Recency	Frequency	Monetary	CLV	
PT. POLYTAMA PROPINDO	96	8	861123629	0.6895672	
PT. MITSUBISHI CHEMICAL INDONESIA	119	6	816972362	0.6368488	
PT. INDRAMA PETROCHEMICALS	112	5	751214097	0.5913139	
PT. PATRA TRADING	97	60	94075983	0.4940676	
PT. CHANDRA ASRI PETROCHEMICAL TBK	108	4	510237144	0.4291341	
PT. REJO MULYA REJEKI	99	48	58437193	0.3983452	
PT. BONGKARAN PETROKIMIA	110	51	35370834	0.3953532	
PT. SUMBER KIMIA PARUJAKAN SAKTI	99	50	17893680	0.3835925	

Showing 1 to 8 of 123 entries

...

Gambar 5.2 Hasil Eksekusi Visualisasi Tabel Peringkat Pelanggan

5.6.2.5 Tampilan grafik box plot

Grafik *box plot* ditampilkan dalam bentuk *box plot* yang menggambarkan nilai matematis tiap variabel dan CLV setiap segmen. *Box plot* mendeskripsikan rata-rata, nilai minimal, nilai maksimal, kuartil I, dan kuartil III tiap variabel tiap segmen. *Box plot* tiap segmen memiliki warna berbeda antar satu dengan yang lain. Data yang dibutuhkan adalah data nilai RFM dikali bobot tiap pelanggan dan hasil *cluster*. Proses visualisasi grafik *box plot* dapat dilihat pada Script 5.12 dan 5.13.

```

#Data box plot
y1 <- subset(data,data$c1.cluster==1)
y2 <- subset(data,data$c1.cluster==2)
y3 <- subset(data,data$c1.cluster==3)

#Box Plot
output$boxclv <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = y1$CLV, name = "1",
                marker = list(color = "#00FFFF"),
                line = list(color = "#00FFFF")) %>%
    add_boxplot(y = y2$CLV, name = "2",
                marker = list(color = "#8A2BE2"),
                line = list(color = "#8A2BE2")) %>%
    add_boxplot(y = y3$CLV, name = "3",
                marker = list(color = "#E9967A"),
                line = list(color = "#E9967A"))
})

output$boxr <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = y1$R, name = "1",
                marker = list(color = "#00FFFF"),
                line = list(color = "#00FFFF")) %>%
    add_boxplot(y = y2$R, name = "2",
                marker = list(color = "#8A2BE2"),
                line = list(color = "#8A2BE2")) %>%
    add_boxplot(y = y3$R, name = "3",
                marker = list(color = "#E9967A"),
                line = list(color = "#E9967A"))
})

```

Script 5.12 Visualisasi Grafik Box-Plot (1)

```

output$boxf <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = y1$F, name = "1",
                marker = list(color = "#00FFFF"),
                line = list(color = "#00FFFF")) %>%
    add_boxplot(y = y2$F, name = "2",
                marker = list(color = "#8A2BE2"),
                line = list(color = "#8A2BE2")) %>%
    add_boxplot(y = y3$F, name = "3",
                marker = list(color = "#E9967A"),
                line = list(color = "#E9967A"))
  })

output$boxm <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = y1$M, name = "1",
                marker = list(color = "#00FFFF"),
                line = list(color = "#00FFFF")) %>%
    add_boxplot(y = y2$M, name = "2",
                marker = list(color = "#8A2BE2"),
                line = list(color = "#8A2BE2")) %>%
    add_boxplot(y = y3$M, name = "3",
                marker = list(color = "#E9967A"),
                line = list(color = "#E9967A"))
  })

```

Script 5.13 Visualisasi Grafik Box-Plot (2)

5.6.3 Halaman Tabel

Halaman tabel data menampilkan informasi tabel detail peringkat seluruh pelanggan. Tabel ini memuat informasi nilai *recency*, *frequency*, *monetary*, CLV, dan segmen tiap

pelanggan. Proses visualisasi tabel data dapat dilihat pada Script 5.14.

```
#Data Tabel Total
output$table = renderDataTable({
  tabel <- data
  tampil <- data.frame(
    tabel$Customer_ID,
    tabel$R,
    tabel$F,
    tabel$M,
    tabel$CLV,
    tabel$c1.cluster
  )
  colnames(tampil) <- c(
    "Customer Code",
    "Recency",
    "Frequency",
    "Monetary",
    "CLV",
    "Segmen"))
  list(lengthMenu= c(5, 10, 15),
  pageLength= 10)})}
```

Script 5.14 Visualisasi Tabel Data

Hasil eksekusi Script 5.14 ditampilkan pada Gambar 5.3.

Customer Code	Recency	Frequency	Monetary	CLV	Segment
2053	97	60	94075983.193	0.494067600995258	3
2138	93	11	4731484	0.146366832783448	2
2166	98	10	356683171.616	0.369051124219629	3
700032	98	27	10145411.867	0.242393206690916	3
700189	204	1	200000	0.0268713074498163	3
700286	100	7	10517521.843	0.122805009384089	3
700385	160	4	2750129.684	0.0690080033363687	3
700493	250	1	151084.7	0.00316905027243236	3
700503	121	11	12610553.377	0.137134696905383	1
700529	98	16	3021794.812	0.172371772325963	3

Customer Code Recency Frequency Monetary CLV Segment

Showing 1 to 10 of 123 entries

Previous 1 2 3 4 5 ... 13 Next

Gambar 5.3 Hasil Eksekusi Visualisasi Tabel Data

Halaman ini sengaja dikosongkan

BAB VI

HASIL PEMBAHASAN

Pada bab ini akan menjelaskan hasil dan pembahasan dari proses pembobotan RFM, perhitungan CLV, *clustering Fuzzy C-Means*, dan uji performa clustering. Bab ini juga berisi analisis hasil proses *clustering* dan visualisasi hasil *clustering*.

6.1 Pembobotan RFM dengan AHP

Pembobotan RFM dilakukan dengan mengolah data hasil kuesioner yang telah direkap pada *excel*.

Tabel 6.1 Hasil Kuesioner Responden 1

Kriteria	Recency	Frequency	Monetary
Recency	1	0,33	0,20
Frequency	3	1	0,50
Monetary	5	2	1

Tabel 6.2 Hasil Kuesioner Responden 2

Kriteria	Recency	Frequency	Monetary
Recency	1	0,20	0,17
Frequency	5	1	0,33
Monetary	6	3	1

Tabel 6.3 Hasil Kuesioner Responden 3

Kriteria	Recency	Frequency	Monetary
Recency	1	0,33	0,25
Frequency	3	1	0,50
Monetary	4	2	1

Data kuesioner kemudian diolah dengan cara merata-rata semua pengisian responden terhadap perbandingan tiga kriteria RFM. Hasil rata-rata ini dijumlahkan setiap kolomnya. Hasil rata-rata dapat dilihat pada Tabel 6.1.

Tabel 6.4 Hasil Rata-Rata Penilaian Kuesioner

Kriteria	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
<i>Recency</i>	1	0,29	0,2
<i>Frequency</i>	3,67	1	0,44
<i>Monetary</i>	5	2,33	1
<i>Jumlah</i>	9,67	3,62	1,65

Setelah itu dilakukan normalisasi matriks pada hasil rata-rata nilai setiap kriteria. Tabel baru yang dihasilkan tersebut disebut sebagai *normal comparisons matrix*. Hasil normalisasi matriks atau *normal comparisons matrix* dapat dilihat pada tabel 6.2. dari tabel ini didapatkan nilai bobot untuk setiap kriteria.

Tabel 6.5 Normal Comparisons Matrix

Kriteria	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>	Bobot
<i>Recency</i>	0.10	0.08	0.12	0.10
<i>Frequency</i>	0.38	0.28	0.27	0.31
<i>Monetary</i>	0.52	0.64	0.61	0.59

Bobot yang diperoleh perlu dilakukan uji konsistensi untuk mengetahui apakah kuesioner yang diisi secara konsisten atau tidak. Uji konsistensi dilakukan dengan cara mengalikan setiap nilai perbandingan pada *normal comparisons matrix* dengan bobot masing-masing kriteria. Lalu setiap baris kriteria dijumlahkan menjadi *overall score*. Hasil *overall score* dapat dilihat pada Tabel 6.3.

Tabel 6.6 Overall Score Bobot

Kriteria	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>	Overall Score
<i>Recency</i>	0.103	0.089	0.121	0.313
<i>Frequency</i>	0.376	0.308	0.262	0.946
<i>Monetary</i>	0.513	0.719	0.589	1.821

Selanjutnya, dicari nilai x dengan merata-rata hasil pembagian *overall score* dengan nilai bobot masing-masing kriteria.

$$Recency = \frac{0.313}{0.10} = 3.048$$

$$Frequency = \frac{0.946}{0.31} = 3.069$$

$$Monetary = \frac{1.821}{0.59} = 3.091$$

$$x = \frac{3.048 + 3.069 + 3.091}{3} = 3.067$$

Setelah mengetahui nilai x yang menjadi masukan untuk menghitung CI, maka dengan mengacu pada Bab II, sehingga kriteria tersebut memiliki CI sebesar :

$$CI = \frac{3.067 - 3}{3 - 1} = 0.035$$

Nilai CI yang diperoleh adalah 0.035, berdasarkan hukum uji konsistensi, karena nilai CI tidak sama dengan nol, maka dibutuhkan uji ketidakkonsistensian. Mengacu pada Bab II, nilai ketidakkonsistensian dihitung dengan membagi nilai CI dengan RI. Nilai RI diperoleh dari Tabel 2.6 pada Bab II, dan diperoleh nilai n = 3, maka nilai RI adalah 0.58. Sehingga nilai CI/RI yang diperoleh adalah :

$$\frac{0.035}{0.58} = 0.060$$

Karena nilai CI/RI adalah 0.060 atau kurang dari 0.1, maka hasil kuesioner telah bersifat konsisten. Oleh karena itu, bobot kriteria telah valid dan dapat digunakan pada perhitungan CLV. Adapun bobot masing-masing kriteria adalah 0.10, 0.31, dan

0.59 berturut-turut untuk *recency*, *frequency*, dan *monetary*. Sehingga kriteria *monetary* memiliki bobot terbesar dibanding kriteria lain.

6.2 Perhitungan CLV

Dari nilai bobot yang diperoleh, selanjutnya dilakukan perhitungan CLV. Perhitungan CLV dilakukan dengan mengalikan masing-masing bobot dengan nilai RFM masing-masing pelanggan, kemudian menjumlahkan ketiga hasil perkalian tersebut. Sebagian hasil perhitungan CLV dapat dilihat pada Tabel 6.4 dan hasil lengkapnya pada Lampiran E.

Tabel 6.7 Sebagian Hasil Perhitungan CLV Pelanggan

No	Customer ID	R x Bobot	F x Bobot	M x Bobot	CLV
1	2053	0.08182	0.35047	0.06178	0.49407
2	2138	0.08387	0.05940	0.00309	0.14637
3	2166	0.08130	0.05346	0.23429	0.36905
4	700032	0.08130	0.15444	0.00665	0.24239
5	700189	0.02676	0	0.00011	0.02687
6	700286	0.08027	0.03564	0.00689	0.12281
7	700385	0.04940	0.01782	0.00179	0.06901
8	700493	0.00309	0	0.00008	0.00317
9	700503	0.06947	0.05940	0.00827	0.13713

6.3 Proses Clustering

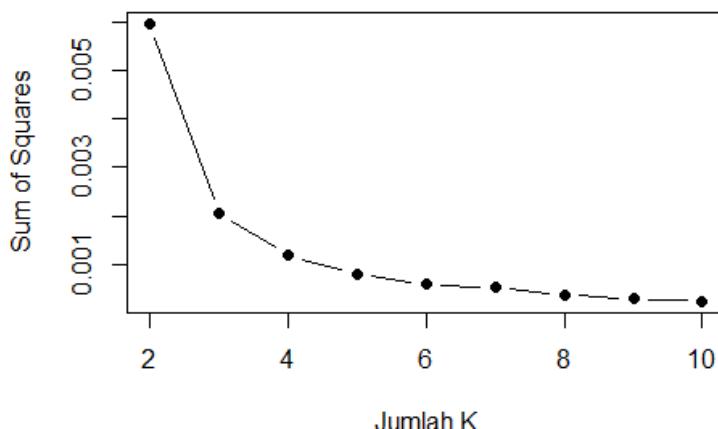
Dalam proses ini dijelaskan mengenai hasil dari penentuan nilai K dengan metode Elbow dan *clustering* dengan metode *Fuzzy C-Means*.

6.3.1 Penentuan Nilai K

Penentuan nilai K dilakukan menggunakan metode Elbow pada RStudio. Hasil perhitungan metode Elbow, ditampilkan dalam plot dua dimensi pada Gambar 6.1 . Nilai pada sumbu X adalah jumlah K yaitu 2 hingga 10, sedangkan sumbu Y adalah nilai SSE yang dihasilkan oleh setiap nilai K. Metode Elbow menjelaskan bahwa nilai K yang diambil adalah pada titik dimana terjadi penurunan yang signifikan dan diikuti oleh nilai yang relative konstan. Pada plot tersebut dapat dilihat bahwa seiring naiknya nilai K, maka nilai SSE akan semakin menurun. Dapat dilihat bahwa titik yang dimaksud adalah 3 atau 4. Akan tetapi pada tabel 6.5, dapat dilihat bahwa penurunan SSE pada $K = 3$ lebih signifikan dibanding dengan $K = 4$, sehingga nilai 3 dipilih sebagai nilai K untuk proses *clustering*.

Tabel 6.8 Selisih Penurunan Nilai Error

K	Selisih error
2	0
3	0,0039072440
4	0,0008537139
4	0,0003812762
5	0,0001778256
6	0,0001843487
7	0,0000731956
8	0,0000654441
9	0,0000547865



Gambar 6.1 Plot Penentuan Nilai K dengan Metode Elbow

6.3.2 Clustering Fuzzy C-Means

Proses *Clustering Fuzzy C-Means* dilakukan pada RStudio. Luaran yang dihasilkan dari eksekusi *script* adalah pusat setiap *cluster*, derajat keanggotaan setiap objek terhadap tiap *cluster*, dan *cluster* tiap objek berdasarkan derajat keanggotaannya yang paling besar terhadap *cluster* tertentu. *Cluster* setiap objek akan disimpan dalam bentuk file .csv agar mempermudah dalam proses visualisasi hasil *clustering*. Dalam menentukan parameter yang paling optimal, dilakukan percobaan dengan mengubah nilai 3 parameter. Parameter tersebut adalah *MaxIter*, *Learning rate*, dan *Weights*. Sebagian hasil percobaan penentuan parameter dengan $K = 3$ dapat dilihat pada Tabel 6.5 dan lebih lengkapnya dapat dilihat pada Lampiran F.

Tabel 6.9 Percobaan Terhadap Perubahan Nilai Parameter

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
16	0,3	0,5	0,0046780651	0,0015782925
17	0,3	0,5	0,0046779366	0,0015782639
18	0,3	0,5	0,0046778290	0,0015782528
19	0,3	0,5	0,0046777379	0,0015782485
20	0,3	0,5	0,0046776599	0,0015782469
16	0,4	0,5	0,0046688554	0,0015782854
17	0,4	0,5	0,0046668182	0,0015782611
18	0,4	0,5	0,0046640166	0,0015782517
19	0,4	0,5	0,0046600555	0,0015782481
20	0,4	0,5	0,0046542470	0,0015782467
16	0,5	0,5	0,0046777017	0,0015782470
17	0,5	0,5	0,0046776287	0,0015782463
18	0,5	0,5	0,0046775653	0,0015782460
19	0,5	0,5	0,0046775099	0,0015782459
20	0,5	0,5	0,0046774611	0,0015782459
16	0,6	0,5	0,0017450240	0,0015782510
17	0,6	0,5	0,0017450215	0,0015782478
18	0,6	0,5	0,0017450206	0,0015782466
19	0,6	0,5	0,0017450202	0,0015782462
20	0,6	0,5	0,0017450201	0,0015782460

Dari tabel tersebut dapat diketahui bahwa perubahan parameter *learning rate* dan *weights* tidak berpengaruh besar terhadap nilai SSE testing. Parameter yang lebih mempengaruhi nilai SSE adalah *max iter*. Nilai SSE terkecil terbentuk ketika parameter *max iter* mencapai 20 ke atas. Sehingga parameter yang digunakan adalah 20 untuk *max iter*, 0,5 untuk *learning rate*, dan 0,5 untuk *weights*. Hasil *clustering* dengan menggunakan parameter yang telah ditentukan dapat dilihat pada Gambar 6.2.

```
Fuzzy c-means clustering with 3 clusters

Cluster centers:
  HasilRFMBobot.R_Bobot  HasilRFMBobot.F_Bobot  HasilRFMBobot.M_Bobot
1          0.09794529          0.18147671          0.022682262
2          0.07708765          0.02399898          0.009440328
3          0.09262034          0.02681061          0.515592341

Memberships:
      1           2           3
[1,] 0.7848373314 0.1657460765 4.941659e-02
[2,] 0.0795906009 0.9152395124 5.169887e-03
[3,] 0.3225459215 0.3859138196 2.915403e-01
[4,] 0.8418054324 0.1509374538 7.257114e-03
[5,] 0.0646414480 0.9263342751 9.024277e-03
[6,] 0.0216197147 0.9764754053 1.904880e-03
[7,] 0.0135796745 0.9849176795 1.502646e-03
[8,] 0.1225209454 0.8585836838 1.889537e-02
[9,] 0.0478566324 0.9489743437 3.169024e-03
[10,] 0.2385606253 0.7515369131 9.902462e-03

Closest hard clustering:
 [1] 1 2 2 1 2 2 2 2 2 1 1 1 2 2 2 2 1 2 2 2 1 3 2 3 3 2 3 2 2
2 2 1 1 2 2 2 2 1 2 2 2 2 2 1 1
[48] 1 2 2 2 2 1 2 2 2 2 2 1 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[95] 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
```

Gambar 6.2 Hasil clustering Fuzzy C-Means untuk 3 cluster

Derajat keanggotaan setiap pelanggan dihitung dengan rumus 5 pada Bab II. Misalkan titik X, dengan R = 0.082, F = 0.35, M = 0.062, sedangkan m = 1 dan w = 0.5. Untuk iterasi awal, titik pusat *cluster* dipilih secara random, titik pusat *cluster* 1 adalah 0.03, 0, 0; *cluster* 2 adalah 0.08, 0.04, 0.01; dan *cluster* 3 adalah 0.07, 0.06, 0.01. sebelum menghitung membership degree, dihitung jarak titik dengan pusat *cluster* dengan rumus *Euclidean*. Jarak titik X dengan *cluster* 1 adalah :

$$\begin{aligned}
 & (X_{11} - V_{11}) \\
 &= \sqrt{\left((0.08 - 0.03) \right)^2 + \left((0.08 - 0) \right)^2 + \left((0.08 - 0) \right)^2} \\
 &= \sqrt{0.0025 + 0.0064 + 0.0064} \\
 &= 0.203635331
 \end{aligned}$$

Jarak titik X dengan *cluster* 2 adalah :

$$\begin{aligned}
 & (X_{11} - V_{12}) \\
 &= \sqrt{((0.08 - 0.08))^2 + ((0.08 - 0.04))^2 + ((0.08 - 0.01))^2} \\
 &= \sqrt{0 + 0.0016 + 0.0049} \\
 &= 0,360088897
 \end{aligned}$$

Jarak titik X dengan *cluster* 2 adalah :

$$\begin{aligned}
 & (X_{11} - V_{13}) \\
 &= \sqrt{((0.08 - 0.07))^2 + ((0.08 - 0.06))^2 + ((0.08 - 0.01))^2} \\
 &= \sqrt{0.001 + 0.004 + 0.0049} \\
 &= 0,296205313
 \end{aligned}$$

Setelah ketiga jarak diketahui, kemudian diimplementasikan rumus berikut :

$$u_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}; i = 1; k = 1, 2, 3$$

Untuk hasil *membership degree* terhadap *cluster* 1 adalah :

$$u_{11} = \frac{[0.203635331^2]^{\frac{-1}{0.5-1}}}{\sum_{k=1}^3 \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{0.5-1}}}$$

$$u_{11} = 0,065556$$

Untuk hasil *membership degree* terhadap *cluster* 2 adalah :

$$u_{12} = \frac{[0,360088897^2]^{\frac{-1}{0.5-1}}}{\sum_{k=1}^3 \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{0.5-1}}}$$

$$u_{12} = 0,64097$$

Untuk hasil *membership degree* terhadap *cluster* 3 adalah :

$$u_{13} = \frac{[0,296205313^2]^{\frac{-1}{0.5-1}}}{\sum_{k=1}^3 \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{0.5-1}}}$$

$$u_{13} = 0,293474$$

Derajat keanggotaan akan terus diperbarui hingga iterasi berhenti saat fungsi objektif = batas *error*. *Cluster* akan dipilih berdasarkan nilai derajat keanggotaan yang paling besar, sehingga titik X termasuk ke dalam *cluster* 2 untuk iterasi pertama.

6.4 Pengujian Kualitas *Clustering*

Pengujian kualitas dilakukan untuk mengetahui seberapa baik kualitas hasil *clustering* yang dihasilkan. Pengujian dilakukan dengan validasi derajat keanggotaan dan SSE.

6.4.1 Validasi Derajat Keanggotaan

Setiap objek hasil *clustering* memiliki derajat keanggotaan terhadap tiap cluster. Derajat keanggotaan ini menunjukkan seberapa besar kedekatan objek dengan cluster tertentu. Semakin besar nilai keanggotaan suatu objek terhadap cluster

tertentu, semakin dekat objek masuk ke dalam *cluster* tersebut. Pemilihan *cluster* dipilih dari nilai keanggotaan paling besar. Validasi dilakukan dengan membandingkan *cluster* dimana nilai keanggotaan terbesar dengan hasil *clustering*. Sebagian hasil validasi derajat keanggotaan dapat dilihat pada Tabel 6.5 dan lebih lengkapnya dapat dilihat pada Lampiran G. Dari keseluruhan hasil *clustering*, 123 objek telah valid atau berada pada cluster dengan nilai keanggotaan terbesar.

Tabel 6.10 Sebagian Hasil Validasi Derajat Keanggotaan

<i>STP</i>	<i>C 1</i>	<i>C 2</i>	<i>C 3</i>	<i>Max</i>	<i>Cluster</i>	<i>Validasi</i>
2053	0.1018	0.1343	0.7639	0.7639	3	Valid
2138	0.0703	0.9164	0.0133	0.9164	2	Valid
2166	0.3358	0.3685	0.2957	0.3685	2	Valid
700032	0.1485	0.3215	0.5300	0.5300	3	Valid
700189	0.8763	0.1089	0.0148	0.8763	1	Valid
700286	0.1089	0.8818	0.0092	0.8818	2	Valid
700385	0.8975	0.0965	0.0061	0.8975	1	Valid
700493	0.7520	0.2046	0.0434	0.7520	1	Valid
700503	0.0767	0.9135	0.0097	0.9135	2	Valid

6.4.2 Pengujian SSE

Pengujian SSE dilakukan menggunakan aplikasi RStudio. Hasil pengujian SSE menampilkan nilai *error* pada hasil *clustering* berdasarkan parameter yang telah digunakan. Nilai SSE yang diperoleh untuk keseluruhan data adalah 0.002047527. Hasil pengujian SSE dapat dilihat pada Gambar 6.3.

6.5 Analisis Hasil Clustering

Proses *clustering* menghasilkan tiga segmen dengan 123 pelanggan yang menjadi anggotanya. Masing-masing segmen memiliki karakteristik yang berbeda. Karakteristik setiap segmen dipengaruhi oleh variabel *recency*, *frequency*, dan *monetary*. Ketiga variabel tersebut juga mempengaruhi nilai

CLV tiap pelanggan. Nilai CLV inilah yang menjadi ukuran tingkat loyalitas setiap pelanggan tiap segmen.

```

Iteration: 5, Error: 0.0038256081
Iteration: 6, Error: 0.0038122805
Iteration: 7, Error: 0.0037980931
Iteration: 8, Error: 0.0037739037
Iteration: 9, Error: 0.0037224473
Iteration: 10, Error: 0.0035983981
Iteration: 11, Error: 0.0032871847
Iteration: 12, Error: 0.0027005930
Iteration: 13, Error: 0.0022598608
Iteration: 14, Error: 0.0020895367
Iteration: 15, Error: 0.0020531986
Iteration: 16, Error: 0.0020481995
Iteration: 17, Error: 0.0020476208
Iteration: 18, Error: 0.0020475464
Iteration: 19, Error: 0.0020475326
Iteration: 20, Error: 0.0020475287
> print(c1$withinerror)
[1] 0.002047529

```

Gambar 6.3 Hasil Pengujian SSE

6.5.1 Pemeringkatan Segmen

Pemeringkatan segmen dilakukan untuk menentukan segmen terbaik hingga terburuk. Tingkatan segmen ditentukan berdasarkan rata-rata nilai CLV pelanggan dalam satu segmen. Segmen yang memiliki nilai CLV terbesar menunjukkan bahwa segmen tersebut beranggotakan pelanggan yang loyal dan layak disebut sebagai segmen terbaik, begitu pula sebaliknya.

Tabel 6.11 Peringkat Segmen Berdasarkan CLV

<i>Segmen</i>	<i>CLV</i>	<i>Peringkat Segmen</i>
1	0.05370356	3
2	0.16737282	2
3	0.35156224	1

Tabel 6.7 merupakan hasil perhitungan rata-rata CLV tiap segmen. Dari nilai tersebut diperoleh bahwa segmen 3 merupakan segmen terbaik dan segmen 1 merupakan segmen yang terburuk.

6.5.2 Analisis Karakteristik Segmen 1

Segmen 1 memiliki anggota 42 pelanggan. Rata-rata nilai CLV segmen 1 adalah 0.054 atau yang paling rendah dari dua segmen lainnya, sehingga segmen 1 merupakan segmen terburuk. Rentang jumlah uang yang dikeluarkan oleh pelanggan antara 26.800 rupiah hingga 51.101.240, dengan rata-rata 6.561.256 rupiah. Hal ini menunjukkan bahwa pelanggan mengeluarkan uang yang relatif rendah pada setiap transaksi, meskipun juga ada pelanggan yang mengeluarkan uang dengan jumlah besar pada sekali transaksi.

Rentang jumlah transaksi yang dilakukan pelanggan antara 1 hingga 6 kali, dengan rata-rata 2 transaksi per pelanggan. Hal ini menunjukkan bahwa pelanggan jarang melakukan transaksi pembelian. Rentang waktu transaksi terakhir kali dilakukan adalah antara 110 hingga 256 hari. Dari hal ini diketahui bahwa pelanggan aktif melakukan transaksi rata-rata 4 hingga 9 bulan yang lalu. Jika dihitung mundur dari bulan Oktober, maka pelanggan terakhir kali melakukan transaksi pada bulan Januari hingga Mei 2017. Karakteristik segmen 1 lebih lengkap dapat dilihat pada Tabel 6.8.

Tabel 6.12 Karakteristik Pelanggan pada Segmen 1

Segmen	1
Peringkat Segmen	3
Jumlah Anggota	42
Karakteristik	Recency 110 - 256 hari (rata-rata : 173 hari)
	Frequency 1 - 6 kali (rata-rata : 2 kali)
	Monetary Rp. 26.800,- - Rp. 51.101.240,- (rata-rata : Rp. 6.561.256,-)

6.5.3 Analisis Karakteristik Segmen 2

Segmen 2 memiliki anggota 64 pelanggan. Rata-rata nilai CLV segmen 2 adalah 0.167 atau tertinggi kedua, sehingga segmen 2 dikategorikan sebagai segmen menengah. CLV yang sedang menggambarkan bahwa pelanggan pada segmen 2 cukup loyal terhadap perusahaan.

Rentang jumlah uang yang dikeluarkan oleh pelanggan antara 133.893 rupiah hingga 816.972.362 rupiah, dengan rata-rata 58.354.478 rupiah. Hal ini menunjukkan bahwa pelanggan mengeluarkan uang yang relatif tinggi pada setiap transaksi.

Rentang jumlah transaksi yang dilakukan pelanggan antara 2 hingga 25 kali, dengan rata-rata 10 transaksi per pelanggan. Hal ini menunjukkan bahwa pelanggan cukup sering melakukan transaksi pada perusahaan. Rentang waktu transaksi terakhir kali dilakukan adalah antara 93 hingga 122 hari. Dari hal ini diketahui bahwa pelanggan aktif melakukan transaksi rata-rata 4 bulan yang lalu. Jika dihitung mundur dari bulan Oktober, maka pelanggan terakhir kali melakukan transaksi pada bulan Juni 2017. Karakteristik segmen 2 lebih lengkap dapat dilihat pada Tabel 6.9.

Tabel 6.13 Karakteristik Pelanggan pada Segmen 2

Segmen	2	
Peringkat Segmen	2	
Jumlah Anggota	64	
Karakteristik	Recency	93 - 122 hari (rata-rata : 104 hari)
	Frequency	2 - 25 kali (rata-rata : 10 kali)
	Monetary	Rp. 133.893,- - Rp. 816.972.362,- (rata-rata : Rp. 58.354.478,-)

6.5.4 Analisis Karakteristik Segmen 3

Segmen 3 memiliki anggota 17 pelanggan. Rata-rata nilai CLV segmen 3 adalah 0.35 atau paling tinggi dari dua segmen lainnya, sehingga segmen 3 dikategorikan sebagai segmen terbaik. CLV yang tinggi menggambarkan bahwa pelanggan pada segmen 3 sangat loyal terhadap perusahaan. Rentang jumlah uang yang dikeluarkan oleh pelanggan antara 7.481.951 rupiah hingga 861.123.629 rupiah, dengan rata-rata 81.600.624 rupiah. Hal ini menunjukkan bahwa pelanggan mengeluarkan uang yang besar pada saat transaksi.

Rentang jumlah transaksi yang dilakukan pelanggan antara 8 hingga 60 kali, dengan rata-rata 38 transaksi per pelanggan. Hal ini menunjukkan bahwa pelanggan sangat sering melakukan transaksi pada perusahaan. Rentang waktu transaksi terakhir kali dilakukan adalah antara 93 hingga 110 hari. Dari hal ini diketahui bahwa pelanggan aktif melakukan transaksi rata-rata kurang dari 4 bulan yang lalu. Jika dihitung mundur dari bulan Oktober, maka pelanggan terakhir kali melakukan transaksi pada bulan Juni 2017. Karakteristik segmen 2 lebih lengkap dapat dilihat pada Tabel 6.10.

Tabel 6.14 Karakteristik Pelanggan pada Segmen 3

Segmen	3
Peringkat Segmen	1
Jumlah Anggota	17
Karakteristik	Recency 93 - 110 hari (rata-rata : 99 hari)
	Frequency 8 - 60 kali (rata-rata : 38 kali)
	Monetary Rp. 7.481.951 – Rp. 861.123.629 (rata-rata : Rp. 81.600.624,-)

6.5.5 Perbandingan Karakteristik Segmen

Perbandingan karakteristik ketiga segmen dapat dilihat pada Tabel 6.11. Dari tabel tersebut terlihat bahwa pada nilai variabel *recency* terdapat banyak *overlap* data, terbukti dengan adanya rentang nilai yang sama pada ketiga segmen. Hal ini dikarenakan variabel *recency* memiliki bobot terkecil, sehingga jarak nilai antar segmen tidak signifikan.

Pada nilai variabel *frequency*, terlihat mulai ada jarak yang signifikan antar segmen, meskipun terdapat *overlap* data. Hal ini disebabkan bobot *frequency* yang lebih besar daripada *recency*.

Variabel *monetary* yang memiliki bobot paling besar menyebabkan jarak nilai *monetary* antar segmen lebih terlihat signifikan. Meskipun ada overlap data, terlihat *monetary* segmen 1 paling rendah, segmen 2 menengah, dan segmen 3 paling tinggi.

Tabel 6.15 Perbandingan Karakteristik 3 Segmen

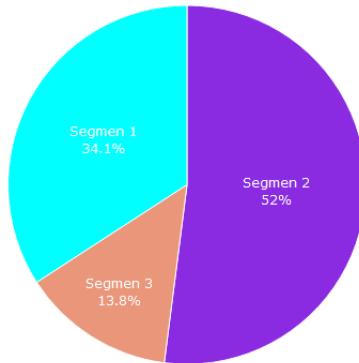
	Segmen 1	Segmen 2	Segmen 3
<i>Peringkat</i>	3	2	1
<i>Jumlah Anggota</i>	42	64	17
<i>Recency</i>	110 - 256 hari	93 - 122 hari	93 - 110 hari
<i>Frequency</i>	1 - 6 kali	2 - 25 kali	8 - 60 kali
<i>Monetary</i>	Rp. 26.800 – Rp. 51.101.240	Rp. 133.893 – Rp. 816.972.362	Rp. 7.481.951 – Rp. 861.123.629

6.6 Analisis Visualisasi Clustering

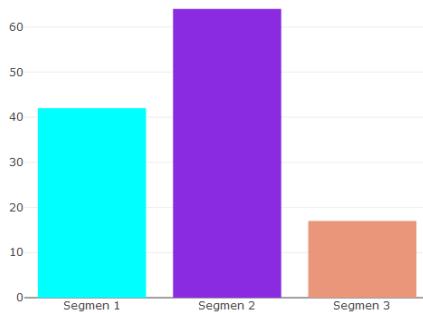
Hasil segmentasi ditampilkan dalam bentuk grafik untuk mempermudah dalam menganalisa perilaku pelanggan pada setiap segmen.

6.6.1 Grafik Bar Chart dan Pie Chart

Grafik bar chart dan pie chart ditampilkan untuk mengtahui perbandingan jumlah anggota setiap segmen.



Gambar 6.4 Grafik Pie Chart 3 Segmen



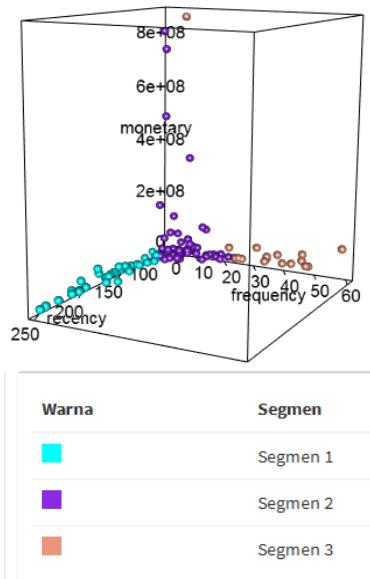
Gambar 6.5 Grafik Bar Chart 3 Segmen

Dari Gambar 6.4 dan 6.5 diketahui bahwa segmen 1 memiliki anggota 42 pelanggan atau sebanyak 34.2%. Segmen 2 memiliki anggota terbanyak, dengan 64 pelanggan atau 52%.

Segmen 3 memiliki anggota paling sedikit, dengan 17 pelanggan atau 13.8%.

6.6.2 Grafik Scatter Plot 3D

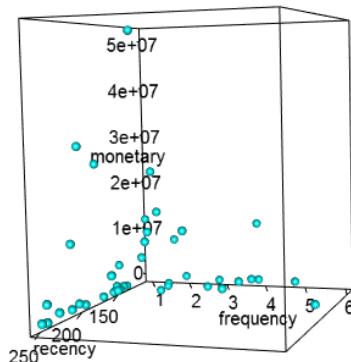
Grafik *Scatter Plot 3D* mendeskripsikan persebaran pelanggan pada ruang tiga dimensi berdasarkan variabel RFM. Sumbu X mewakili variabel *recency*, sumbu Y mewakili variabel *frequency*, dan sumbu Z mewakili variabel *monetary*. Segmen 1 digambarkan dengan warna biru, segmen 2 digambarkan dengan warna ungu, sedangkan segmen 3 digambarkan dengan warna cokelat. Hasil visualisasi grafik *Scatter Plot* ditampilkan pada Gambar 6.6.



Gambar 6.6 Grafik Scatter Plot 3D Keseluruhan Segmen

Dari Gambar tersebut dapat diketahui bahwa pelanggan pada segmen 1 cenderung berkumpul pada variabel *recency* dan mengarah pada nilai *recency* yang semakin tinggi. Pelanggan pada segmen 2 cenderung berkumpul mendekati titik 0 dan mengarah pada variabel *monetary* yang semakin tinggi. Pelanggan pada segmen 3 cenderung berkumpul pada variabel *frequency* dan mengarah ke nilai *frequency* yang semakin tinggi.

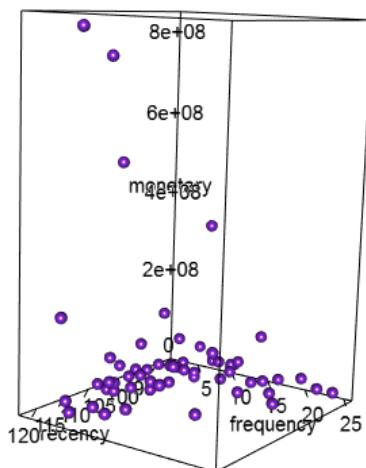
Gambar 6.7 menjelaskan lebih detail mengenai persebaran pelanggan pada segmen 1. Dari gambar tersebut dapat dilihat bahwa pelanggan sebagian besar menyebar pada nilai *recency* yang tinggi, nilai *monetary* yang relatif rendah, dan nilai *frequency* yang relatif rendah. Hal ini menunjukkan bahwa segmen 1 memiliki karakteristik *recency* yang tinggi atau pelanggan tidak melakukan transaksi dalam waktu yang lama.



Gambar 6.7 Grafik Scatter Plot 3D Segmen 1

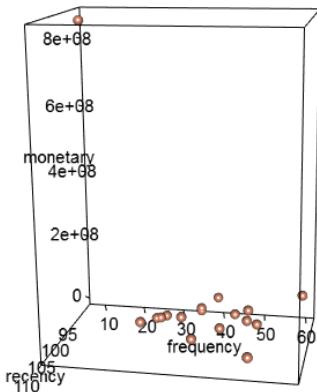
Gambar 6.8 menjelaskan lebih detail mengenai persebaran pelanggan pada segmen 2. Dari gambar tersebut dapat dilihat bahwa pelanggan sebagian besar menyebar pada nilai *recency*

yang cukup rendah, nilai *monetary* yang rendah, dan nilai *frequency* yang cukup tinggi. Hal ini menunjukkan bahwa segmen 2 memiliki karakteristik *frequency* yang cukup tinggi atau pelanggan cukup sering melakukan transaksi. Selain itu, sebagian besar pelanggan berkumpul pada nilai *monetary* yang rendah tidak banyak mengeluarkan uang dalam melakukan transaksi, meskipun terdapat beberapa pelanggan yang tersebar pada nilai *monetary* yang tinggi.



Gambar 6.8 Grafik Scatter Plot 3D Segmen 2

Gambar 6.9 menjelaskan lebih detail mengenai persebaran pelanggan pada segmen 3. Dari gambar tersebut dapat dilihat bahwa pelanggan menyebar pada nilai *frequency* tinggi, nilai *monetary* yang rendah, dan nilai *recency* yang relatif rendah. Hal ini menunjukkan bahwa segmen 3 memiliki karakteristik *frequency* yang tinggi atau pelanggan sering melakukan transaksi.



Gambar 6.9 Grafik Scatter Plot 3D Segmen 3

6.6.3 Grafik Box Plot

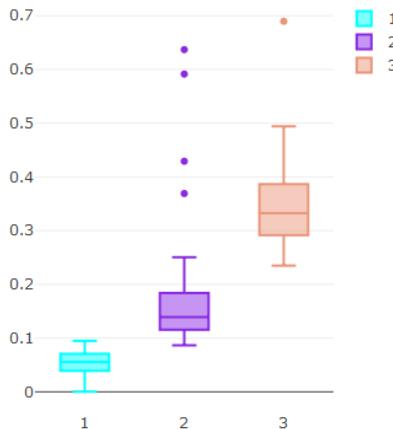
Visualisasi *Box Plot* mendeskripsikan bentuk distribusi data, ukuran tendensi sentral, dan ukuran penyebaran data. Pada grafik *box plot* disajikan data berupa nilai minimum, kuartil 1, median, kuartil 3 dan nilai maksimum. Selain itu, grafik *box plot* juga mampu mendeteksi ada tidaknya data yang *outlier*. Grafik *box plot* berguna untuk membandingkan persebaran pelanggan antar segmen.

6.6.3.1 Grafik box plot CLV

Gambar 6.10 adalah grafik *box plot* untuk nilai CLV. Segmen 1 memiliki ukuran kotak yang paling pendek dibanding dengan dua segmen lain. Hal ini menunjukkan bahwa sebaran data pada segmen 1 juga kecil. Nilai median segmen 1 terletak di tengah kotak dan garis *whisker* sama panjang, yang berarti segmen 1 bersifat simetris. Garis *whisker* bagian atas sedikit beririsan dengan segmen 2, sehingga menyebabkan sedikit *overlap*.

Segmen 2 memiliki kotak yang lebih panjang, sehingga sebaran data pada segmen 2 lebih lebar. Nilai median segmen 2 tidak terletak di tengah kotak dan panjang garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan bahwa segmen 2 tidak bersifat simetris dan cenderung ke *positive skewness*. Selain itu terdapat beberapa data yang *outlier*, sehingga *overlap* terhadap segmen 3.

Segmen 3 memiliki kotak yang paling panjang, sehingga persebaran data menjadi lebar. Nilai median segmen 3 tidak terletak di tengah kotak dan panjang garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan bahwa segmen 3 tidak simetris dan cenderung ke *positive skewness*. Selain itu terdapat data *outlier* di bagian atas kotak.



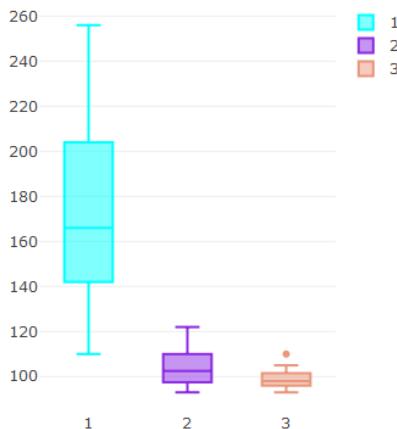
Gambar 6.10 Grafik Box Plot CLV Semua Segmen

6.6.3.2 Grafik box plot recency

Gambar 6.11 adalah grafik *box plot* untuk nilai *recency*. Segmen 1 memiliki ukuran kotak yang paling panjang

dibanding dengan dua segmen lain, sehingga persebaran data menjadi lebar. Nilai median segmen 1 tidak berada di tengah kotak dan garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan menunjukkan bahwa segmen 2 tidak bersifat simetris dan cenderung ke *positive skewness*. Garis *whisker* bagian bawah sedikit beririsan dengan segmen 2 sehingga menyebabkan sedikit *overlap*.

Segmen 2 memiliki kotak lebih pendek, sehingga persebaran data lebih kecil dibanding segmen 1. Nilai median segmen 2 tidak terletak di tengah kotak dan panjang garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan bahwa segmen 2 tidak bersifat simetris dan cenderung ke *positive skewness*. Sebagian kotak segmen 2 beririsan dengan segmen 3, sehingga menyebabkan banyak overlap.



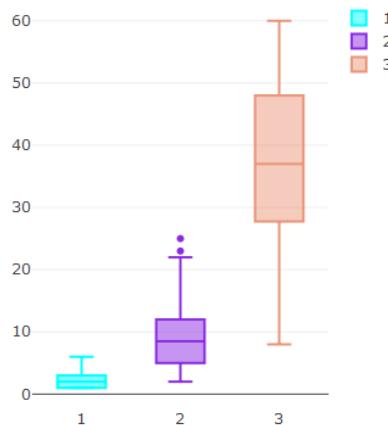
Gambar 6.11 Grafik Box Plot Recency Semua Segmen

Segmen 3 memiliki kotak yang paling pendek, sehingga persebaran data menjadi paling kecil. Nilai median segmen 3 tidak terletak di tengah kotak dan panjang garis *whisker* bagian

atas lebih panjang. Hal ini menunjukkan bahwa segmen 3 tidak simetris dan cenderung ke *positive skewness*. Selain itu terdapat data *outlier* di bagian atas kotak.

6.6.3.3 Grafik box plot frequency

Gambar 6.12 adalah grafik *box plot* untuk nilai *frequency*. Segmen 1 memiliki ukuran kotak yang paling pendek dibanding dengan dua segmen lain. Hal ini menunjukkan bahwa sebaran data pada segmen 1 juga kecil. Nilai median segmen 1 terletak di tengah kotak, tetapi garis *whisker* bagian atas lebih panjang. Sebagian kotak beriris dengan garis *whisker* segmen 2, sehingga terjadi sedikit *overlap*.



Gambar 6.12 Grafik Box Plot Frequency Semua Segmen

Segmen 2 memiliki kotak yang lebih panjang, sehingga sebaran data pada segmen 2 lebih lebar. Nilai median segmen 2 terletak di tengah kotak, tetapi garis *whisker* bagian atas lebih panjang. Selain itu terdapat beberapa data yang *outlier* di bagian atas

kotak. Sebagian kotak beririsan dengan garis *whisker* segmen 3, sehingga terjadi sedikit *overlap*.

Segmen 3 memiliki kotak yang paling panjang, sehingga persebaran data menjadi paling lebar. Nilai median segmen 3 tidak terletak di tengah kotak dan panjang garis *whisker* bagian bawah lebih panjang. Hal ini menunjukkan bahwa segmen 3 tidak simetris dan cenderung ke *negative skewness*.

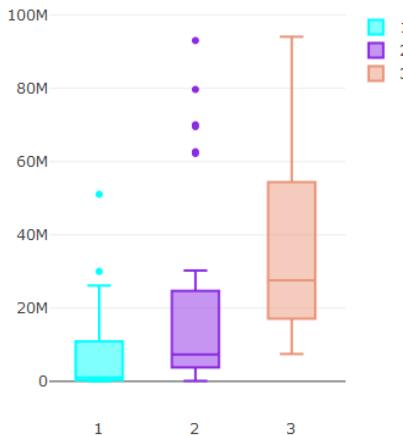
6.6.3.4 Grafik box plot monetary

Gambar 6.12 adalah grafik *box plot* untuk nilai *monetary*. Segmen 1 memiliki ukuran kotak yang paling pendek dibanding dengan dua segmen lain. Hal ini menunjukkan bahwa sebaran data pada segmen 1 juga paling kecil. Nilai median segmen 1 tidak terletak di tengah kotak dan garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan menunjukkan bahwa segmen 1 tidak bersifat simetris dan cenderung ke *positive skewness*. Selain itu, terdapat beberapa data yang *outlier* di bagian atas kotak. Sebagian kotak beririsan dengan kotak segmen 2 dan segmen 3, sehingga terjadi banyak *overlap*.

Segmen 2 memiliki kotak yang lebih panjang, sehingga sebaran data pada segmen 2 lebih lebar. Nilai median segmen 2 tidak terletak di tengah kotak dan garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan menunjukkan bahwa segmen 2 tidak bersifat simetris dan cenderung ke *positive skewness*. Selain itu terdapat beberapa data yang *outlier* di bagian atas kotak. Sebagian kotak beririsan dengan kotak segmen 1 dan 3, sehingga terjadi banyak *overlap*.

Segmen 3 memiliki kotak yang paling panjang, sehingga persebaran data menjadi paling lebar. Nilai median segmen 3 tidak terletak di tengah kotak dan panjang garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan bahwa segmen 3 tidak

simetris dan cenderung ke *positive skewness*. Sebagian kotak beririsan dengan kotak segmen 2 sehingga terjadi banyak *overlap*.



Gambar 6.13 Grafik Box Plot Monetary Semua Segmen

6.7 Kesimpulan Analisis

Dari proses *clustering* dan analisis visualisasi yang telah dilakukan, diperoleh rangkuman sebagai berikut :

1. Hasil AHP menunjukkan bahwa variabel *monetary* memiliki bobot tertinggi yaitu sebesar 0.59. Hal ini berarti bahwa variabel *monetary* menjadi pertimbangan paling penting dalam melakukan *clustering*. Sedangkan variabel *recency* dan *monetary* masing-masing memiliki bobot 0.31 dan 0.10.
2. Berdasarkan metode Elbow, nilai K yang paling baik untuk digunakan adalah 3, sehingga jumlah *cluster* yang dihasilkan pada proses *clustering* berjumlah 3 *cluster*.

3. Dari 123 pelanggan yang di-*cluster*, 42 pelanggan tergabung pada *cluster* 1, 64 pelanggan tergabung pada *cluster* 2, dan 17 pelanggan tergabung pada *cluster* 3.
4. *Cluster* 1 mempunyai rataan nilai CLV paling rendah. Hal ini didukung dengan nilai *recency* yang sangat tinggi yang mengindikasikan bahwa pelanggan sudah lama tidak melakukan transaksi. Sehingga dapat disimpulkan bahwa *cluster* 1 adalah *cluster* terburuk. Dengan karakteristik tersebut, perusahaan dapat menerapkan strategi promosi yang bertujuan untuk menarik kembali pelanggan, misalkan *advertisement campaign* via *email*.
5. *Cluster* 2 memiliki rataan CLV sedang, begitu juga dengan nilai ketiga variabel. Sehingga dapat disimpulkan bahwa *cluster* 2 adalah *cluster* urutan menengah. Dengan karakteristik tersebut, perusahaan dapat menerapkan strategi promosi yang bertujuan untuk meningkatkan frekuensi pembelian dan jumlah pembelian, misalkan memberikan penawaran promosi *up-selling* atau *cross-selling*.
6. *Cluster* 3 memiliki rataan CLV tertinggi. Hal ini didukung dengan nilai *frequency* dan *monetary* yang tinggi yang mengindikasikan pelanggan sering melakukan transaksi dan mengeluarkan banyak uang. Sehingga dapat disimpulkan bahwa *cluster* 3 adalah *cluster* terbaik. Dengan karakteristik tersebut, perusahaan dapat menerapkan strategi promosi yang bertujuan untuk menjaga mempertahankan loyalitas pelanggan, misalkan memberikan diskon terhadap produk.

Halaman ini sengaja dikosongkan

BAB VII

KESIMPULAN DAN SARAN

Bab ini memaparkan kesimpulan yang diperoleh dari rangkaian pengerjaan tugas akhir yang telah dilakukan dan menyertakan saran untuk pengembangan tugas akhir dengan topik serupa.

7.1 Kesimpulan

Dari hasil penelitian ini, dapat diperoleh beberapa kesimpulan antara lain :

1. Penggalian data dengan *clustering* dapat membantu pihak perusahaan dalam mengetahui segmentasi pelanggan yang dimiliki. *Clustering* mampu mengelompokkan pelanggan-pelanggan dalam segmen-semen tertentu berdasarkan perilaku bisnis yang sama. Metode *clustering Fuzzy C-Means* yang diimplementasikan menghasilkan 3 segmen pelanggan dengan karakteristik berbeda. Karakteristik yang menjadi pertimbangan adalah rentang waktu terakhir transaksi, jumlah transaksi, dan total pemasukan. Dengan pembagian segmen-semen ini, perusahaan dapat menerapkan strategi pemasaran yang tepat untuk setiap segmen dengan karakteristik tertentu.
2. Hasil *clustering* yang divisualisasikan berbasis web dapat mempermudah pengguna dalam membaca grafik yang ditampilkan. Pengguna dapat dengan mudah melihat persebaran anggota segmen dalam grafik *scatter plot* 3D. Pengguna juga dapat melihat rentang nilai tiap segmen pada grafik *box-plot*. Pada visualisasi ini ditampilkan juga ranking pelanggan tertinggi dan terendah pada tiap segmen berdasarkan nilai *customer lifetime value*.

7.2 Saran

Selama pengerjaan penelitian ini, tentunya terdapat kekurangan yang dapat diperbaiki dan dikembangkan pada penelitian selanjutnya. Saran yang dapat dipertimbangkan dari penulis antara lain :

1. Penelitian ini dapat dikembangkan dengan memperhatikan atribut lain selain atribut RFM untuk memberikan variasi hasil *clustering*, misalkan dengan menggunakan lokasi pelanggan.
2. Rentang data transaksi sebagai masukan lebih baik diperbanyak lebih dari 6 bulan transaksi agar hasil *clustering* lebih akurat, misalkan selama 1 tahun.
3. Penelitian ini dapat dilanjutkan dengan menggunakan metode *clustering* yang lain selain *Fuzzy C-Means* atau dikombinasikan dengan metode lain, misalkan ANFIS.
4. Penentuan nilai K dapat dikembangkan dengan metode lain, misalkan SOM, agar memberikan nilai K yang lebih optimal.

DAFTAR PUSTAKA

- [1] A. Pingit, "Inilah Industri Yang Rawan Kena Dampak MEA", Tempo.co, 8 Januari 2016. [Online]. Available : <https://www.tempo.co/read/news/2016/01/08/090734434/inilah-industri-yang-rawan-kena-dampak-mea>. [Diakses 30 Agustus 2017].
- [2] F. Yulyanna, "Ekonomi Tumbuh, Industri Tambang Malah Melorot", CNN Indonesia, 5 Mei 2017. [Online]. Available : <https://www.cnindonesia.com/ekonomi/20170505144748-85-212583/ekonomi-tumbuh-industri-tambang-malah-melorot>. [Diakses 30 Agustus 2017].
- [3] Petrochemical Trading, "Monthly Report Petrochemical Trading PT. XYZ Bulan Januari s/d Juni 2017", 2017.
- [4] Chan. C. C. H., "Online Auction Customer Segmentation Using A Neural Network Model", *International Journal of Applied Science and Engineering*, vol. 3 Issue-II, pp. 101- 109, 2005.
- [5] Y.-S. Chen, C.-H. Cheng, . C.-J. Lai, C.-Y. Hsu dan . H.-J. Syu, "Identifying patients in target customer segments using a two-stage clustering-classification approach: A hospital-based assessment," *Computers in Biology and Medicine*, p. 213–221, 2012.
- [6] A.M. Hughes, Strategic Database Marketing, Chicago, Probus Publishing Company, 2012.
- [7] Y.-S. C. Ching-Hsue Cheng, "Classifying the segmentation of customer value via RFM model and RS theory," *Expert system with application*, vol. 36, pp. 4176-4184, 2009.
- [8] S.-Y. L. H.-H. W. Jo-Ting Wei, "A review of the application of RFM model," *African journal of business management*, vol. 4, no. 19, 2010.

- [9] J. K. Yogendra dan B. K. Santosh, "Min Max Normalization Based Data Perturbation Method for Privacy Protection," *International Journal of Computer & communication Technology*, vol. 2 Issue-VIII, 2011.
- [10] K. Z. ,. S. A. ,. S. Mahboubeh Khajvand, "Estimating Customer Lifetime Value Based on RFM Analysis of Customer Purchase Behavior : Case Study," *Procedia Computer Science* , vol. 3, pp. 57-63, 2011.
- [11] R. Ramakrishnan, "Customer Lifetime Value", National Seminar on- "Changing Scenario of Consumerism", Department of Commerce, Bharathidasan University, 7-8 January 2006.
- [12] K. Sri. Purnomo, Aplikasi Logika Fuzzy untuk Pendukung Keputusan. Yogyakarta, Graha Ilmu, 2004.
- [13] B. W. Taylor, Introduction to Management Science, United States of America: Pearson Education, Inc, 2013.
- [14] S. M. Rezaeinia, A. Keramati dan A. Albadvi , "An Integrated AHP-RFM Method to Banking Customer Segmentation," *Electronic Customer Relationship Management*, vol. 6, pp. 153-168, 2012.
- [15] Anonim, "What is clustering in data mining," [Online], Available: <http://bigdata-madesimple.com/what-is-clustering-in-data-mining/>. [Diakses 30 Agustus 2017].
- [16] C. Fraley dan A. E. Fraley , "Technical Report No. 329," dalam How Many Clusters? Which Clustering Method?Answers, Department of Statisticstics University of Washington, 1998.
- [17] P. N. Tan, M. Steinbach dan V. Kumar, "Lecture Note for Chapter I : Introduction to Data Mining", Pearson, 2004.

- [18] Hoffer, Jeffrey A., Mary B. Prescott dan Fred R. McFadden. "Modern Database Management 8th Edition", 2007.
- [19] M. J. A. Berry dan G. S. Linoff, "Mastering Data Mining The Art and Science of Customer Relationship Management", Canada, John Wiley & Sons, Inc, 2000.
- [20] Ngai EWT, "Customer relationship management research (1992-2002)," *Marketing Intelligence & Planning*, Vol. 23, No. 6, pp. 582-605, 2005
- [21] G. Paul, "CRM at the Speed of Light, Fourth Edition", 2010.
- [22] P. Helen, C. Martin, dan C. Moira, Relationship Marketing – Strategy and Implementation, Burlington, Butterworth Heinemann, 2004.
- [23] Abbott J, "Data data everywhere – and not a byte of use?," *Qualitative Market Research: An International Journal*, Vol. 4, No. 3, pp.182-192, 2007.
- [24] M. Viljoen. et all, "The Use Of Technology In Customer Relationship Management (CRM)," *Acta Commercii*, 2005.
- [25] M. P. William, et all., "Marketing Principles", 2nd Edition, South Melbourne, Vic Cengage Learning, 2015.
- [26] W. Hugh, D. Elizabeth dan M. Malcolm, "Factors for Success in Customer Relationship Management (CRM) Systems," *Journal of Marketing Management*, Vol. 18, pp. 193-219, 2002.
- [27] M. J. A. Berry dan G. S. Linoff, "Mastering Data Mining The Art and Science of Customer Relationship Management", Canada, John Wiley & Sons, Inc, 2000.

- [28] S. Babak dan K. Amir, "Customer Lifetime Value (CLV) Measurement Based on RFM Model". *Iranian Accounting & Auditing Review*, Vol. 14 No. 47, pp 7- 20, 2007.
- [29] T. Luis, B. Chitta, dan K. Seungchan, "Fuzzy C-Means Clustering With Prior Biological Knowledge," *Journal of Biomedical Informatics*, Vol. 42, pp. 74-81, 2009.
- [30] J. Han, M. Kamber dan J. Pei, Data Mining Concepts and Techniques Third Edition, San Fransisco: Morgan Kauffman, 2006.
- [31] P. N. Tan, M. Steinbach dan V. Kumar, Introduction to Data Mining, Boston: Addison-Wesley Longman Publishing Co, 2005.
- [32] T. S. Madulatha, "An Overview On Clustering Methods," *IOSR Journal of Engineering*, Vol. 2 No. 4, pp. 719-725, 2012.
- [33] B. Purnima dan K. Arvind, "EBK-Means : A Clustering Techniques Based On Elbow Method and K-Means in WSN," *International Journal of Computer Application*, Vol. 9 No. 105, pp. 17-24. 2014.
- [34] G. Sonakshi dan K. Parminder, "Medical Image Segmentation Using Fuzzzy C-Means for MRI Images," *IOSR Journal of Electronics and Communication Engineering – National Conference on Advances in Engineering, Technology & Management (AETM'15)*, Vol. 1, pp. 35-40, 2015
- [35] J. Slawomir, "Removing Inconsistency in Pairwise Comparison Matrix in the AHP," *Central European Journal of Social Sciences and Humanities*, Vol 11, pp. 63-76, 2016.

- [36] B, Alireza dan S. David, “Data segmentation and Model Selection for Computer Vision : a statistical approach,” New York, Springer Science and Business Media, 2012.

BIODATA PENULIS



Penulis bernama lengkap Denny Bintang Saputra, lahir di Lamongan, 14 Oktober 1996. Penulis merupakan anak kelima dari enam bersaudara. Penulis mengenyam pendidikan formal Sekolah Dasar di SD Negeri Babat 7 dan melanjutkan ke jenjang Sekolah Menengah Pertama di SMP Negeri 1 Babat. Lulus tahun 2011, penulis melanjutkan di SMA Negeri

Model Terpadu Bojonegoro. Tiga tahun berselang, penulis mendaftar kuliah di Departemen Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi, Institut Teknologi Sepuluh Nopember Surabaya melalui jalur SBMPTN dengan beasiswa Bidikmisi. Selama menjadi mahasiswa, penulis aktif berorganisasi dan mengikuti kepanitiaan di tingkat departemen, fakultas, maupun institut. Penulis terdaftar sebagai staf ahli *Competition Test Maker, Information System Expo 2016*; staf ahli *Research and Technology Department* BEM FTIf 2016/2017; dan Ketua Departemen PSDM BIMITS 2017. Dalam penggeraan Tugas Akhir, penulis memilih laboratorium Rekayasa Data dan Intelelegensi Bisnis, dengan bimbingan Bapak Edwin Riksakomara, S.Kom., M.T., dengan topik mengenai penggalian data. Penulis dapat dihubungi melalui email dennybintangsaputra@gmail.com untuk keperluan penelitian. Semoga penulisan Tugas Akhir ini dapat memberikan kontribusi positif bagi semua pihak terkait.

LAMPIRAN A

KUESIONER SURVEY PEMERINGKATAN KEPENTINGAN KRITERIA PERILAKU PELANGGAN

Kuesioner ini bertujuan untuk mendapatkan masukan pendapat dari Narasumber terkait dengan pemeringkatan kriteria yang akan digunakan dalam pembobotan variabel segmentasi pelanggan. Kami menjamin kerahasiaan data/informasi yang diberikan serta tidak akan membawa dampak apapun bagi responden..

IDENTITAS RESPONDEN

Nama : _____

Jabatan : _____

PETUNJUK PENGISIAN KUESIONER

1. Berikan tanda centang (✓) pada kriteria yang menurut persepsi bapak/ibu lebih penting dibanding dengan kriteria lainnya, dengan cara membandingkan kriteria pada sebelah kiri dengan kriteria pada sebelah kanan. Kriteria yang dimaksud adalah perilaku pelanggan yang melakukan transaksi. Skala numerik akan menunjukkan suatu perbandingan dari tingkat kepentingan dua kriteria dengan penjelasan setiap skalanya yaitu:

Skala Nilai	Tingkat Preferensi	Penjelasan
1	Sama Pentingnya	Kedua kriteria memiliki pengaruh yang sama
3	Sedikit lebih penting	Penilaian dan pengalaman sedikit memihak pada salah satu kriteria tertentu dibanding kriteria pasangannya
5	Lebih penting	Penilaian dan pengalaman memihak pada salah satu kriteria tertentu disbanding kriteria pasangannya

A-2

7	Jelas lebih penting	Salah satu kriteria lebih diprioritaskan dan relatif lebih penting dibandingkan kriteria pasangannya.
9	Mutlak sangat penting	Salah satu kriteria sangat jelas lebih penting dibandingkan kriteria pasangannya.
2,4,6,8		Diberikan bila ada keraguan penilaian diantara dua tingkat kepentingan yang berdekatan.

2. Jika kriteria pada sebelah kiri lebih penting dibandingkan dengan kriteria sebelah kanan, maka pilih skala yang lebih dekat pada kriteria sebelah kiri, begitu pula sebaliknya.
3. Kriteria pelanggan yang dibandingkan yaitu :
 - a. *Recency* : jarak dari waktu transaksi terakhir kali dilakukan dengan waktu saat ini.
 - b. *Frequency* : total jumlah transaksi yang dilakukan selama periode tertentu
 - c. *Monetary* : total nilai produk dalam bentuk uang yang diberikan pelanggan pada perusahaan

CONTOH PENGISIAN KUESIONER

Kriteria	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Kriteria
Recency												✓						Frequency

Artinya :

Pada pernyataan tersebut, tanda centang (✓) di kolom jawaban sebelah kiri menunjukan bahwa: *recency* 1/5 kali lebih penting dibandingkan dengan *frequency* dalam hal peringkatan kriteria untuk retailer. Dengan kata lain *frequency* 5 kali lebih penting dibandingkan dengan *recency* dalam hal pemeringkatan kriteria untuk pelanggan.

DAFTAR PERTANYAAN

Berikan tanda centang (✓) untuk menunjukkan persepsi atau penilaian bapak/ibu terhadap perbandingan kriteria pelanggan yang dipilih lebih penting dengan melihat perilaku pelanggan dalam melakukan transaksi dengan perusahaan. “Manakah kriteria yang lebih penting antara keriteria pada kolom sebelah kiri dengan kriteria pada kolom sebelah kanan?”

Kriter ia																		Kriter ia
Recen cy	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Frequ ency
Frequ ency	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Mone tary
Recen cy	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Mone tary

Terima kasih atas kesediaan bapak/ibu untuk mengisi kuesioner ini

Halaman ini sengaja dikosongkan

LAMPIRAN B

Tabel B.1 Use Case Melihat Grafik Scatter Plot 3D

UC-2 Melihat Grafik Scatter Plot 3D		
Purpose	Mengetahui persebaran pelanggan dalam grafik <i>scatter plot</i> 3 dimensi, berdasarkan <i>recency</i> , <i>frequency</i> , dan <i>monetary</i> .	
Overview	Dimulai ketika pengguna mengakses aplikasi	
Actors	Pengguna aplikasi	
Pre Condition	Pengguna telah berada pada halaman utama	
Post Condition	Grafik <i>scatter plot</i> 3D ditampilkan	
Typical Course Event	Actor  1. Mengklik ikon <i>slide bar</i>  3. Mengeklik menu “Visualisasi Segmen” 5. Mengklik salah satu pilihan segmen pada menu <i>dropdown</i>  7. Mengklik ikon pada kotak dengan judul “Grafik 3 Dimensi” 9. Mengklik dan memutar kubus grafik	System 2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman “Visualisasi Segmen” 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan grafik <i>scatter plot</i> 10. Menampilkan grafik dari sisi yang berbeda
Alternate Flow Event	Langkah 5 : jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan grafik semua segmen secara default	
Exceptional Flow Event	Aplikasi tertutup karena <i>error</i>	

Tabel B.2 Use Case Melihat Grafik Box-Plot

UC-3 Melihat Persebaran Pada Grafik Box-Plot						
Purpose	Mengetahui persebaran nilai variabel <i>recency</i> , <i>frequency</i> , dan <i>monetary</i> dalam grafik <i>box-plot</i>					
Overview	Dimulai ketika pengguna mengakses aplikasi					
Actors	Pengguna aplikasi					
Pre Condition	Pengguna telah berada pada halaman utama					
Post Condition	Grafik <i>box-plot</i> ditampilkan					
Typical Course Event	<table> <thead> <tr> <th><i>Actor</i></th><th><i>System</i></th></tr> </thead> <tbody> <tr> <td> 1. Mengklik ikon  3. Mengklik menu “Visualisasi Segmen” 5. Mengklik salah satu pilihan segmen pada menu <i>dropdown</i> 7. Mengklik ikon  </td><td> 2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman “Visualisasi Segmen” 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan grafik <i>scatter plot</i> <i>recency</i>, <i>frequency</i>, atau <i>monetary</i> </td></tr> </tbody> </table>	<i>Actor</i>	<i>System</i>	1. Mengklik ikon  3. Mengklik menu “Visualisasi Segmen” 5. Mengklik salah satu pilihan segmen pada menu <i>dropdown</i> 7. Mengklik ikon 	2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman “Visualisasi Segmen” 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan grafik <i>scatter plot</i> <i>recency</i> , <i>frequency</i> , atau <i>monetary</i>	
<i>Actor</i>	<i>System</i>					
1. Mengklik ikon  3. Mengklik menu “Visualisasi Segmen” 5. Mengklik salah satu pilihan segmen pada menu <i>dropdown</i> 7. Mengklik ikon 	2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman “Visualisasi Segmen” 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan grafik <i>scatter plot</i> <i>recency</i> , <i>frequency</i> , atau <i>monetary</i>					
Alternate Flow Event	Langkah 5 : jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan grafik semua segmen secara default					
Exceptional Flow Event	Aplikasi tertutup karena <i>error</i>					

Tabel B.3 Use Case Melihat Rentang Nilai Segmen

UC-4 Melihat rentang nilai segmen						
Purpose	Mengetahui rentang nilai variabel RFM pada setiap segmen					
Overview	Dimulai ketika pengguna mengakses aplikasi					
Actors	Pengguna aplikasi					
Pre Condition	Pengguna telah berada pada halaman utama					
Post Condition	Rentang nilai ditampilkan					
Typical Course Event	<table> <thead> <tr> <th><i>Actor</i></th><th><i>System</i></th></tr> </thead> <tbody> <tr> <td> 1. Mengklik ikon  3. Mengklik menu “Visualisasi Segmen” 5. Mengklik salah satu pilihan segmen pada menu <i>dropdown</i> 7. Mengklik ikon  </td><td> 2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman “Visualisasi Segmen” 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan rentang nilai ketiga variabel pada segmen </td></tr> </tbody> </table>	<i>Actor</i>	<i>System</i>	1. Mengklik ikon  3. Mengklik menu “Visualisasi Segmen” 5. Mengklik salah satu pilihan segmen pada menu <i>dropdown</i> 7. Mengklik ikon 	2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman “Visualisasi Segmen” 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan rentang nilai ketiga variabel pada segmen	
<i>Actor</i>	<i>System</i>					
1. Mengklik ikon  3. Mengklik menu “Visualisasi Segmen” 5. Mengklik salah satu pilihan segmen pada menu <i>dropdown</i> 7. Mengklik ikon 	2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman “Visualisasi Segmen” 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan rentang nilai ketiga variabel pada segmen					
Alternate Flow Event	Langkah 5 : jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan pilihan semua segmen secara default					
Exceptional Flow Event	Aplikasi tertutup karena <i>error</i>					

Tabel B.4 Use Case Melihat Pelanggan Teratas dan Terbawah

UC-5 Melihat pelanggan teratas dan terbawah		
Purpose	Mengetahui pelanggan dengan nilai CLV teratas dan terbawah setiap segmen	
Overview	Dimulai ketika pengguna mengakses aplikasi	
Actors	Pengguna aplikasi	
Pre Condition	Pengguna telah berada pada halaman utama	
Post Condition	Pelanggan teratas dan terbawah ditampilkan	
Typical Course Event	<p>Actor</p> <p>1. Mengklik ikon  pada slide bar 2. Menampilkan pilihan menu pada slide bar</p> <p>3. Mengklik menu “Visualisasi Segmen” 4. Menampilkan halaman “Visualisasi Segmen”</p> <p>5. Mengklik salah satu pilihan segmen pada menu dropdown 6. Menampilkan visualisasi sesuai pilihan</p> <p>7. Mengklik ikon  pada kotak dengan judul “Pelanggan Teratas dan Terbawah” 8. Menampilkan pelanggan teratas dan terbawah pada segmen</p>	<p>System</p> <p>Menampilkan pilihan menu pada slide bar</p> <p>Menampilkan halaman “Visualisasi Segmen”</p> <p>Menampilkan visualisasi sesuai pilihan</p> <p>Menampilkan pelanggan teratas dan terbawah pada segmen</p>
Alternate Flow Event	Langkah 5 : jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan pilihan semua segmen secara default	
Exceptional Flow Event	Aplikasi tertutup karena <i>error</i>	

Tabel B.5 Use Case Melihat Rataan Variabel RFM

UC-6 Melihat rataan variabel RFM setiap cluster		
Purpose	Mengetahui nilai rata-rata variabel <i>recency</i> , <i>frequency</i> , dan <i>monetary</i> setiap segmen	
Overview	Dimulai ketika pengguna mengakses aplikasi	
Actors	Pengguna aplikasi	
Pre Condition	Pengguna telah berada pada halaman utama	
Post Condition	Nilai rata-rata ketiga variabel ditampilkan	
Typical Course Event	<p>Actor</p> <ol style="list-style-type: none"> Mengklik ikon  Mengeklik menu “Visualisasi Segmen” Mengklik salah satu pilihan segmen pada menu <i>dropdown</i> 	<p>System</p> <ol style="list-style-type: none"> Menampilkan pilihan menu pada <i>slide bar</i> Menampilkan halaman “Visualisasi Segmen” Menampilkan rataan nilai ketiga variabel sesuai pilihan segmen
Alternate Flow Event	Langkah 5 : jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan pilihan semua segmen secara default	
Exceptional Flow Event	Aplikasi tertutup karena <i>error</i>	

Tabel B.6 Use Case Melihat Peringkat Pelanggan Setiap Cluster

UC-7 Melihat peringkat pelanggan setiap cluster						
Purpose	Mengetahui tabel peringkat pelanggan pada setiap cluster segmen berdasarkan nilai CLV pelanggan					
Overview	Dimulai ketika pengguna mengakses aplikasi					
Actors	Pengguna aplikasi					
Pre Condition	Pengguna telah berada pada halaman utama					
Post Condition	Tabel peringkat pelanggan ditampilkan					
Typical Course Event	<table> <thead> <tr> <th><i>Actor</i></th><th><i>System</i></th></tr> </thead> <tbody> <tr> <td> 1. Mengklik ikon  3. Mengklik menu “Visualisasi Segmen” 5. Mengklik salah satu pilihan segmen pada menu dropdown 7. Mengklik ikon  pada kotak dengan judul “Peringkat Pelanggan” </td><td> 2. Menampilkan pilihan menu pada slide bar 4. Menampilkan halaman “Visualisasi Segmen” 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan tabel peringkat pelanggan </td></tr> </tbody> </table>	<i>Actor</i>	<i>System</i>	1. Mengklik ikon  3. Mengklik menu “Visualisasi Segmen” 5. Mengklik salah satu pilihan segmen pada menu dropdown 7. Mengklik ikon  pada kotak dengan judul “Peringkat Pelanggan”	2. Menampilkan pilihan menu pada slide bar 4. Menampilkan halaman “Visualisasi Segmen” 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan tabel peringkat pelanggan	
<i>Actor</i>	<i>System</i>					
1. Mengklik ikon  3. Mengklik menu “Visualisasi Segmen” 5. Mengklik salah satu pilihan segmen pada menu dropdown 7. Mengklik ikon  pada kotak dengan judul “Peringkat Pelanggan”	2. Menampilkan pilihan menu pada slide bar 4. Menampilkan halaman “Visualisasi Segmen” 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan tabel peringkat pelanggan					
Alternate Flow Event	Langkah 5 : jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan pilihan semua segmen secara default					
Exceptional Flow Event	Aplikasi tertutup karena <i>error</i>					

Tabel B.7 Use Case Melihat Tabel Data

UC-8 Melihat tabel data <i>cluster</i> pelanggan		
Purpose	Mengetahui secara detail tabel data <i>cluster</i> pelanggan secara keseluruhan	
Overview	Dimulai ketika pengguna mengakses aplikasi	
Actors	Pengguna aplikasi	
Pre Condition	Pengguna telah berada pada halaman utama	
Post Condition	Tabel data <i>cluster</i> pelanggan ditampilkan	
Typical Course Event	<i>Actor</i>	<i>System</i>
	1. Mengklik ikon  2. Menampilkan pilihan menu pada <i>slide bar</i> 3. Mengklik menu “Tabel”	2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan tabel data <i>cluster</i> pelanggan
Alternate Flow Event	-	
Exceptional Flow Event	Aplikasi tertutup karena <i>error</i>	

Halaman ini sengaja dikosongkan

LAMPIRAN C

Tabel C.1 Pemodelan RFM Keseluruhan Pelanggan

<i>Customer ID</i>	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
2053	97	60	94.075.983
2138	93	11	4.731.484
2166	98	10	356.683.172
700032	98	27	10.145.412
700189	204	1	200.000
700286	100	7	10.517.522
700385	160	4	2.750.130
700493	250	1	151.085
700503	121	11	12.610.553
700529	98	16	3.021.795
700793	97	11	27.860.136
701121	103	34	65.581.103
701976	99	48	58.437.193
702223	105	37	23.094.523
702643	110	4	62.681.101
702761	122	4	12.318.043
702920	216	2	3.979.716
704054	107	9	4.244.696
704402	104	22	5.645.602
704622	95	9	14.460.571
704818	115	14	62.221.259
704838	105	5	79.689.999
704904	96	44	14.786.319
705329	119	6	816.972.362
705416	94	11	2.395.857
705759	112	5	751.214.097
705762	96	8	861.123.629
706383	121	5	199.309.712
706737	108	4	510.237.144
706806	100	5	136.046.782
710854	101	5	13.723.096
710856	106	4	1.495.264
710858	97	6	6.572.262
710868	95	23	3.821.480

<i>Customer ID</i>	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
710869	101	42	18.955.704
711781	97	17	5.802.074
711810	142	3	327.625
711831	113	6	26.247.416
712559	98	6	69.550.744
712561	97	28	7.481.952
712562	122	7	2.746.530
712563	100	6	9.855.156
712564	120	5	9.511.411
712565	180	1	103.160
712566	110	7	4.471.623
712717	100	48	36.589.453
712746	99	50	17.893.680
712748	93	39	35.005.558
712785	100	13	7.052.128
712843	94	12	19.708.522
712872	105	9	3.897.821
713115	112	5	783.708
713136	95	25	3.463.566
713167	120	4	724.670
713249	110	2	242.550
713253	93	20	7.394.317
713255	101	20	7.316.712
713320	146	1	150.000
714954	112	6	1.164.559
715470	98	26	7.721.476
719930	93	6	30.270.365
719971	135	4	1.693.943
720145	96	36	27.589.588
720236	158	2	15.903.063
721918	130	2	9.042.186
721954	157	1	4.985.287
722149	105	2	15.172.797
722171	109	17	121.423.217
728260	173	2	12.606.554
728447	114	18	10.297.772
729665	104	25	53.048.622
730112	120	4	677.986
730648	256	1	392.071
730948	96	12	3.085.379

<i>Customer ID</i>	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
731081	102	4	510.977
731090	142	2	184.935
731737	193	2	51.101.240
735810	117	11	4.104.686
737744	115	1	12.426.150
738593	188	1	26.163.919
739995	119	8	5.059.845
743352	96	36	20.309.572
746205	177	2	10.920.435
751408	101	10	14.322.453
757273	100	8	7.787.607
758186	182	2	8.143.134
762727	166	2	24.107.869
766881	94	15	93.044.715
777121	120	3	324.000
777897	221	1	427.693
788346	183	3	13.503.303
793566	109	9	23.175.433
794521	210	1	30.019.821
795940	236	1	929.366
804967	105	5	3.799.334
806038	224	2	7.163.209
806791	129	5	1.130.741
807046	94	11	3.799.716
819874	98	5	616.815
820764	109	12	70.048.285
821655	95	18	5.184.628
821868	110	51	35.370.834
822231	183	6	582.274
823661	141	2	645.482
823940	109	9	4.342.613
826451	110	4	891.006
828392	166	1	86.000
832323	151	1	353.649
832593	190	4	3.303.871
848041	99	6	2.150.480
850911	112	6	19.974.781
853537	103	17	5.462.694
858218	221	1	12.479.874
869214	161	1	1.000.000

C-4

<i>Customer ID</i>	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
883122	94	11	9.089.730
887934	250	1	3.345.000
888124	100	6	10.107.613
890697	103	3	1.819.766
892055	212	1	900.000
892303	101	10	3.887.791
893393	154	2	26.800
893562	111	3	133.893
893809	156	1	55.900

LAMPIRAN D

Tabel D.1 Hasil Keseluruhan Normalisasi Variabel RFM

<i>Customer ID</i>	<i>R_Norm</i>	<i>R_Rev</i>	<i>F_Norm</i>	<i>M_Norm</i>
2053	0,02454	0,97546	1	0,10922
2138	0	1	0,16949	0,00546
2166	0,03067	0,96933	0,15254	0,41419
700032	0,03067	0,96933	0,44068	0,01175
700189	0,68098	0,31902	0	0,00020
700286	0,04294	0,95706	0,10169	0,01218
700385	0,41104	0,58896	0,05085	0,00316
700493	0,96319	0,03681	0	0,00014
700503	0,17178	0,82822	0,16949	0,01461
700529	0,03067	0,96933	0,25424	0,00348
700793	0,02454	0,97546	0,16949	0,03232
701121	0,06135	0,93865	0,55932	0,07613
701976	0,03681	0,96319	0,79661	0,06783
702223	0,07362	0,92638	0,61017	0,02679
702643	0,10429	0,89571	0,05085	0,07276
702761	0,17791	0,82209	0,05085	0,01427
702920	0,75460	0,24540	0,01695	0,00459
704054	0,08589	0,91411	0,13559	0,00490
704402	0,06748	0,93252	0,35593	0,00653
704622	0,01227	0,98773	0,13559	0,01676
704818	0,13497	0,86503	0,22034	0,07223
704838	0,07362	0,92638	0,06780	0,09251
704904	0,01840	0,98160	0,72881	0,01714
705329	0,15951	0,84049	0,08475	0,94873
705416	0,00613	0,99387	0,16949	0,00275
705759	0,11656	0,88344	0,06780	0,87236
705762	0,01840	0,98160	0,11864	1,00000
706383	0,17178	0,82822	0,06780	0,23143
706737	0,09202	0,90798	0,05085	0,59251
706806	0,04294	0,95706	0,06780	0,15796
710854	0,04908	0,95092	0,06780	0,01591
710856	0,07975	0,92025	0,05085	0,00171
710858	0,02454	0,97546	0,08475	0,00760

<i>Customer ID</i>	<i>R_Norm</i>	<i>R_Rev</i>	<i>F_Norm</i>	<i>M_Norm</i>
710868	0,01227	0,98773	0,37288	0,00441
710869	0,04908	0,95092	0,69492	0,02198
711781	0,02454	0,97546	0,27119	0,00671
711810	0,30061	0,69939	0,03390	0,00035
711831	0,12270	0,87730	0,08475	0,03045
712559	0,03067	0,96933	0,08475	0,08074
712561	0,02454	0,97546	0,45763	0,00866
712562	0,17791	0,82209	0,10169	0,00316
712563	0,04294	0,95706	0,08475	0,01141
712564	0,16564	0,83436	0,06780	0,01101
712565	0,53374	0,46626	0	0,00009
712566	0,10429	0,89571	0,10169	0,00516
712717	0,04294	0,95706	0,79661	0,04246
712746	0,03681	0,96319	0,83051	0,02075
712748	0	1	0,64407	0,04062
712785	0,04294	0,95706	0,20339	0,00816
712843	0,00613	0,99387	0,18644	0,02286
712872	0,07362	0,92638	0,13559	0,00450
713115	0,11656	0,88344	0,06780	0,00088
713136	0,01227	0,98773	0,40678	0,00399
713167	0,16564	0,83436	0,05085	0,00081
713249	0,10429	0,89571	0,01695	0,00025
713253	0	1	0,32203	0,00856
713255	0,04908	0,95092	0,32203	0,00847
713320	0,32515	0,67485	0	0,00014
714954	0,11656	0,88344	0,08475	0,00132
715470	0,03067	0,96933	0,42373	0,00894
719930	0	1	0,08475	0,03512
719971	0,25767	0,74233	0,05085	0,00194
720145	0,01840	0,98160	0,59322	0,03201
720236	0,39877	0,60123	0,01695	0,01844
721918	0,22699	0,77301	0,01695	0,01047
721954	0,39264	0,60736	0	0,00576
722149	0,07362	0,92638	0,01695	0,01759
722171	0,09816	0,90184	0,27119	0,14098
728260	0,49080	0,50920	0,01695	0,01461
728447	0,12883	0,87117	0,28814	0,01193
729665	0,06748	0,93252	0,40678	0,06157
730112	0,16564	0,83436	0,05085	0,00076

<i>Customer ID</i>	<i>R_Norm</i>	<i>R_Rev</i>	<i>F_Norm</i>	<i>M_Norm</i>
730648	1	0	0	0,00042
730948	0,01840	0,98160	0,18644	0,00355
731081	0,05521	0,94479	0,05085	0,00056
731090	0,30061	0,69939	0,01695	0,00018
731737	0,61350	0,38650	0,01695	0,05931
735810	0,14724	0,85276	0,16949	0,00474
737744	0,13497	0,86503	0	0,01440
738593	0,58282	0,41718	0	0,03035
739995	0,15951	0,84049	0,11864	0,00584
743352	0,01840	0,98160	0,59322	0,02355
746205	0,51534	0,48466	0,01695	0,01265
751408	0,04908	0,95092	0,15254	0,01660
757273	0,04294	0,95706	0,11864	0,00901
758186	0,54601	0,45399	0,01695	0,00943
762727	0,44785	0,55215	0,01695	0,02797
766881	0,00613	0,99387	0,23729	0,10802
777121	0,16564	0,83436	0,03390	0,00035
777897	0,78528	0,21472	0	0,00047
788346	0,55215	0,44785	0,03390	0,01565
793566	0,09816	0,90184	0,13559	0,02688
794521	0,71779	0,28221	0	0,03483
795940	0,87730	0,12270	0	0,00105
804967	0,07362	0,92638	0,06780	0,00438
806038	0,80368	0,19632	0,01695	0,00829
806791	0,22086	0,77914	0,06780	0,00128
807046	0,00613	0,99387	0,16949	0,00438
819874	0,03067	0,96933	0,06780	0,00069
820764	0,09816	0,90184	0,18644	0,08132
821655	0,01227	0,98773	0,28814	0,00599
821868	0,10429	0,89571	0,84746	0,04105
822231	0,55215	0,44785	0,08475	0,00065
823661	0,29448	0,70552	0,01695	0,00072
823940	0,09816	0,90184	0,13559	0,00501
826451	0,10429	0,89571	0,05085	0,00100
828392	0,44785	0,55215	0	0,00007
832323	0,35583	0,64417	0	0,00038
832593	0,59509	0,40491	0,05085	0,00381
848041	0,03681	0,96319	0,08475	0,00247
850911	0,11656	0,88344	0,08475	0,02317

D-4

<i>Customer ID</i>	<i>R_Norm</i>	<i>R_Rev</i>	<i>F_Norm</i>	<i>M_Norm</i>
853537	0,06135	0,93865	0,27119	0,00631
858218	0,78528	0,21472	0	0,01446
869214	0,41718	0,58282	0	0,00113
883122	0,00613	0,99387	0,16949	0,01052
887934	0,96319	0,03681	0	0,00385
888124	0,04294	0,95706	0,08475	0,01171
890697	0,06135	0,93865	0,03390	0,00208
892055	0,73006	0,26994	0	0,00101
892303	0,04908	0,95092	0,15254	0,00448
893393	0,37423	0,62577	0,01695	0,00000
893562	0,11043	0,88957	0,03390	0,00012
893809	0,38650	0,61350	0	0,00003

LAMPIRAN E

Tabel E.1 Hasil Keseluruhan Perhitungan CLV dan Clustering

<i>Customer ID</i>	<i>R_Bobot</i>	<i>F_Bobot</i>	<i>M_Bobot</i>	<i>CLV</i>	<i>Cluster</i>
2053	0,08182	0,35047	0,06178	0,49407	3
2138	0,08387	0,05940	0,00309	0,14637	2
2166	0,08130	0,05346	0,23429	0,36905	2
700032	0,08130	0,15444	0,00665	0,24239	3
700189	0,02676	0	0,00011	0,02687	1
700286	0,08027	0,03564	0,00689	0,12281	2
700385	0,04940	0,01782	0,00179	0,06901	1
700493	0,00309	0	0,00008	0,00317	1
700503	0,06947	0,05940	0,00827	0,13713	2
700529	0,08130	0,08910	0,00197	0,17237	2
700793	0,08182	0,05940	0,01828	0,15950	2
701121	0,07873	0,19603	0,04306	0,31782	3
701976	0,08079	0,27919	0,03837	0,39835	3
702223	0,07770	0,21385	0,01515	0,30670	3
702643	0,07513	0,01782	0,04116	0,13411	2
702761	0,06895	0,01782	0,00807	0,09485	1
702920	0,02058	0,00594	0,00260	0,02912	1
704054	0,07667	0,04752	0,00277	0,12696	2
704402	0,07821	0,12474	0,00369	0,20665	2
704622	0,08285	0,04752	0,00948	0,13985	2
704818	0,07255	0,07722	0,04086	0,19063	2
704838	0,07770	0,02376	0,05233	0,15379	2
704904	0,08233	0,25543	0,00970	0,34745	3
705329	0,07050	0,02970	0,53665	0,63685	2
705416	0,08336	0,05940	0,00156	0,14432	2
705759	0,07410	0,02376	0,49346	0,59131	2
705762	0,08233	0,04158	0,56566	0,68957	3
706383	0,06947	0,02376	0,13091	0,22414	2
706737	0,07616	0,01782	0,33516	0,42913	2
706806	0,08027	0,02376	0,08935	0,19338	2
710854	0,07976	0,02376	0,00900	0,11252	2
710856	0,07719	0,01782	0,00096	0,09597	2
710858	0,08182	0,02970	0,00430	0,11582	2

<i>Customer ID</i>	<i>R_Bobot</i>	<i>F_Bobot</i>	<i>M_Bobot</i>	<i>CLV</i>	<i>Cluster</i>
710868	0,08285	0,13068	0,00249	0,21602	2
710869	0,07976	0,24355	0,01243	0,33574	3
711781	0,08182	0,09504	0,00379	0,18065	2
711810	0,05866	0,01188	0,00020	0,07074	1
711831	0,07358	0,02970	0,01722	0,12051	2
712559	0,08130	0,02970	0,04567	0,15667	2
712561	0,08182	0,16038	0,00490	0,24710	3
712562	0,06895	0,03564	0,00179	0,10638	2
712563	0,08027	0,02970	0,00646	0,11643	2
712564	0,06998	0,02376	0,00623	0,09997	2
712565	0,03911	0	0,00005	0,03916	1
712566	0,07513	0,03564	0,00292	0,11369	2
712717	0,08027	0,27919	0,02402	0,38348	3
712746	0,08079	0,29107	0,01174	0,38359	3
712748	0,08387	0,22573	0,02298	0,33258	3
712785	0,08027	0,07128	0,00461	0,15617	2
712843	0,08336	0,06534	0,01293	0,16163	2
712872	0,07770	0,04752	0,00254	0,12776	2
713115	0,07410	0,02376	0,00050	0,09836	2
713136	0,08285	0,14256	0,00226	0,22767	2
713167	0,06998	0,01782	0,00046	0,08826	1
713249	0,07513	0,00594	0,00014	0,08121	1
713253	0,08387	0,11286	0,00484	0,20158	2
713255	0,07976	0,11286	0,00479	0,19741	2
713320	0,05660	0	0,00008	0,05668	1
714954	0,07410	0,02970	0,00075	0,10455	2
715470	0,08130	0,14850	0,00505	0,23486	3
719930	0,08387	0,02970	0,01987	0,13344	2
719971	0,06226	0,01782	0,00110	0,08118	1
720145	0,08233	0,20791	0,01811	0,30834	3
720236	0,05043	0,00594	0,01043	0,06680	1
721918	0,06484	0,00594	0,00592	0,07670	1
721954	0,05094	0	0,00326	0,05420	1
722149	0,07770	0,00594	0,00995	0,09359	2
722171	0,07564	0,09504	0,07975	0,25043	2
728260	0,04271	0,00594	0,00826	0,05691	1
728447	0,07307	0,10098	0,00675	0,18080	2
729665	0,07821	0,14256	0,03483	0,25561	3
730112	0,06998	0,01782	0,00043	0,08823	1

<i>Customer ID</i>	<i>R_Bobot</i>	<i>F_Bobot</i>	<i>M_Bobot</i>	<i>CLV</i>	<i>Cluster</i>
730648	0	0	0,00024	0,00024	1
730948	0,08233	0,06534	0,00201	0,14968	2
731081	0,07924	0,01782	0,00032	0,09738	2
731090	0,05866	0,00594	0,00010	0,06470	1
731737	0,03242	0,00594	0,03355	0,07191	1
735810	0,07152	0,05940	0,00268	0,13361	2
737744	0,07255	0	0,00815	0,08070	1
738593	0,03499	0	0,01717	0,05216	1
739995	0,07050	0,04158	0,00331	0,11538	2
743352	0,08233	0,20791	0,01332	0,30356	3
746205	0,04065	0,00594	0,00716	0,05375	1
751408	0,07976	0,05346	0,00939	0,14261	2
757273	0,08027	0,04158	0,00510	0,12695	2
758186	0,03808	0,00594	0,00533	0,04935	1
762727	0,04631	0,00594	0,01582	0,06807	1
766881	0,08336	0,08316	0,06110	0,22763	2
777121	0,06998	0,01188	0,00020	0,08206	1
777897	0,01801	0	0,00026	0,01827	1
788346	0,03756	0,01188	0,00885	0,05830	1
793566	0,07564	0,04752	0,01521	0,13837	2
794521	0,02367	0	0,01970	0,04337	1
795940	0,01029	0	0,00059	0,01088	1
804967	0,07770	0,02376	0,00248	0,10394	2
806038	0,01647	0,00594	0,00469	0,02709	1
806791	0,06535	0,02376	0,00073	0,08984	1
807046	0,08336	0,05940	0,00248	0,14524	2
819874	0,08130	0,02376	0,00039	0,10545	2
820764	0,07564	0,06534	0,04600	0,18698	2
821655	0,08285	0,10098	0,00339	0,18722	2
821868	0,07513	0,29701	0,02322	0,39535	3
822231	0,03756	0,02970	0,00036	0,06763	1
823661	0,05918	0,00594	0,00041	0,06552	1
823940	0,07564	0,04752	0,00284	0,12600	2
826451	0,07513	0,01782	0,00057	0,09352	2
828392	0,04631	0	0,00004	0,04635	1
832323	0,05403	0	0,00021	0,05424	1
832593	0,03396	0,01782	0,00215	0,05393	1
848041	0,08079	0,02970	0,00140	0,11188	2
850911	0,07410	0,02970	0,01310	0,11690	2

<i>Customer ID</i>	<i>R_Bobot</i>	<i>F_Bobot</i>	<i>M_Bobot</i>	<i>CLV</i>	<i>Cluster</i>
853537	0,07873	0,09504	0,00357	0,17734	2
858218	0,01801	0	0,00818	0,02619	1
869214	0,04888	0	0,00064	0,04952	1
883122	0,08336	0,05940	0,00595	0,14872	2
887934	0,00309	0	0,00218	0,00527	1
888124	0,08027	0,02970	0,00662	0,11660	2
890697	0,07873	0,01188	0,00118	0,09179	2
892055	0,02264	0	0,00057	0,02321	1
892303	0,07976	0,05346	0,00254	0,13576	2
893393	0,05249	0,00594	0	0,05843	1
893562	0,07461	0,01188	0,00007	0,08656	2
893809	0,05146	0	0,00002	0,05148	1

LAMPIRAN F

Tabel F.1 Keseluruhan Hasil Percobaan Perubahan Nilai Parameter

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
16	0,1	0,1	0,0041660987	0,0015782602
17	0,1	0,1	0,0025282876	0,0015782514
18	0,1	0,1	0,0018133676	0,0015782480
19	0,1	0,1	0,0017496160	0,0015782467
20	0,1	0,1	0,0017453833	0,0015782462
16	0,2	0,1	0,0046781852	0,0015955338
17	0,2	0,1	0,0046780361	0,0015880502
18	0,2	0,1	0,0046779125	0,0015832561
19	0,2	0,1	0,0046778088	0,0015805853
20	0,2	0,1	0,0046777206	0,0015792651
16	0,3	0,1	0,0046776787	0,0016851874
17	0,3	0,1	0,0046776088	0,0016851296
18	0,3	0,1	0,0046775480	0,0016850930
19	0,3	0,1	0,0046774947	0,0016850698
20	0,3	0,1	0,0046774476	0,0016850551
16	0,4	0,1	0,0046784808	0,0016906859
17	0,4	0,1	0,0046782769	0,0016890747
18	0,4	0,1	0,0046781116	0,0016878452
19	0,4	0,1	0,0046779754	0,0016869489
20	0,4	0,1	0,0046778617	0,0016863175
16	0,5	0,1	0,0046487751	0,0016850332
17	0,5	0,1	0,0046364081	0,0016850318
18	0,5	0,1	0,0046142966	0,0016850309
19	0,5	0,1	0,0045673890	0,0016850303
20	0,5	0,1	0,0044281701	0,0016850299
16	0,6	0,1	0,0017454813	0,0015782470
17	0,6	0,1	0,0017450836	0,0015782463
18	0,6	0,1	0,0017450380	0,0015782460
19	0,6	0,1	0,0017450263	0,0015782459
20	0,6	0,1	0,0017450223	0,0015782459
16	0,7	0,1	0,0017450205	0,0015782542
17	0,7	0,1	0,0017450202	0,0015782491
18	0,7	0,1	0,0017450201	0,0015782471

F-2

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
19	0,7	0,1	0,0017450201	0,0015782463
20	0,7	0,1	0,0017450200	0,0015782461
16	0,8	0,1	0,0017450599	0,0016850470
17	0,8	0,1	0,0017450346	0,0016850406
18	0,8	0,1	0,0017450253	0,0016850365
19	0,8	0,1	0,0017450220	0,0016850339
20	0,8	0,1	0,0017450207	0,0016850320
16	0,9	0,1	0,0017450324	0,0015783461
17	0,9	0,1	0,0017450246	0,0015782850
18	0,9	0,1	0,0017450217	0,0015782610
19	0,9	0,1	0,0017450206	0,0015782517
20	0,9	0,1	0,0017450203	0,0015782481
16	1	0,1	0,0046781623	0,0015782477
17	1	0,1	0,0046780173	0,0015782466
18	1	0,1	0,0046778968	0,0015782461
19	1	0,1	0,0046777955	0,0015782460
20	1	0,1	0,0046777093	0,0015782459
16	0,1	0,2	0,0017450235	0,0016850404
17	0,1	0,2	0,0017450213	0,0016850363
18	0,1	0,2	0,0017450205	0,0016850338
19	0,1	0,2	0,0017450202	0,0016850321
20	0,1	0,2	0,0017450201	0,0016850311
16	0,2	0,2	0,0017551396	0,0015783200
17	0,2	0,2	0,0017457172	0,0015782747
18	0,2	0,2	0,0017451112	0,0015782570
19	0,2	0,2	0,0017450451	0,0015782501
20	0,2	0,2	0,0017450287	0,0015782475
16	0,3	0,2	0,0017450819	0,0015782494
17	0,3	0,2	0,0017450426	0,0015782472
18	0,3	0,2	0,0017450283	0,0015782464
19	0,3	0,2	0,0017450230	0,0015782461
20	0,3	0,2	0,0017450211	0,0015782460
16	0,4	0,2	0,0046771872	0,0015784569
17	0,4	0,2	0,0046771685	0,0015783292
18	0,4	0,2	0,0046771511	0,0015782783
19	0,4	0,2	0,0046771348	0,0015782584
20	0,4	0,2	0,0046771194	0,0015782507
16	0,5	0,2	0,0017450225	0,0015791703
17	0,5	0,2	0,0017450209	0,0015786277

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
18	0,5	0,2	0,0017450204	0,0015783988
19	0,5	0,2	0,0017450202	0,0015783059
20	0,5	0,2	0,0017450201	0,0015782692
16	0,6	0,2	0,0017450282	0,0015782471
17	0,6	0,2	0,0017450230	0,0015782463
18	0,6	0,2	0,0017450211	0,0015782461
19	0,6	0,2	0,0017450204	0,0015782459
20	0,6	0,2	0,0017450202	0,0015782459
16	0,7	0,2	0,0017450252	0,0015782491
17	0,7	0,2	0,0017450219	0,0015782471
18	0,7	0,2	0,0017450207	0,0015782463
19	0,7	0,2	0,0017450203	0,0015782461
20	0,7	0,2	0,0017450201	0,0015782459
16	0,8	0,2	0,0046709451	0,0015784161
17	0,8	0,2	0,0046695362	0,0015783128
18	0,8	0,2	0,0046676629	0,0015782719
19	0,8	0,2	0,0046651118	0,0015782559
20	0,8	0,2	0,0046615430	0,0015782497
16	0,9	0,2	0,0017450205	0,0015794231
17	0,9	0,2	0,0017450202	0,0015787377
18	0,9	0,2	0,0017450201	0,0015784443
19	0,9	0,2	0,0017450201	0,0015783241
20	0,9	0,2	0,0017450200	0,0015782763
16	1	0,2	0,0017450257	0,0015783070
17	1	0,2	0,0017450221	0,0015782696
18	1	0,2	0,0017450208	0,0015782550
19	1	0,2	0,0017450203	0,0015782494
20	1	0,2	0,0017450201	0,0015782472
16	0,1	0,3	0,0017450809	0,0016850711
17	0,1	0,3	0,0017450423	0,0016850559
18	0,1	0,3	0,0017450281	0,0016850462
19	0,1	0,3	0,0017450230	0,0016850401
20	0,1	0,3	0,0017450211	0,0016850361
16	0,2	0,3	0,0046772069	0,0015782466
17	0,2	0,3	0,0046771868	0,0015782461
18	0,2	0,3	0,0046771681	0,0015782460
19	0,2	0,3	0,0046771507	0,0015782459
20	0,2	0,3	0,0046771344	0,0015782459
16	0,3	0,3	0,0039270613	0,0015782705

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
17	0,3	0,3	0,0021150268	0,0015782553
18	0,3	0,3	0,0017746643	0,0015782495
19	0,3	0,3	0,0017469331	0,0015782473
20	0,3	0,3	0,0017451934	0,0015782464
16	0,4	0,3	0,0017450247	0,0015802723
17	0,4	0,3	0,0017450218	0,0015791198
18	0,4	0,3	0,0017450207	0,0015786059
19	0,4	0,3	0,0017450203	0,0015783898
20	0,4	0,3	0,0017450200	0,0015783023
16	0,5	0,3	0,0046780242	0,0015783274
17	0,5	0,3	0,0046779026	0,0015782776
18	0,5	0,3	0,0046778003	0,0015782581
19	0,5	0,3	0,0046777134	0,0015782506
20	0,5	0,3	0,0046776388	0,0015782477
16	0,6	0,3	0,0046782567	0,0016850999
17	0,6	0,3	0,0046780948	0,0016850742
18	0,6	0,3	0,0046779614	0,0016850579
19	0,6	0,3	0,0046778499	0,0016850475
20	0,6	0,3	0,0046777557	0,0016850409
16	0,7	0,3	0,0017450211	0,0016851888
17	0,7	0,3	0,0017450204	0,0016851305
18	0,7	0,3	0,0017450202	0,0016850936
19	0,7	0,3	0,0017450201	0,0016850702
20	0,7	0,3	0,0017450201	0,0016850553
16	0,8	0,3	0,0046780233	0,0016850326
17	0,8	0,3	0,0046779018	0,0016850314
18	0,8	0,3	0,0046777996	0,0016850306
19	0,8	0,3	0,0046777128	0,0016850301
20	0,8	0,3	0,0046776383	0,0016850298
16	0,9	0,3	0,0017793066	0,0016851858
17	0,9	0,3	0,0017566388	0,0016851286
18	0,9	0,3	0,0017492890	0,0016850924
19	0,9	0,3	0,0017465966	0,0016850694
20	0,9	0,3	0,0017455999	0,0016850548
16	1	0,3	0,0017452127	0,0015782469
17	1	0,3	0,0017450905	0,0015782463
18	1	0,3	0,0017450458	0,0015782460
19	1	0,3	0,0017450294	0,0015782459
20	1	0,3	0,0017450235	0,0015782459

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
16	0,1	0,4	0,0046783632	0,0015784888
17	0,1	0,4	0,0046781817	0,0015783421
18	0,1	0,4	0,0046780333	0,0015782834
19	0,1	0,4	0,0046779101	0,0015782604
20	0,1	0,4	0,0046778067	0,0015782514
16	0,2	0,4	0,0046778626	0,0015782470
17	0,2	0,4	0,0046777665	0,0015782463
18	0,2	0,4	0,0046776844	0,0015782460
19	0,2	0,4	0,0046776137	0,0015782459
20	0,2	0,4	0,0046775523	0,0015782459
16	0,3	0,4	0,0046779827	0,0015782473
17	0,3	0,4	0,0046778678	0,0015782464
18	0,3	0,4	0,0046777708	0,0015782461
19	0,3	0,4	0,0046776882	0,0015782460
20	0,3	0,4	0,0046776170	0,0015782459
16	0,4	0,4	0,0046775624	0,0016851070
17	0,4	0,4	0,0046775073	0,0016850786
18	0,4	0,4	0,0046774587	0,0016850607
19	0,4	0,4	0,0046774156	0,0016850493
20	0,4	0,4	0,0046773771	0,0016850420
16	0,5	0,4	0,0017450539	0,0015782541
17	0,5	0,4	0,0017450324	0,0015782490
18	0,5	0,4	0,0017450245	0,0015782471
19	0,5	0,4	0,0017450217	0,0015782463
20	0,5	0,4	0,0017450206	0,0015782461
16	0,6	0,4	0,0046783627	0,0015782813
17	0,6	0,4	0,0046781812	0,0015782595
18	0,6	0,4	0,0046780328	0,0015782511
19	0,6	0,4	0,0046779098	0,0015782479
20	0,6	0,4	0,0046778064	0,0015782466
16	0,7	0,4	0,0046780434	0,0015790854
17	0,7	0,4	0,0046779185	0,0015785911
18	0,7	0,4	0,0046778138	0,0015783837
19	0,7	0,4	0,0046777249	0,0015782999
20	0,7	0,4	0,0046776487	0,0015782668
16	0,8	0,4	0,0046778287	0,0016852301
17	0,8	0,4	0,0046777377	0,0016851566
18	0,8	0,4	0,0046776597	0,0016851101
19	0,8	0,4	0,0046775923	0,0016850807

F-6

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
20	0,8	0,4	0,0046775336	0,0016850620
16	0,9	0,4	0,0017450204	0,0016850357
17	0,9	0,4	0,0017450202	0,0016850334
18	0,9	0,4	0,0017450201	0,0016850319
19	0,9	0,4	0,0017450201	0,0016850309
20	0,9	0,4	0,0017450200	0,0016850303
16	1	0,4	0,0046781462	0,0015782661
17	1	0,4	0,0046780040	0,0015782537
18	1	0,4	0,0046778857	0,0015782489
19	1	0,4	0,0046777860	0,0015782470
20	1	0,4	0,0046777012	0,0015782463
16	0,1	0,5	0,0046760446	0,0015794365
17	0,1	0,5	0,0046759703	0,0015787436
18	0,1	0,5	0,0046758859	0,0015784467
19	0,1	0,5	0,0046757896	0,0015783251
20	0,1	0,5	0,0046756791	0,0015782767
16	0,2	0,5	0,0017450591	0,0015783473
17	0,2	0,5	0,0017450343	0,0015782854
18	0,2	0,5	0,0017450252	0,0015782612
19	0,2	0,5	0,0017450219	0,0015782517
20	0,2	0,5	0,0017450207	0,0015782481
16	0,3	0,5	0,0046780651	0,0015782925
17	0,3	0,5	0,0046779366	0,0015782639
18	0,3	0,5	0,0046778290	0,0015782528
19	0,3	0,5	0,0046777379	0,0015782485
20	0,3	0,5	0,0046776599	0,0015782469
16	0,4	0,5	0,0046688554	0,0015782854
17	0,4	0,5	0,0046668182	0,0015782611
18	0,4	0,5	0,0046640166	0,0015782517
19	0,4	0,5	0,0046600555	0,0015782481
20	0,4	0,5	0,0046542470	0,0015782467
16	0,5	0,5	0,0046777017	0,0015782470
17	0,5	0,5	0,0046776287	0,0015782463
18	0,5	0,5	0,0046775653	0,0015782460
19	0,5	0,5	0,0046775099	0,0015782459
20	0,5	0,5	0,0046774611	0,0015782459
16	0,6	0,5	0,0017450240	0,0015782510
17	0,6	0,5	0,0017450215	0,0015782478
18	0,6	0,5	0,0017450206	0,0015782466

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
19	0,6	0,5	0,0017450202	0,0015782462
20	0,6	0,5	0,0017450201	0,0015782460
16	0,7	0,5	0,0017450224	0,0015782461
17	0,7	0,5	0,0017450209	0,0015782460
18	0,7	0,5	0,0017450204	0,0015782459
19	0,7	0,5	0,0017450202	0,0015782459
20	0,7	0,5	0,0017450201	0,0015782459
16	0,8	0,5	0,0017450217	0,0015782656
17	0,8	0,5	0,0017450206	0,0015782535
18	0,8	0,5	0,0017450203	0,0015782488
19	0,8	0,5	0,0017450201	0,0015782470
20	0,8	0,5	0,0017450201	0,0015782463
16	0,9	0,5	0,0017450586	0,0015783260
17	0,9	0,5	0,0017450324	0,0015782770
18	0,9	0,5	0,0017450245	0,0015782579
19	0,9	0,5	0,0017450216	0,0015782505
20	0,9	0,5	0,0017450206	0,0015782476
16	1	0,5	0,0046775600	0,0015782777
17	1	0,5	0,0046775052	0,0015782581
18	1	0,5	0,0046774569	0,0015782506
19	1	0,5	0,0046774140	0,0015782477
20	1	0,5	0,0046773756	0,0015782466
16	0,1	0,6	0,0046769743	0,0015782477
17	0,1	0,6	0,0046769653	0,0015782466
18	0,1	0,6	0,0046769566	0,0015782461
19	0,1	0,6	0,0046769481	0,0015782460
20	0,1	0,6	0,0046769397	0,0015782459
16	0,2	0,6	0,0017450315	0,0016851142
17	0,2	0,6	0,0017450242	0,0016850833
18	0,2	0,6	0,0017450215	0,0016850636
19	0,2	0,6	0,0017450206	0,0016850511
20	0,2	0,6	0,0017450202	0,0016850432
16	0,3	0,6	0,0017450836	0,0016856437
17	0,3	0,6	0,0017450432	0,0016853249
18	0,3	0,6	0,0017450285	0,0016851986
19	0,3	0,6	0,0017450231	0,0016851332
20	0,3	0,6	0,0017450212	0,0016850947
16	0,4	0,6	0,0017450219	0,0015782827
17	0,4	0,6	0,0017450207	0,0015782601

F-8

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
18	0,4	0,6	0,0017450203	0,0015782513
19	0,4	0,6	0,0017450201	0,0015782480
20	0,4	0,6	0,0017450201	0,0015782467
16	0,5	0,6	0,0046770204	0,0015782645
17	0,5	0,6	0,0046770094	0,0015782530
18	0,5	0,6	0,0046769993	0,0015782486
19	0,5	0,6	0,0046769898	0,0015782469
20	0,5	0,6	0,0046769805	0,0015782463
16	0,6	0,6	0,0046780445	0,0015798381
17	0,6	0,6	0,0046779195	0,0015789221
18	0,6	0,6	0,0046778146	0,0015785217
19	0,6	0,6	0,0046777256	0,0015783554
20	0,6	0,6	0,0046776493	0,0015782887
16	0,7	0,6	0,0017451421	0,0016851837
17	0,7	0,6	0,0017450646	0,0016851273
18	0,7	0,6	0,0017450363	0,0016850915
19	0,7	0,6	0,0017450260	0,0016850689
20	0,7	0,6	0,0017450222	0,0016850545
16	0,8	0,6	0,0046776090	0,0015782470
17	0,8	0,6	0,0046775482	0,0015782463
18	0,8	0,6	0,0046774948	0,0015782460
19	0,8	0,6	0,0046774477	0,0015782459
20	0,8	0,6	0,0046774058	0,0015782459
16	0,9	0,6	0,0046779349	0,0015782477
17	0,9	0,6	0,0046778277	0,0015782466
18	0,9	0,6	0,0046777368	0,0015782461
19	0,9	0,6	0,0046776589	0,0015782460
20	0,9	0,6	0,0046775916	0,0015782459
16	1	0,6	0,0046781943	0,0016851698
17	1	0,6	0,0046780437	0,0016851185
18	1	0,6	0,0046779189	0,0016850860
19	1	0,6	0,0046778141	0,0016850653
20	1	0,6	0,0046777252	0,0016850522
16	0,1	0,7	0,0017450275	0,0015782613
17	0,1	0,7	0,0017450228	0,0015782518
18	0,1	0,7	0,0017450210	0,0015782481
19	0,1	0,7	0,0017450204	0,0015782467
20	0,1	0,7	0,0017450202	0,0015782462
16	0,2	0,7	0,0017450238	0,0015782478

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
17	0,2	0,7	0,0017450214	0,0015782466
18	0,2	0,7	0,0017450205	0,0015782462
19	0,2	0,7	0,0017450202	0,0015782460
20	0,2	0,7	0,0017450201	0,0015782459
16	0,3	0,7	0,0046777463	0,0015782516
17	0,3	0,7	0,0046776671	0,0015782481
18	0,3	0,7	0,0046775987	0,0015782467
19	0,3	0,7	0,0046775391	0,0015782462
20	0,3	0,7	0,0046774869	0,0015782460
16	0,4	0,7	0,0018945329	0,0015782571
17	0,4	0,7	0,0017555202	0,0015782502
18	0,4	0,7	0,0017457203	0,0015782475
19	0,4	0,7	0,0017451046	0,0015782465
20	0,4	0,7	0,0017450423	0,0015782461
16	0,5	0,7	0,0046777228	0,0015782625
17	0,5	0,7	0,0046776469	0,0015782523
18	0,5	0,7	0,0046775812	0,0015782483
19	0,5	0,7	0,0046775238	0,0015782468
20	0,5	0,7	0,0046774734	0,0015782462
16	0,6	0,7	0,0046770394	0,0015782678
17	0,6	0,7	0,0046770287	0,0015782543
18	0,6	0,7	0,0046770183	0,0015782491
19	0,6	0,7	0,0046770083	0,0015782471
20	0,6	0,7	0,0046769985	0,0015782463
16	0,7	0,7	0,0046782545	0,0015782480
17	0,7	0,7	0,0046780932	0,0015782467
18	0,7	0,7	0,0046779601	0,0015782462
19	0,7	0,7	0,0046778488	0,0015782460
20	0,7	0,7	0,0046777548	0,0015782459
16	0,8	0,7	0,0046790792	0,0015783444
17	0,8	0,7	0,0046787488	0,0015782843
18	0,8	0,7	0,0046784906	0,0015782607
19	0,8	0,7	0,0046782848	0,0015782516
20	0,8	0,7	0,0046781179	0,0015782481
16	0,9	0,7	0,0017450447	0,0015784343
17	0,9	0,7	0,0017450290	0,0015783201
18	0,9	0,7	0,0017450233	0,0015782747
19	0,9	0,7	0,0017450212	0,0015782570
20	0,9	0,7	0,0017450205	0,0015782501

<i>Max Iter</i>	<i>Learning Rate</i>	<i>Weights</i>	<i>SSE Training</i>	<i>SSE Testing</i>
16	1	0,7	0,0017450211	0,0015782673
17	1	0,7	0,0017450204	0,0015782541
18	1	0,7	0,0017450202	0,0015782490
19	1	0,7	0,0017450201	0,0015782471
20	1	0,7	0,0017450201	0,0015782463
16	0,1	0,8	0,0017451835	0,0015782603
17	0,1	0,8	0,0017450798	0,0015782514
18	0,1	0,8	0,0017450418	0,0015782480
19	0,1	0,8	0,0017450280	0,0015782467
20	0,1	0,8	0,0017450229	0,0015782462
16	0,2	0,8	0,0017450225	0,0015782691
17	0,2	0,8	0,0017450209	0,0015782548
18	0,2	0,8	0,0017450204	0,0015782493
19	0,2	0,8	0,0017450202	0,0015782472
20	0,2	0,8	0,0017450201	0,0015782464
16	0,3	0,8	0,0046784686	0,0015782484
17	0,3	0,8	0,0046782669	0,0015782468
18	0,3	0,8	0,0046781032	0,0015782462
19	0,3	0,8	0,0046779684	0,0015782460
20	0,3	0,8	0,0046778558	0,0015782459
16	0,4	0,8	0,0017450204	0,0016852164
17	0,4	0,8	0,0017450202	0,0016851480
18	0,4	0,8	0,0017450201	0,0016851047
19	0,4	0,8	0,0017450201	0,0016850772
20	0,4	0,8	0,0017450200	0,0016850598
16	0,5	0,8	0,0046777685	0,0015782670
17	0,5	0,8	0,0046776862	0,0015782540
18	0,5	0,8	0,0046776153	0,0015782490
19	0,5	0,8	0,0046775537	0,0015782471
20	0,5	0,8	0,0046774997	0,0015782463
16	0,6	0,8	0,0046775660	0,0015782465
17	0,6	0,8	0,0046775105	0,0015782461
18	0,6	0,8	0,0046774615	0,0015782460
19	0,6	0,8	0,0046774181	0,0015782459
20	0,6	0,8	0,0046773793	0,0015782459
16	0,7	0,8	0,0017450713	0,0015783147
17	0,7	0,8	0,0017450388	0,0015782726
18	0,7	0,8	0,0017450269	0,0015782562
19	0,7	0,8	0,0017450225	0,0015782498

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
20	0,7	0,8	0,0017450209	0,0015782474
16	0,8	0,8	0,0017450223	0,0016855852
17	0,8	0,8	0,0017450209	0,0016853151
18	0,8	0,8	0,0017450203	0,0016851977
19	0,8	0,8	0,0017450201	0,0016851337
20	0,8	0,8	0,0017450201	0,0016850952
16	0,9	0,8	0,0046706505	0,0015782577
17	0,9	0,8	0,0046691934	0,0015782504
18	0,9	0,8	0,0046672414	0,0015782476
19	0,9	0,8	0,0046645676	0,0015782465
20	0,9	0,8	0,0046608063	0,0015782461
16	1	0,8	0,0046745561	0,0015782548
17	1	0,8	0,0046741686	0,0015782493
18	1	0,8	0,0046737197	0,0015782472
19	1	0,8	0,0046731765	0,0015782464
20	1	0,8	0,0046725010	0,0015782461
16	0,1	0,9	0,0017450809	0,0015782527
17	0,1	0,9	0,0017450423	0,0015782485
18	0,1	0,9	0,0017450281	0,0015782469
19	0,1	0,9	0,0017450230	0,0015782463
20	0,1	0,9	0,0017450211	0,0015782460
16	0,2	0,9	0,0046765851	0,0015785255
17	0,2	0,9	0,0046765578	0,0015783570
18	0,2	0,9	0,0046765332	0,0015782893
19	0,2	0,9	0,0046765089	0,0015782626
20	0,2	0,9	0,0046764837	0,0015782523
16	0,3	0,9	0,0046778239	0,0015784192
17	0,3	0,9	0,0046777335	0,0015783140
18	0,3	0,9	0,0046776561	0,0015782723
19	0,3	0,9	0,0046775892	0,0015782560
20	0,3	0,9	0,0046775308	0,0015782498
16	0,4	0,9	0,0046783180	0,0015783722
17	0,4	0,9	0,0046781448	0,0015782953
18	0,4	0,9	0,0046780028	0,0015782650
19	0,4	0,9	0,0046778847	0,0015782532
20	0,4	0,9	0,0046777852	0,0015782487
16	0,5	0,9	0,0046762386	0,0015782518
17	0,5	0,9	0,0046761878	0,0015782481
18	0,5	0,9	0,0046761313	0,0015782467

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
19	0,5	0,9	0,0046760679	0,0015782462
20	0,5	0,9	0,0046759965	0,0015782460
16	0,6	0,9	0,0046361605	0,0015783240
17	0,6	0,9	0,0046140316	0,0015782763
18	0,6	0,9	0,0045669156	0,0015782576
19	0,6	0,9	0,0044265353	0,0015782504
20	0,6	0,9	0,0037175791	0,0015782476
16	0,7	0,9	0,0017450658	0,0015782527
17	0,7	0,9	0,0017450367	0,0015782485
18	0,7	0,9	0,0017450261	0,0015782469
19	0,7	0,9	0,0017450223	0,0015782463
20	0,7	0,9	0,0017450208	0,0015782460
16	0,8	0,9	0,0017450228	0,0015798112
17	0,8	0,9	0,0017450210	0,0015789101
18	0,8	0,9	0,0017450204	0,0015785166
19	0,8	0,9	0,0017450202	0,0015783534
20	0,8	0,9	0,0017450201	0,0015782878
16	0,9	0,9	0,0046779317	0,0016851517
17	0,9	0,9	0,0046778248	0,0016851070
18	0,9	0,9	0,0046777343	0,0016850787
19	0,9	0,9	0,0046776568	0,0016850607
20	0,9	0,9	0,0046775898	0,0016850493
16	1	0,9	0,0046779068	0,0016851817
17	1	0,9	0,0046778039	0,0016851260
18	1	0,9	0,0046777165	0,0016850907
19	1	0,9	0,0046776414	0,0016850684
20	1	0,9	0,0046775764	0,0016850542
16	0,1	1	0,0046769814	0,0015791229
17	0,1	1	0,0046769723	0,0015786073
18	0,1	1	0,0046769635	0,0015783904
19	0,1	1	0,0046769549	0,0015783026
20	0,1	1	0,0046769465	0,0015782678
16	0,2	1	0,0046779741	0,0015786565
17	0,2	1	0,0046778606	0,0015784106
18	0,2	1	0,0046777648	0,0015783106
19	0,2	1	0,0046776830	0,0015782710
20	0,2	1	0,0046776125	0,0015782555
16	0,3	1	0,0046783239	0,0015783058
17	0,3	1	0,0046781498	0,0015782691

Max Iter	Learning Rate	Weights	SSE Training	SSE Testing
18	0,3	1	0,0046780069	0,0015782548
19	0,3	1	0,0046778881	0,0015782493
20	0,3	1	0,0046777881	0,0015782472
16	0,4	1	0,0017450250	0,0015782468
17	0,4	1	0,0017450218	0,0015782462
18	0,4	1	0,0017450207	0,0015782460
19	0,4	1	0,0017450203	0,0015782459
20	0,4	1	0,0017450201	0,0015782459
16	0,5	1	0,0017450819	0,0015782462
17	0,5	1	0,0017450426	0,0015782460
18	0,5	1	0,0017450283	0,0015782459
19	0,5	1	0,0017450230	0,0015782459
20	0,5	1	0,0017450211	0,0015782459
16	0,6	1	0,0017450819	0,0015782462
17	0,6	1	0,0017450426	0,0015782460
18	0,6	1	0,0017450283	0,0015782459
19	0,6	1	0,0017450230	0,0015782459
20	0,6	1	0,0017450211	0,0015782459
16	0,7	1	0,0046776992	0,0015783208
17	0,7	1	0,0046776265	0,0015782750
18	0,7	1	0,0046775634	0,0015782571
19	0,7	1	0,0046775082	0,0015782502
20	0,7	1	0,0046774596	0,0015782475
16	0,8	1	0,0046772066	0,0015782546
17	0,8	1	0,0046771866	0,0015782492
18	0,8	1	0,0046771680	0,0015782472
19	0,8	1	0,0046771506	0,0015782464
20	0,8	1	0,0046771343	0,0015782461
16	0,9	1	0,0017451516	0,0016851801
17	0,9	1	0,0017450681	0,0016851250
18	0,9	1	0,0017450376	0,0016850901
19	0,9	1	0,0017450264	0,0016850680
20	0,9	1	0,0017450224	0,0016850539
16	1	1	0,0017450274	0,0015782536
17	1	1	0,0017450227	0,0015782488
18	1	1	0,0017450210	0,0015782470
19	1	1	0,0017450204	0,0015782463
20	1	1	0,0017450202	0,0015782460

Halaman ini sengaja dikosongkan

LAMPIRAN G

Tabel G.1 Keseluruhan Validasi Derajat Keanggotaan Pelanggan

<i>Customer ID</i>	<i>C 1</i>	<i>C 2</i>	<i>C 3</i>	<i>Max</i>	<i>Cluster</i>	<i>Validasi</i>
2053	0,102	0,134	0,764	0,764	3	Valid
2138	0,070	0,916	0,013	0,916	2	Valid
2166	0,336	0,369	0,296	0,369	2	Valid
700032	0,148	0,322	0,530	0,530	3	Valid
700189	0,876	0,109	0,015	0,876	1	Valid
700286	0,109	0,882	0,009	0,882	2	Valid
700385	0,897	0,096	0,006	0,897	1	Valid
700493	0,752	0,205	0,043	0,752	1	Valid
700503	0,077	0,914	0,010	0,914	2	Valid
700529	0,155	0,777	0,068	0,777	2	Valid
700793	0,021	0,975	0,004	0,975	2	Valid
701121	0,022	0,038	0,940	0,940	3	Valid
701976	0,035	0,050	0,915	0,915	3	Valid
702223	0,022	0,036	0,942	0,942	3	Valid
702643	0,388	0,586	0,026	0,586	2	Valid
702761	0,513	0,473	0,014	0,513	1	Valid
702920	0,849	0,132	0,020	0,849	1	Valid
704054	0,078	0,914	0,008	0,914	2	Valid
704402	0,199	0,543	0,258	0,543	2	Valid
704622	0,033	0,962	0,004	0,962	2	Valid
704818	0,168	0,778	0,054	0,778	2	Valid
704838	0,348	0,615	0,037	0,615	2	Valid
704904	0,032	0,048	0,920	0,920	3	Valid
705329	0,331	0,337	0,332	0,337	2	Valid
705416	0,078	0,907	0,015	0,907	2	Valid
705759	0,333	0,339	0,328	0,339	2	Valid
705762	0,329	0,335	0,336	0,336	3	Valid
706383	0,404	0,437	0,158	0,437	2	Valid
706737	0,343	0,352	0,305	0,352	2	Valid
706806	0,391	0,515	0,094	0,515	2	Valid
710854	0,240	0,745	0,015	0,745	2	Valid
710856	0,377	0,603	0,019	0,603	2	Valid
710858	0,181	0,805	0,014	0,805	2	Valid

G-2

<i>Customer ID</i>	<i>C 1</i>	<i>C 2</i>	<i>C 3</i>	<i>Max</i>	<i>Cluster</i>	<i>Validasi</i>
710868	0,191	0,501	0,307	0,501	2	Valid
710869	0,023	0,036	0,941	0,941	3	Valid
711781	0,166	0,746	0,089	0,746	2	Valid
711810	0,794	0,195	0,010	0,794	1	Valid
711831	0,189	0,802	0,010	0,802	2	Valid
712559	0,260	0,710	0,030	0,710	2	Valid
712561	0,136	0,284	0,580	0,580	3	Valid
712562	0,256	0,731	0,013	0,731	2	Valid
712563	0,173	0,814	0,013	0,814	2	Valid
712564	0,395	0,591	0,014	0,591	2	Valid
712565	0,926	0,067	0,007	0,926	1	Valid
712566	0,165	0,824	0,011	0,824	2	Valid
712717	0,039	0,056	0,905	0,905	3	Valid
712746	0,057	0,081	0,862	0,862	3	Valid
712748	0,009	0,014	0,977	0,977	3	Valid
712785	0,091	0,886	0,023	0,886	2	Valid
712843	0,047	0,942	0,011	0,942	2	Valid
712872	0,077	0,915	0,008	0,915	2	Valid
713115	0,345	0,638	0,017	0,638	2	Valid
713136	0,176	0,415	0,409	0,415	2	Valid
713167	0,498	0,485	0,017	0,498	1	Valid
713249	0,530	0,448	0,022	0,530	1	Valid
713253	0,190	0,630	0,180	0,630	2	Valid
713255	0,195	0,628	0,177	0,628	2	Valid
713320	0,851	0,139	0,010	0,851	1	Valid
714954	0,264	0,721	0,015	0,721	2	Valid
715470	0,163	0,368	0,469	0,469	3	Valid
719930	0,134	0,854	0,012	0,854	2	Valid
719971	0,673	0,314	0,013	0,673	1	Valid
720145	0,022	0,038	0,940	0,940	3	Valid
720236	0,973	0,025	0,002	0,973	1	Valid
721918	0,731	0,256	0,013	0,731	1	Valid
721954	0,924	0,071	0,005	0,924	1	Valid
722149	0,485	0,494	0,021	0,494	2	Valid
722171	0,261	0,533	0,206	0,533	2	Valid
728260	0,984	0,015	0,001	0,984	1	Valid
728447	0,190	0,700	0,110	0,700	2	Valid
729665	0,160	0,369	0,472	0,472	3	Valid
730112	0,498	0,485	0,017	0,498	1	Valid

<i>Customer ID</i>	<i>C 1</i>	<i>C 2</i>	<i>C 3</i>	<i>Max</i>	<i>Cluster</i>	<i>Validasi</i>
730648	0,738	0,214	0,048	0,738	1	Valid
730948	0,086	0,896	0,019	0,896	2	Valid
731081	0,356	0,624	0,021	0,624	2	Valid
731090	0,818	0,171	0,010	0,818	1	Valid
731737	0,869	0,117	0,014	0,869	1	Valid
735810	0,097	0,890	0,013	0,890	2	Valid
737744	0,616	0,365	0,019	0,616	1	Valid
738593	0,949	0,046	0,005	0,949	1	Valid
739995	0,153	0,836	0,011	0,836	2	Valid
743352	0,029	0,049	0,921	0,921	3	Valid
746205	0,975	0,023	0,002	0,975	1	Valid
751408	0,023	0,974	0,003	0,974	2	Valid
757273	0,078	0,914	0,008	0,914	2	Valid
758186	0,959	0,038	0,004	0,959	1	Valid
762727	0,994	0,006	0,000	0,994	1	Valid
766881	0,220	0,659	0,120	0,659	2	Valid
777121	0,566	0,416	0,018	0,566	1	Valid
777897	0,829	0,147	0,024	0,829	1	Valid
788346	0,967	0,030	0,003	0,967	1	Valid
793566	0,007	0,992	0,001	0,992	2	Valid
794521	0,877	0,107	0,015	0,877	1	Valid
795940	0,788	0,178	0,033	0,788	1	Valid
804967	0,290	0,693	0,017	0,693	2	Valid
806038	0,828	0,148	0,024	0,828	1	Valid
806791	0,518	0,467	0,015	0,518	1	Valid
807046	0,073	0,913	0,014	0,913	2	Valid
819874	0,270	0,710	0,019	0,710	2	Valid
820764	0,170	0,787	0,042	0,787	2	Valid
821655	0,177	0,707	0,115	0,707	2	Valid
821868	0,057	0,079	0,864	0,864	3	Valid
822231	0,824	0,161	0,014	0,824	1	Valid
823661	0,812	0,177	0,011	0,812	1	Valid
823940	0,081	0,911	0,008	0,911	2	Valid
826451	0,407	0,574	0,019	0,574	2	Valid
828392	0,924	0,069	0,006	0,924	1	Valid
832323	0,879	0,113	0,008	0,879	1	Valid
832593	0,903	0,088	0,009	0,903	1	Valid
848041	0,203	0,781	0,016	0,781	2	Valid
850911	0,190	0,800	0,010	0,800	2	Valid

G-4

<i>Customer ID</i>	<i>C 1</i>	<i>C 2</i>	<i>C 3</i>	<i>Max</i>	<i>Cluster</i>	<i>Validasi</i>
853537	0,170	0,743	0,087	0,743	2	Valid
858218	0,843	0,135	0,022	0,843	1	Valid
869214	0,919	0,075	0,006	0,919	1	Valid
883122	0,054	0,936	0,010	0,936	2	Valid
887934	0,755	0,203	0,043	0,755	1	Valid
888124	0,173	0,815	0,013	0,815	2	Valid
890697	0,424	0,554	0,022	0,554	2	Valid
892055	0,856	0,126	0,019	0,856	1	Valid
892303	0,066	0,925	0,009	0,925	2	Valid
893393	0,895	0,098	0,007	0,895	1	Valid
893562	0,484	0,496	0,020	0,496	2	Valid
893809	0,899	0,093	0,007	0,899	1	Valid

LAMPIRAN H

Keseluruhan Script Server.R

```
library(shiny)
library(leaflet)
library(plot3drgl)
library(sqldf)
library(shinydashboard)
library(plotly)
dashboardPage(skin = "red",
              dashboardHeader(title = "PT. XYZ Surabaya"),
              dashboardSidebar(
                sidebarMenu(
                  menuItem("Halaman Utama", tabName = "home", icon = icon("home")),
                  menuItem("visualisasi Segmen", tabName = "detail", icon = icon("dashboard")),
                  menuItem("Tabel", tabName = "tabel", icon = icon("table"))
                )
              ),
              dashboardBody(
                tabItems(
                  tabItem("home",
                          fluidRow(),
                          fluidRow(
                            column(width = 12,
                                   h1("Visualisasi Segmentasi Pelanggan PT. XYZ Surabaya"))
                          )),
                  fluidRow(
                    column(width = 6,
                           box(title = "Jumlah Anggota per Segmen",
                               status = "info",
                               width = 50,
                               plotlyOutput("barchart")))
                ),
                tabItem("detail",
                        fluidRow(
                          column(width = 6,
                                 box(title = "Persentase Jumlah Anggota per Segmen",
                                     status = "info",
                                     width = 50,
                                     plotlyOutput("pie")))
                ),
                tabItem("tabel")
              ))
)
```

H-2

```
tabItem("detail",
fluidPage(
  column(width=12,
    box(width = NULL,
        background = "navy",
        h4(strong(textOutput('judul'))),
        br(),
        selectInput("dataset", "Pilih Jenis Segmen",
          choices = c(
            "Semua Segmen",
            "Segmen 1",
            "Segmen 2",
            "Segmen 3")
      )
    ),
    column(width=12,
      fluidRow(
        infoBoxOutput("recency"),
        infoBoxOutput("frequency"),
        infoBoxOutput("monetary")
      ),
      box( width = 100,
        title = "Rentang Segmen",
        background = "navy",
        solidHeader = TRUE,
        collapsible = TRUE,
        collapsed = FALSE,
        strong(textOutput('text')),
        verbatimTextOutput('range')
      ),
      box( width = 100,
        title = "Pelanggan Teratas dan Terbawah",
        background = "navy",
        solidHeader = TRUE,
        collapsible = TRUE,
        collapsed = FALSE,
        valueBoxOutput("teratas"),
        valueBoxOutput("terbawah")
      )
    )
  fluidRow(
    column(width=12,
      box(width = 400,
          title = "Grafik 3 Dimensi",
          solidHeader = TRUE,
          collapsible = TRUE,
          collapsed = FALSE,
          rglwidgetOutput("scatter", height = 500, width
= 500),
          box(width = NULL,
              uioutput("tabelsegmen")
            )
      )
    )
  )
),
```

```
fluidRow(  
  column(width=12,  
    box(width = 800,  
      title = "Peringkat Pelanggan",  
      solidHeader = TRUE,  
      collapsible = TRUE,  
      collapsed = FALSE,  
      dataTableOutput("clv")  
    )  
  ),  
  fluidRow(column(width=6,  
    box( width = 300,  
      title = "CLV",  
      background = "navy",  
      solidHeader = TRUE,  
      collapsible = TRUE,  
      collapsed = FALSE,  
      plotlyOutput ("boxclv")  
  )),  
  column(width=6,  
    box( width = 300,  
      title = "Recency",  
      background = "navy",  
      solidHeader = TRUE,  
      collapsible = TRUE,  
      collapsed = FALSE,  
      plotlyOutput ("boxr")  
  ))  
,  
  fluidRow(column(width=6,  
    box( width = 300,  
      title = "Frequency",  
      background = "navy",  
      solidHeader = TRUE,  
      collapsible = TRUE,  
      collapsed = FALSE,  
      plotlyOutput ("boxf")  
  )),  
  column(width=6,  
    box( width = 300,  
      title = "Monetary",  
      background = "navy",  
      solidHeader = TRUE,  
      collapsible = TRUE,  
      collapsed = FALSE,  
      plotlyOutput ("boxm")  
  ))  
)  
)  
)  
,  
tabItem("table",  
  dataTableOutput('table')  
}  
}  
}
```

Keseluruhan Script UI.R

```

library(shiny)
library(leaflet)
library(plot3drgl)
library(sqldf)
library(shinydashboard)
library(plotly)
function(input, output, session) {

  #ALL
  alldata <- HasilRFMBobot
  data <- HasiKlaster
  pilihdata <- reactive({
    switch(input$dataset,
      "Semua Segmen" = data,
      "Segmen 1"= subset(data,data$c1.cluster==1),
      "Segmen 2"= subset(data,data$c1.cluster==2),
      "Segmen 3"= subset(data,data$c1.cluster==3)
    )
  })

  #headline1
  output$ket1<- renderValueBox({
    valueBox(
      value = "123",
      subtitle = "Pelanggan",
      icon = icon(""),
      color = "blue"
    )
  })
  #headline2

  output$ket2<- renderValueBox({
    valueBox(
      value = "100.000.000",
      subtitle = "Pemasukan",
      icon = icon("money"),
      color = "blue"
    )
  })
  #headline3
  output$ket3<- renderValueBox({
    valueBox(
      value = "1340",
      subtitle = "Transaksi",
      icon = icon("shopping-cart"),
      color = "blue"
    )
  })

  #Data Masukan
  segmen <- c("Segmen 1", "Segmen 2", "Segmen 3")
  value <- c(sum(data$c1.cluster == "1"),
            sum(data$c1.cluster == "2"),
            sum(data$c1.cluster == "3"))
}

```

```

#Bar Chart
output$bchart <- renderPlotly ({
  plot_ly(
    x = segmen,
    y = value,
    name = "Jenis Segmen",
    type = "bar",
    hoverinfo = 'text',
    text = paste(value, ' Pelanggan'),
    marker = list(color = warna)
  )
})

#Kode Warna
warna <- c("#00FFFF", "#8A2BE2", "#E9967A")

#Pie Chart
output$pie <- renderPlotly ({
  plot_ly(labels = segmen, values = value, type = 'pie',
    textposition = 'inside',
    textinfo = 'label+percent',
    insidetextfont = list(color = '#FFFFFF'),
    hoverinfo = 'text',
    text = paste(value, ' Pelanggan'),
    marker = list(colors = warna,
      line = list(color = '#FFFFFF', width = 1))
  )
})

#Dropdown menu segmen
output$judul <- renderText({
  paste("Menampilkan Detail", input$dataset)
})

#Rata-rata Recency
output$recency <- renderInfoBox({
  data <- pilihdata()
  infoBox(
    title = "Recency Average",
    value = paste(round(mean(data$R)), "hari"),
    subtitle = "Terakhir transaksi",
    icon = icon("calendar"),
    color = "navy",
    fill = TRUE
  )
})

#Rata-rata Frequency
output$frequency <- renderInfoBox({
  data <- pilihdata()
  infoBox(
    title = "Frequency Average",
    value = paste(round(mean(data$F)), "kali"),
    subtitle = "Banyak transaksi",
    icon = icon("area-chart"),
    color = "navy",
    fill = TRUE
  )
})

```

```

#Rata-rata Monetary
output$monetary <- renderInfoBox({
  data <- pilihdata()
  infoBox(
    title = "Monetary_Average",
    value = paste("Rp", format(round((mean(data$M)),1), nsmall = 0, big.mark = "."),
    subtitle = "Jumlah biaya",
    icon = icon("money"),
    color = "navy",
    fill = TRUE
  )
})

#Text Rentang Nilai
output$text <- renderText({
  paste("Rentang Nilai", input$dataset)
})

#Rentang RFM
output$range <- renderPrint({
  rentang <- pilihdata()
  list(Recency = paste(min(rentang$R), "-", max(rentang$R), "hari"),
       Frequency = paste(min(rentang$F), "-", max(rentang$F), "kali"),
       Monetary = paste ("Rp", format(round(min(rentang$M),1), nsmall = 0, big.mark = "."),
       subtitle = "-","Rp", format(round(max(rentang$M),1), nsmall = 0, big.mark = "."))
  )
})

#Pelanggan Teratas
output$teratas<- renderValueBox({
  inputan <- pilihdata()
  cari <- sqldf("Select Customer From(select Customer,max(CustomerID) From inputan")
  )
  valueBox(
    value = tags$h5( "PELANGGAN TERATAS"),
    subtitle = tags$h4(tags$strong (cari)),
    icon = icon("thumbs-o-up"),
    color = "navy"
  )
})

#Pelanggan Terbawah
output$terbawah<- renderValueBox({
  inputan <- pilihdata()
  cari <- sqldf("Select Customer From(select Customer,min(CustomerID) From inputan")
  )
  valueBox(
    value = tags$h5( "PELANGGAN TERBAWAH"),
    subtitle = tags$h4(tags$strong (cari)),
    icon = icon("thumbs-o-down"),
    color = "navy"
  )
})

```

```

#Scatter Plot 3 Dimensi
output$scatter <- renderRglwidget ({
  splot <- pilihdata()
  open3d()
  plot3d(x=splot$R, y=splot$F, z=splot$M,
          col = warna[splot$c].cluster,
          type = 's', size = 1,
          xlab = "recency",
          ylab = "frequency",
          zlab = "monetary",
          colkey = list(length = 0.5, width = 0.5,
                        cex.clab = 0.75)
  )
  scene1 <- scene3d()
  rgl.close()
  save <- options(rgl.inShiny = TRUE)
  on.exit(options(save))
  rglwidget(scene1)
})

#Tabel peringkat
output$clv <- renderDataTable({
  tabel <- pilihdata()
  tampil <- data.frame(sqlDF("Select Customer, R, F, M, CLV
From tabel order by CLV DESC")
  )
  colnames(tampil) <- c("Customer", "Recency", "Frequency", "Monetary", "CLV")
  print(tampil)
},
list(lengthMenu = c(5, 10, 15), pageLength = 8))

#Keterangan Warna
output$tabelsegmen <- renderUI({
  dirColors <-c("1"="#00FFFF", "2"="#8A2BE2", "3"="#E9967A")
}
tags$table(class = "table",
           tags$thead(tags$str(
             tags$th("Warna"),
             tags$th("Segmen")
           )),
           tags$tbody(
             tags$str(
               tags$td(span(style = sprintf(
                 "width:1.1em; height:1.1em; background-color:%s; display:
                   inline-block;",
                 dirColors[1]
               ))),
               tags$td("Segmen 1")
             ),
             tags$str(
               tags$td(span(style = sprintf(
                 "width:1.1em; height:1.1em; background-color:%s; display:
                   inline-block;",
                 dirColors[2]
               ))),
               tags$td("Segmen 2")
             )
           )
}

```

```

tags$tr(
  tags$td(span(style = sprintf(
    "width:1.1em; height:1.1em; background-c
olor:%s; display:
      inline-block;",
      dirColors[3]
    ))),
  tags$td("Segmen 3")
)
}
}

#Data box plot
y1 <- subset(data,data$c1.cluster==1)
y2 <- subset(data,data$c1.cluster==2)
y3 <- subset(data,data$c1.cluster==3)

#Box Plot
output$boxr <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = y1$R, name = "1",
                marker = list(color = "#00FFFF"),
                line = list(color = "#00FFFF")) %>%
    add_boxplot(y = y2$R, name = "2",
                marker = list(color = "#8A2BE2"),
                line = list(color = "#8A2BE2")) %>%
    add_boxplot(y = y3$R, name = "3",
                marker = list(color = "#E9967A"),
                line = list(color = "#E9967A"))
})
output$boxf <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = y1$F, name = "1",
                marker = list(color = "#00FFFF"),
                line = list(color = "#00FFFF")) %>%
    add_boxplot(y = y2$F, name = "2",
                marker = list(color = "#8A2BE2"),
                line = list(color = "#8A2BE2")) %>%
    add_boxplot(y = y3$F, name = "3",
                marker = list(color = "#E9967A"),
                line = list(color = "#E9967A"))
})
output$boxm <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = y1$M, name = "1",
                marker = list(color = "#00FFFF"),
                line = list(color = "#00FFFF")) %>%
    add_boxplot(y = y2$M, name = "2",
                marker = list(color = "#8A2BE2"),
                line = list(color = "#8A2BE2")) %>%
    add_boxplot(y = y3$M, name = "3",
                marker = list(color = "#E9967A"),
                line = list(color = "#E9967A"))
})

```

```
output$boxclv <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = y1$CLV, name = "1",
                marker = list(color = "#00FFFF"),
                line = list(color = "#00FFFF")) %>%
    add_boxplot(y = y2$CLV, name = "2",
                marker = list(color = "#8A2BE2"),
                line = list(color = "#8A2BE2")) %>%
    add_boxplot(y = y3$CLV, name = "3",
                marker = list(color = "#E9967A"),
                line = list(color = "#E9967A"))
})

#Data Tabel Keseluruhan
output$table = renderDataTable({
  tabel <- data
  tampil <- data.frame(tabel$Customer_ID,
                        tabel$Customer,
                        tabel$R,
                        tabel$F,
                        tabel$M,
                        tabel$CLV,
                        tabel$c1.cluster
  )
  colnames(tampil) <- c("Customer_Code",
                        "Customer",
                        "Recency",
                        "Frequency",
                        "Monetary",
                        "CLV",
                        "Segmen"
  )
  print(tampil)
},
list(lengthMenu = c(5, 10, 15), pageLength = 10))
})
```