



**TUGAS AKHIR - KS141501**

**IMPLEMENTASI *FUZZY C-MEANS* DAN MODEL LRFM UNTUK SEGMENTASI PELANGGAN (STUDI KASUS : PT. ALIF DUTA PERSADA)**

***IMPLEMENTATION OF FUZZY C-MEANS AND LRFM MODEL FOR CUSTOMER SEGMENTATION (CASE STUDY : PT. ALIF DUTA PERSADA)***

Nolan Firdaus Badrid Duja  
NRP 0521144000002

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  
LRFM UNTUK SEGMENTASI PELANGGAN (STUDI  
KASUS : PT. ALIF DUTA PERSADA)**

**Nolan Firdaus Badrid Duja  
NRP 0521144000002**

**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  
LRFM MODEL FOR CUSTOMER SEGMENTATION  
(CASE STUDY : PT. ALIF DUTA PERSADA)**

**Nolan Firdaus Badrid Duja  
NRP 0521144000002**

**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  
LRFM UNTUK SEGMENTASI PELANGGAN (STUDI  
KASUS PT. ALIF DUTA PERSADA)**

**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:

**NOLAN FIRDAUS BADRID DUJA**

**NRP. 0521144000002**

Surabaya, Juli 2018

**KETUA  
DEPARTEMEN SISTEM INFORMASI**



**Dr. Ir. Aris Tjahyanto, M.Kom.**

**NIP. 196503101991021001**



**IMPLEMENTASI FUZZY C-MEANS DAN MODEL  
LRFM UNTUK SEGMENTASI PELANGGAN (STUDI  
KASUS : PT. ALIF DUTA PERSADA)**

**Nama Mahasiswa : Nolan Firdaus Badrid Duja**  
**NRP : 0521144000002**  
**Departemen : Sistem Informasi**  
**Pembimbing I : Edwin Riksakomara, S.Kom, M.T.**

**ABSTRAK**

*Pada era digital saat ini data menjadi sumber daya yang sangat penting dalam berbagai sektor bisnis, termasuk sektor perdagangan. Dalam sektor perdagangan, data transaksi pelanggan menjadi komponen yang sangat penting, untuk menghasilkan banyak informasi penting bagi perusahaan. Hal ini memaksa perusahaan-perusahaan yang bergerak di sektor perdagangan berlomba-lomba untuk memanfaatkan dan mengolah data transaksi pelanggan yang mereka miliki. Dalam hal ini perusahaan yang mampu memanfaatkan data mereka dengan baik maka akan mendapatkan banyak keuntungan dan perusahaan yang tidak memanfaatkan data dalam bisnisnya maka akan tertinggal. Hal tersebut juga di sadari oleh PT. Alif Duta Persada, salah satu distributor PT. Unilever Indonesia, Tbk. Sehingga merasa perlu untuk memanfaatkan data transaksi pelanggan yang mereka miliki untuk menentukan strategi bisnis atau pemasaran mereka.*

*Untuk menyelesaikan masalah tersebut, salah satu solusi yang dapat dilakukan yaitu dengan melakukan segmentasi pelanggan. Segmentasi yang dilakukan menggunakan metode clustering Fuzzy C-Means. Sedangkan untuk menentukan jumlah K, menggunakan metode Elbow. Dalam proses clustering, diterapkan variable LRFM (Length, Recency,*

*Frequency, dan Monetary) yang menggambarkan perilaku pelanggan dalam melakukan transaksi. Dan ketika clustering selesai dilakukan, maka tahap selanjutnya yaitu memvisualisasikan setiap cluster agar mudah untuk dianalisa.*

*Luaran dari penelitian ini adalah berupa segmen-segmen pelanggan yang terbentuk berdasarkan hasil dari kemiripan perilaku pelanggan pada setiap cluster. Dengan adanya segmentasi ini diharapkan dapat membantu PT. Alif Duta Persada untuk menerapkan strategi pemasaran dan pemberian promo pada pelanggan mereka, berdasarkan hasil segmentasi mereka. Hasil dari clustering dengan Fuzzy C-Means dan model LRFM adalah tiga segmen pelanggan. Segmen 1 adalah segmen terbaik dengan rata-rata nilai CLV 0.187. Segmen 3 adalah segmen menengah dengan rata-rata nilai CLV 0.114. Segmen 2 adalah segmen terburuk dengan rata-rata nilai CLV 0.07. Segmen 1 memiliki 140 anggota, segmen 2 memiliki 378 anggota, dan segmen 3 memiliki 644 anggota.*

***Kata Kunci : Segmentasi Pelanggan, Clustering, Fuzzy C-Means, Model LRFM***

**IMPLEMENTATION OF FUZZY C-MEANS AND  
LRFM MODEL FOR CUSTOMER SEGMENTATION  
(CASE STUDY : PT. ALIF DUTA PERSADA)**

**Student Name** : Nolan Firdaus Badrid Duja  
**NRP** : 0521144000002  
**Department** : Sistem Informasi  
**Supervisor I** : Edwin Riksakomara, S.Kom, M.T.

**ABSTRACT**

*In today's digital era, data becomes a very important resource in many business sectors, including the trade sector. In the trading sector, customer transaction data becomes a very important component, to generate a lot of important information for the company. This forced the companies engaged in the trading sector competing to utilize and process customer transaction data that they have. In this case companies that are able to utilize their data well it will get a lot of profits and companies that do not use the data in the business it will be left behind. It is also realized by PT. Alif Duta Persada, one of the distributor of PT. Unilever Indonesia, Tbk. So feel the need to utilize the customer transaction data they have to determine their business or marketing strategy.*

*To solve the problem, there is one solution that can be done is to do customer segmentation. Segmentation is done using Fuzzy C-Means clustering method. Meanwhile, to determine the number of K, using Elbow method. In the clustering process, applied LRFM (Length, Recency, Frequency, and Monetary) variables that describe the behavior of customers in conducting*

*transactions. And when the clustering is done, the next step is to visualize each cluster for easy analysis.*

*The output of this research is in the form of customer segments formed based on the result of the resemblance of customer behavior on each cluster. With this segmentation is expected to help PT. Alif Duta Persada to implement marketing strategies and promo giving to their customers, based on their segmentation results. Segment 1 is the worst segment which has 0.187 CLV average. Segment 3 is the medium segment which has 0.144 CLV average. Segment 2 is the best segment which has 0.07 CLV average. Segment 1 has 140 members, segment 2 has 387 members, and segment 3 has 644 members. This research resulted web-based visualization that presents graphics of clustering results.*

***Keywords : Clustering, Fuzzy C-Means, LRFM 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 LRFM UNTUK SEGMENTASI PELANGGAN (STUDI KASUS : PT. ALIF DUTA PERSADA)**, 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, adik, dan seluruh keluarga yang senantiasa memberikan dukungan materi, semangat dan doa untuk penulis.
3. Pihak PT. Alif Duta Persada Probolinggo, 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 Achmad Holil Noor Ali selaku dosen wali yang telah memberikan dorongan, perwalian dan semangat sejak awal perkuliahan hingga pengerjaan 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. Rekan-rekan pengurus BEM FTIf, DPM ITS, dan organisasi lain yang telah menjadi tempat penulis dalam mengembangkan *softskill*.
9. Rekan-rekan pejuang wisuda 118 laboratorium RDIB yang memberikan bantuan dan semangat.
10. Teman-teman OSIRIS yang menjadi keluarga tempat berbagi susah dan senang selama penulis menempuh perkuliahan.
11. 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, Juni 2018

Penulis

## DAFTAR ISI

ABSTRAK .....	vi
ABSTRACT .....	viii
KATA PENGANTAR .....	x
DAFTAR ISI .....	xii
DAFTAR TABEL .....	xvi
DAFTAR GAMBAR .....	xviii
DAFTAR SCRIPT .....	xx
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Tugas Akhir .....	6
1.5 Manfaat Tugas Akhir .....	6
1.6 Relevansi Tugas Akhir.....	6
1.7 Sistematika Penulisan Tugas Akhir.....	7
BAB II TINJAUAN PUSTAKA.....	11
2.1 Studi Sebelumnya.....	11
2.2 Dasar Teori.....	13
2.2.1 PT. Alif Duta Persada.....	14
2.2.2 Customer Relationship Management .....	14
2.2.3 Segmentasi Pelanggan.....	16
2.2.4 <i>Clustering</i> .....	18
2.2.5 <i>Fuzzy C-Means (FCM)</i> .....	19
2.2.6 Model LRFM.....	21
2.2.7 Normalisasi <i>Min-Max</i> .....	22
2.2.8 Metode <i>Elbow</i> .....	23
2.2.9 Analytical Hierarchy Process .....	24
2.2.10 Customer Lifetime Value (CLV) .....	27
2.2.11 Uji Performa.....	28
BAB III METODOLOGI PENELITIAN.....	31

3.1	Metodologi .....	31
3.1.1	Identifikasi Masalah .....	32
3.1.2	Studi Literatur .....	33
3.1.3	Pengumpulan Data .....	33
3.1.4	Praproses Data.....	34
3.1.5	Pembobotan LRFM.....	36
3.1.6	Perhitungan CLV.....	36
3.1.7	Clustering Data.....	36
3.1.8	Pengujian Kualitas <i>Cluster</i> .....	37
3.1.9	Visualisasi <i>Cluster</i> .....	37
3.1.10	Analisis Segmentasi .....	37
3.1.11	Penyusunan Laporan Tugas Akhir .....	38
BAB IV PERANCANGAN .....		39
4.1	Pengumpulan data .....	39
4.2	Praproses data.....	40
4.3	Pembobotan LRFM.....	42
4.4	Penentuan jumlah <i>cluster</i> .....	42
4.5	Rancangan visualisasi .....	42
4.5.1	Use Case Diagram.....	43
4.5.2	<i>Use Case Description</i> .....	44
4.5.3	Desain Antarmuka.....	45
BAB V IMPLEMENTASI.....		49
5.1	Pra Proses Data.....	49
5.1.1	Data Cleaning .....	49
5.1.2	Pemodelan LRFM .....	50
5.1.3	Normalisasi.....	53
5.2	Pembobotan LRFM.....	54
5.3	Perhitungan CLV.....	55
5.4	Proses Clustering.....	55
5.4.1	Penentuan Jumlah K.....	56
5.4.2	Clustering dengan Fuzzy C-Means .....	57
5.5	Pengujian Kualitas <i>Clustering</i> .....	58
5.6	Pembuatan Visualisasi.....	59

5.6.1	Halaman Utama.....	59
5.6.2	Halaman Visualisasi Segmen.....	61
5.6.3	Halaman Tabel.....	71
BAB VI HASIL PEMBAHASAN.....		73
6.1	Pembobotan LRFM dengan AHP.....	73
6.2	Perhitungan CLV.....	76
6.3	Proses <i>Clustering</i> .....	77
6.3.1	Penentuan Nilai K.....	77
6.3.2	<i>Clustering Fuzzy C-Means</i> .....	78
6.4	Pengujian Kualitas <i>Clustering</i> .....	82
6.4.1	Validasi Derajat Keanggotaan.....	82
6.4.2	Pengujian SSE.....	83
6.5	Analisis Hasil Clustering.....	84
6.5.1	Pemeringkatan Segmen.....	84
6.5.2	Analisis Karakteristik Segmen 1.....	85
6.5.3	Analisis Karakteristik Segmen 2.....	86
6.5.4	Analisis Karakteristik Segmen 3.....	87
6.5.5	Perbandingan Karakteristik Segmen.....	88
6.6	Analisis Visualisasi Clustering.....	89
6.6.1	Grafik <i>Bar Chart</i> dan <i>Pie Chart</i> .....	89
6.6.2	Grafik <i>Scatter Plot 3D</i> .....	97
6.6.3	Grafik <i>Box Plot</i> .....	101
6.7	Kesimpulan Analisis.....	108
BAB VII KESIMPULAN DAN SARAN.....		124
7.1	Kesimpulan.....	125
7.2	Saran.....	126
DAFTAR PUSTAKA.....		127
BIODATA PENULIS.....		132
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 .....	11
Tabel 2.2 Penelitian Sebelumnya Judul Paper 2 .....	11
Tabel 2.3 Penelitian Sebelumnya Judul Paper 3 .....	12
Tabel 2.4 Penelitian Sebelumnya Judul Paper 4 .....	13
Tabel 2.5 Skala Preferensi Penilaian Pairwise Comparison...	25
Tabel 2.6 Nilai Random Indeks.....	27
Tabel 4.1 Data Mentah Transaksi Penjualan.....	39
Tabel 4.2 Use case melihat grafik bar chart dan pie chart.....	44
Tabel 5.1 Atribut Terpilih .....	49
Tabel 5.2 Baris Data Tidak Valid.....	50
Tabel 5.3 Sebagian Hasil Pemodelan LRFM.....	52
Tabel 5.4 Rentang Nilai Atribut.....	53
Tabel 6.1 Hasil Kuesioner Responden 1 .....	73
Tabel 6.2 Hasil Kuesioner Responden 2 .....	73
Tabel 6.3 Hasil Kuesioner Responden 3 .....	73
Tabel 6.4 Hasil Rata-Rata Penilaian Kuesioner .....	74
Tabel 6.5 Normal Comparisons Matrix.....	74
Tabel 6.6 Overall Score Bobot.....	75
Tabel 6.7 Sebagian Hasil Perhitungan CLV Pelanggan.....	76
Tabel 6.8 Selisih Penurunan Nilai Error .....	76
Tabel 6.9 Percobaan Terhadap Perubahan Nilai Parameter ...	79
Tabel 6.10 Sebagian Hasil Validasi Derajat Keanggotaan.....	83
Tabel 6.11 Peringkat Segmen Berdasarkan CLV.....	85
Tabel 6.12 Karakteristik Pelanggan pada Segmen 1 .....	85
Tabel 6.13 Karakteristik Pelanggan pada Segmen 2.....	87
Tabel 6.14 Karakteristik Pelanggan pada Segmen 3.....	88
Tabel 6.15 Perbandingan Karakteristik 3 Segmen .....	89
Tabel 6.16 Perbandingan Persebaran Wilayah 3 Segmen.....	91
Tabel 6.17 Perbandingan Karakteristik 3 Segmen (tanpa bobot) .....	115

Tabel 6.18 Perbandingan Karakteristik 4 Segmen (K=4) ....	117
Tabel 6.19 Perbandingan Karakteristik 5 Segmen (K=5) ....	119
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 LRFM.....	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 LRFM	D-1
Tabel E.1 Hasil Keseluruhan Perhitungan CLV dan Clustering .....	E-2
Tabel F.1 Keseluruhan Hasil Percobaan Perubahan Nilai Parameter.....	F-1
Tabel G.1 Keseluruhan Validasi Derajat Keanggotaan Pelanggan .....	G-2

## DAFTAR GAMBAR

Gambar 3.1 Alur Metodologi Penelitian .....	31
Gambar 4.1 Use Case Diagram .....	44
Gambar 4.2 Halaman Utama Dashboard.....	46
Gambar 4.3 Halaman Visualisasi Segmen .....	47
Gambar 4.4 Halaman Tabel Data.....	47
Gambar 5.1 Hasil Eksekusi Visualisasi Rata-Rata Variabel LRFM.....	63
Gambar 5.2 Hasil Eksekusi Visualisasi Tabel Peringkat Pelanggan .....	68
Gambar 5.3 Hasil Eksekusi Visualisasi Tabel Data .....	72
Gambar 6.1 Plot Penentuan Nilai K dengan Metode Elbow ..	78
Gambar 6.2 Hasil clustering Fuzzy C-Means untuk 3 cluster	80
Gambar 6.3 Hasil Pengujian SSE.....	80
Gambar 6.4 Grafik Pie Chart 3 Segmen.....	89
Gambar 6.5 Grafik Bar Chart 3 Segmen .....	90
Gambar 6.6 Grafik Stacked Bar Chart Persebaran 3 Segmen Pada Semua wilayah .....	91
Gambar 6.7 Grafik Bar Persebaran Wilayah cluster Pada Semua Segmen.....	91
Gambar 6.8 Grafik Pie Chart Persebaran Wilayah cluster Pada Semua Segmen .....	91
Gambar 6.9 Grafik Bar Chart Persebaran Wilayah Cluster Pada Segmen 1 .....	92
Gambar 6.10 Grafik Pie Chart Persebaran Wilayah Cluster Pada Segmen 1 .....	93
Gambar 6.11 Grafik Bar Chart Persebaran Wilayah Cluster Pada Segmen 2 .....	94
Gambar 6.12 Grafik Pie Chart ersebaran Wilayah Cluster Pada Segmen 2.....	94

Gambar 6.13 Grafik Bar Chart ersebaran Wilayah Cluster Pada Segmen 3.....	95
Gambar 6.14 Grafik Pie Chart ersebaran Wilayah Cluster Pada Segmen 3.....	95
Gambar 6.15 Grafik Scatter Plot 3D RFM.....	96
Gambar 6.16 Grafik Scatter Plot 3D LRF.....	97
Gambar 6.17 Grafik Scatter Plot 3D LRM.....	98
Gambar 6.18 Grafik Scatter Plot 3D LFM.....	99
Gambar 6.19 Grafik Box Plot CLV Semua Segmen.....	100
Gambar 6.20 Grafik Box Plot Length Semua Segmen .....	101
Gambar 6.21 Grafik Box Plot Recency Semua Segmen .....	103
Gambar 6.22 Grafik Box Plot Frequency Semua Segmen... ..	106
Gambar 6.23 Grafik Box Plot Monetary Semua Segmen ....	105
Gambar 6.24 Grafik Scatter Plot 2D LR .....	108
Gambar 6.25 Grafik Scatter Plot 2D LF .....	109
Gambar 6.26 Grafik Scatter Plot 2D LM .....	111
Gambar 6.27 Grafik Scatter Plot 2D RF .....	112
Gambar 6.28 Grafik Scatter Plot 2D RM.....	113
Gambar 6.29 Grafik Scatter Plot 2D FM .....	114
Gambar 6.30 Grafik Barchart 3 segmen (tanpa bobot) .....	116
Gambar 6.31 Grafik Bar chart Persebaran Wilayah pada segmen 1 (tanpa bobot) .....	116
Gambar 6.32 Grafik Barchart 4 segmen (K=4).....	118
Gambar 6.33 Grafik Bar chart Persebaran Wilayah pada segmen 3 (K=4).....	119
Gambar 6.34 Grafik Barchart 5 segmen (K=5).....	120
Gambar 6.35 Grafik Bar chart Persebaran Wilayah pada segmen 2 (K=5).....	121

## DAFTAR SCRIPT

Script 5.1 Query Ekstraksi Data .....	52
Script 5.2 Normalisasi Data .....	54
Script 5.3 Penentuan Jumlah K .....	57
Script 5.4 Proses Clustering .....	58
Script 5.5 Perhitungan SSE .....	59
Script 5.6 Visualisasi Bar Chart .....	60
Script 5.7 Visualisasi Pie Chart.....	61
Script 5.8 Visualisasi rata-rata variabel LRFM (1) .....	62
Script 5.9 Visualisasi rata-rata variabel LRFM (2) .....	63
Script 5.10 Visualisasi Pelanggan Teratas dan Terbawah.....	64
Script 5.11 Visualisasi Grafik Scatter Plot 3D RFM.....	65
Script 5.12 Visualisasi Grafik Scatter Plot 3D LRF.....	66
Script 5.13 Visualisasi Grafik Scatter Plot 3D LRM .....	66
Script 5.14 Visualisasi Grafik Scatter Plot 3D LFM.....	67
Script 5.15 Visualisasi Tabel Peringkat Pelanggan.....	67
Script 5.16 Visualisasi Grafik Box-Plot (1) .....	69
Script 5.17 Visualisasi Grafik Box-Plot (2) .....	70
Script 5.18 Visualisasi Tabel Data .....	71

*Halaman ini sengaja dikosongkan*



# **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**

Pada era digital saat ini data menjadi sumber daya yang sangat penting dalam berbagai sektor bisnis. Data merupakan sumber daya baru sekaligus menjadi tambang baru, dahulu yang menjadi kaya adalah yang menguasai tambang emas, batubara, minyak, maka pada era digital ini yang disebut sebagai tambang adalah tambang data [1]. Dalam sektor perdagangan, data transaksi pelanggan menjadi komponen yang sangat penting, untuk menghasilkan banyak informasi penting bagi perusahaan. Hal ini memaksa perusahaan-perusahaan yang bergerak di sektor perdagangan berlomba-lomba untuk memanfaatkan dan mengolah data transaksi pelanggan yang mereka miliki. Dalam hal ini perusahaan yang mampu memanfaatkan data mereka dengan baik maka akan mendapatkan banyak keuntungan dan perusahaan yang tidak memanfaatkan data dalam bisnisnya maka akan tertinggal. Dengan melakukan penggalian informasi dari data mereka tidak serta merta membuat perusahaan mendapatkan keuntungan secara mudah, karena banyak factor yang mempengaruhi pendapatan perusahaan, salah satunya yaitu tingkat konsumsi masyarakat. Badan Pusat Statistik menyebutkan masyarakat masih akan menahan konsumsinya pada kuartal pertama tahun 2018, mengikuti tren yang terjadi pada sepanjang tahun lalu.

Tren ini terlihat dari prediksi Indeks Tendensi Konsumsi (ITK) selama tiga bulan pertama tahun 2018, yang diperkirakan akan menyentuh angka 101,35 atau jauh lebih rendah dibandingkan kuartal ke IV 2017 yakni pada angka 107 [2].

PT. Alif Duta Persada, merupakan distributor PT. Unilever Indonesia, Tbk untuk wilayah Kabupaten dan Kota Probolinggo. Berbagai macam produk makanan dan produk rumah tangga dipasarkan diantaranya adalah adalah grup produk *Personal Care* dan *Home Care Food*. Seperti halnya rendahnya tingkat konsumsi masyarakat pada kuartal I 2018, berdampak pada pendapatan PT. Unilever Indonesia, Tbk yang mengalami peningkatan omset yang rendah, yaitu 2,9% dari hasil penjualan tahun sebelumnya [3].

Terkait dengan peningkatan omset yang rendah, PT. Alif Duta Persada telah melakukan beberapa kegiatan promosi untuk meningkatkan omset perusahaan, kegiatan promosi yang dilakukan antara lain, mengelompokkan pelanggan mereka kedalam lima kelompok berdasarkan omset mereka. Pemberian harga promosi dan bonus kepada pelanggan yang diberikan tiap beberapa waktu sekali. Permasalahan yang terjadi selama ini adalah pihak PT. Alif Duta Persada mengelompokkan pelanggannya secara manual setiap satu atau dua bulan sekali, dan pendistribusian bonus yang diberikan oleh PT. Unilever Indonesia, Tbk ditentukan oleh tiap-tiap sales yang mengelolanya sehingga tidak ada standar yang pasti dalam pemberian bonus tersebut, hanya berdasarkan kebijakan pada setiap sales, hal tersebut dapat menyebabkan pemberian promo yang tidak tepat sasaran dan tidak optimal.

Untuk menyelesaikan permasalahan tersebut solusi yang dapat dilakukan adalah dengan mengotomatisasi proses segmentasi pelanggan dan melakukan *clustering* berdasarkan beberapa faktor untuk menentukan siapa saja pelanggan yang pantas dan sesuai untuk menerima bonus sebagai bentuk CRM (*Customer Relationship Management*) perusahaan. Hal itu berguna untuk memastikan bahwa salah satu bentuk CRM tersebut tepat guna dan diharapkan efektif untuk meningkatkan kepuasan pelanggan dan meningkatkan pendapatan yang diperoleh dari pelanggan tersebut.

*Clustering* adalah pengelompokan objek data berdasarkan informasi yang hanya ditemukan pada data yang menggambarkan objek dan hubungannya. Tujuannya adalah bahwa objek dalam suatu kelompok serupa (terkait) satu sama lain dan berbeda (tidak terkait dengan) dari objek dalam kelompok lain [5]. Studi-studi tentang segmentasi pasar yang berbeda sudah pernah dilakukan sebelumnya, salah satu yang telah dilakukan adalah *clustering* pada kelompok pelancong yang homogen. Pengelompokan ini telah terbukti kegunaannya dalam menentukan kebijakan yang lebih spesifik dan efektif yang bertujuan untuk mempromosikan perubahan perilaku dan keberlanjutan sistem transportasi [6].

Pada dasarnya banyak metode yang bisa digunakan untuk melakukan segmentasi pelanggan salah satunya adalah penggalian data pada riwayat transaksi pembelian pelanggan pada periode tertentu. Kemudian dari data transaksi tersebut dapat dirumuskan beberapa variabel yang dapat ditentukan menjadi kriteria dalam melakukan segmentasi pelanggan. Variabel yang umum dijadikan tolak ukur adalah model RFM (*Recency, Frequency, dan Monetary*), model tersebut adalah metode *customer*

*value analysis* yang telah banyak digunakan untuk mensegmentasi pelanggan [7]. Model RFM tidak dapat membedakan mana pelanggan yang memiliki hubungan jangka panjang atau jangka pendek dengan perusahaan [8]. Atas dasar tersebut maka dikembangkan model RFM dengan menambahkan variabel baru, salah satunya adalah penambahan *Length* (L) sebagai pertimbangan untuk mengungkapkan loyalitas pelanggan sehingga terbentuk model LRFM yang merupakan ekstensi dari model RFM.

Algoritma *clustering* yang digunakan adalah *Fuzzy C-Means* (FCM). FCM merupakan salah satu teknik untuk menentukan *cluster* optimal pada suatu ruang vektor yang didasarkan pada bentuk normal *Euclidean* untuk jarak antar vektor. FCM 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. Selain itu, dengan algoritma ini bisa ditentukan jumlah cluster yang akan dibentuk. Dengan penentuan jumlah cluster diawal, bisa diatur keragaman nilai akhir sesuai dengan cluster-nya. Kelebihan algoritma ini adalah penempatan pusat *cluster* yang lebih tepat dibandingkan dengan metode *cluster* lain [9]. Caranya adalah dengan memperbaiki pusat cluster secara berulang, maka akan dapat dilihat bahwa pusat cluster akan bergerak menuju lokasi yang tepat. FCM juga memiliki tingkat akurasi yang tinggi dan waktu komputasi yang cepat [10]. Kemudian dalam suatu penelitian terdahulu yang membandingkan *Fuzzy C-Means* dan *K-Means* dalam studi kasus *Pengelompokan Data User Knowledge Modeling* menunjukkan bahwa FCM adalah metode yang lebih baik daripada *K-Means*

untuk melakukan clustering pada data *user knowledge modeling* dikarenakan nilai validasinya bernilai mendekati 1 [11].

Luaran dari penelitian ini adalah berupa segmen-segmen pelanggan yang terbentuk berdasarkan hasil dari kemiripan perilaku pelanggan pada setiap *cluster*. Dengan adanya segmentasi ini diharapkan dapat membantu PT. Alif Duta Persada untuk menerapkan strategi pemasaran dan pemberian promo pada pelanggan mereka, berdasarkan hasil segmentasi mereka.

## 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 LRFM?
2. Bagaimana akurasi hasil segmentasi menggunakan metode *clustering Fuzzy C-Means*?
3. Bagaimana visualisasi dapat mempermudah analisis segmentasi pelanggan?
4. Bagaimana analisis segmentasi dilakukan terhadap hasil *clustering*?

## 1.3 Batasan Masalah

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

1. Data yang akan diolah yaitu data transaksi harian penjualan produk PT. Alif Duta Persada pada rentang waktu 1 Desember 2017 hingga 31 Januari 2018.

2. Variabel yang digunakan sebagai parameter adalah interval waktu transaksi pembelian pertama (*Length*), jarak waktu pembelian terakhir (*Recency*), frekuensi pembelian (*Frequency*), dan jumlah biaya pembelian (*Monetary*) yang mengacu pada metode LRFM.
3. Perangkat lunak yang digunakan sebagai alat bantu *clustering* adalah R Studio
4. Visualisasi *cluster* 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. Mengetahui akurasi dari segmentasi yang telah dihasilkan oleh metode *Fuzzy C-Means*
3. Membuat visualisasi *clustering* sehingga mempermudah analisis segmentasi
4. Menghasilkan analisa segmentasi dari *clustering* yang telah dilakukan

#### **1.5 Manfaat Tugas Akhir**

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

#### **1.6 Relevansi Tugas Akhir**

Persaingan usaha yang setiap hari semakin meningkat menyebabkan semakin susah untuk meningkatkan omset perusahaan, bahkan dapat mengalami penurunan omset perusahaan. Untuk mengatasi masalah tersebut, perusahaan perlu menciptakan strategi-strategi bisnis yang tepat sasaran. Untuk mempermudah dalam menciptakan strategi dan kebijakan yang dibutuhkan,

maka diperlukan adanya sebuah sistem cerdas yang dapat melakukan segmentasi terhadap pelanggan mereka sesuai dengan kemiripan sifat atau perilakunya berdasarkan model LRFM (*Length, Recency, Frequency* dan *Monetary*). Dengan menggunakan metode *Fuzzy C-Means*, diberikan solusi berupa segmentasi pelanggan dengan model LRFM untuk mempermudah perusahaan dalam memberikan bonus kepada pelanggan mereka sesuai dengan segmentasinya.

Topik tugas akhir ini adalah sistem pendukung keputusan untuk membantu merumuskan kebijakan pemerintah dalam menjaga stabilitas pangan di Indonesia. Topik ini sangat relevan untuk menjadi tugas akhir Strata Satu (S1) karena merupakan bentuk implementasi mata kuliah pada bidang keilmuan seperti Sistem Cerdas, Penggalian Data dan Analitika Bisnis dan Manajemen Rantai Pasok dan Hubungan Pelanggan. Tugas akhir ini juga relevan dengan bidang keilmuan laboratorium RDIB (Rekayasa Data dan Intelegensia Bisnis).

## **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 LRFM, 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 LRFM, perhitungan CLV, *clustering* data, pengujian kualitas *clustering*, pembuatan visualisasi, analisa hasil segmentasi, dan penyusunan laporan. Semua tahapan akan dideskripsikan dan digambarkan dalam jadwal pengerjaan 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 LRFM, 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 LRFM, 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 LRFM, 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*

<b>Judul Penelitian 1</b>	<i>A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business [12]</i>
<b>Penulis; Tahun</b>	Der-Chiang Li, Wen-Li Dai, Wan-Ting Tseng; 2011
<b>Deskripsi Umum</b>	Penelitian ini menggunakan model LRFM, yang terdiri dari empat dimensi: relasi <i>length</i> (L), <i>recency</i> (R), <i>frequency</i> (F), dan <i>monetary</i> (M), untuk melakukan <i>clustering</i> pada pelanggan.
<b>Keterkaitan dengan Tugas Akhir</b>	Referensi penggunaan metode <i>cluster</i> , Model LRFM untuk segmentasi pelanggan, dan penjelasan CLV secara umum

*Tabel 2.2 Penelitian Sebelumnya Judul Paper 2*

<b>Judul Penelitian 2</b>	<i>Investigating Two Customer Lifetime Value Models from Segmentation Perspective [13]</i>
<b>Penulis; Tahun</b>	Abdulkadir Hiziroglu, Serkan Sengul; 2012

<b>Deskripsi Umum</b>	Makalah ini bertujuan untuk memberikan klasifikasi model terkini berdasarkan karakteristik dasarnya dan membuat perbandingan antara dua model representatif dari kelas yang berbeda dengan menggunakan <i>database</i> yang sama. Evaluasi dari perspektif segmentasi dilakukan dan hasilnya menunjukkan bahwa model yang mewakili kelas model perilaku pelanggan masa depan dan masa lalu ternyata lebih unggul daripada rekannya yang menggunakan database dan variabel yang sama.
<b>Keterkaitan dengan Tugas Akhir</b>	Referensi dalam perhitungan CLV menggunakan model RFM terhadap hasil <i>clustering</i> data untuk mengetahui keuntungan yang diperoleh dalam siklus hidup pelanggan.

Tabel 2.3 Penelitian Sebelumnya Judul Paper 3

<b>Judul Penelitian 3</b>	IMPLEMENTASI FUZZY C-MEANS DAN MODEL RFM UNTUK SEGMENTASI PELANGGAN (STUDI KASUS: PT. XYZ) [14]
<b>Penulis; Tahun</b>	Denny Bintang Saputra; 2017
<b>Deskripsi Umum</b>	Pengimplementasian metode <i>fuzzy c-means</i> terhadap data transaksi pelanggan untuk <i>clustering</i> dengan tujuan melakukan segmentasi pelanggan PT. XYZ. Permasalahan yang dialami PT. XYZ adalah tidak adanya strategi promosi spesifik yang bisa ditentukan untuk kelompok atau <i>cluster</i> tertentu karena tidak mengenali karakteristik pelanggannya. Segmentasi dilakukan dengan metode <i>clustering fuzzy c-means</i> , serta model RFM yang dapat menggambarkan perilaku pelanggan dalam melakukan transaksi. Lalu hasil <i>clustering</i> tersebut divisualisasikan untuk memudahkan analisis segmentasi. Hasilnya didapat 3 segmen pelanggan dengan karakteristik yang berbeda-

	beda, dan dengan hasil tersebut perusahaan dapat menerapkan strategi pemasaran yang tepat untuk setiap segmen sesuai dengan karakteristiknya.
<b>Keterkaitan dengan Tugas Akhir</b>	Sebagai referensi untuk melakukan segmentasi pelanggan menggunakan <i>clustering Fuzzy C-Means</i> sebagai panduan dalam menentukan metodologi pengerjaan tugas akhir.

Tabel 2.4 Penelitian Sebelumnya Judul Paper 4

<b>Judul Penelitian 4</b>	<i>Fuzzy c-means clustering with prior biological knowledge</i> [15]
<b>Penulis; Tahun</b>	Luis Tari, Chitta Baral, Seungchan Kim; 2008
<b>Deskripsi Umum</b>	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</i> (GO) sebagai <i>domain cluster</i> yang akan dibentuk. Oleh karena itu, metode <i>clustering</i> ini disebut sebagai <i>semi-supervised clustering</i> . Hasilnya, modifikasi metode <i>Fuzzy C-Means</i> diklaim mampu membentuk <i>cluster</i> yang konsisten dengan menentukan cluster awal dengan <i>prior knowledge</i> .
<b>Keterkaitan dengan Tugas Akhir</b>	Referensi penggunaan metode <i>Fuzzy C-Means</i> dimodifikasi untuk <i>clustering</i> .

## 2.2 Dasar Teori

Dasar teori yang memiliki keterkaitan dengan tugas akhir ini terdiri dari Profil Perusahaan, *Customer Relationship Management*, Segmentasi Pelanggan, *Clustering, Fuzzy C-Means*, Model LRFM, Normalisasi

*Min-Max, Metode Elbow, Analytical Hierarchical Process (AHP), Customer Lifetime Value (CLV), dan Uji Performa Clustering.*

### **2.2.1 PT. Alif Duta Persada**

PT. Alif Duta Persada merupakan perusahaan distributor dari PT. Unilever Indonesia, Tbk di wilayah Probolinggo yang mulai beroperasi sejak tahun 2012. PT. Alif Duta Persada sebagai distributor dari PT. Unilever Indonesia, Tbk bertanggung jawab untuk memasarkan produk-produk dari PT. Unilever Indonesia, Tbk (hanya khusus pada kategori *Personal Care* dan *Home Care Food*) untuk seluruh wilayah Kota dan Kabupaten Probolinggo.

PT. Alif Duta Persada memiliki divisi marketing yang bertugas untuk membantu proses pemasaran dan pendistribusian produk-produk *Personal Care* dan *Home Care Food* dari PT. Unilever Indonesia, Tbk. Dari kedua kategori tersebut dibagi menjadi beberapa *brand* produk yang dijual:

- a. *Personal Care*  
*AXE, Close Up, Dove, LUX, Pond's, Rexona, Sunsilk, Tresemme, Vaseline, Citra, Clear, Fair & Lovely, Lifebuoy.*
- b. *Homecare and Food*  
*Rinso, Molto, Sunlight, Wipol, Super Pell, Vixal, Cif, Blueband, Sariwangi, Bango.*

### **2.2.2 Customer Relationship Management**

*Customer Relationship Management (CRM)* adalah sebuah pendekatan yang komprehensif untuk, menciptakan, memelihara, dan

memperluas hubungan pelanggan [16]. Proses CRM menyediakan struktur untuk menciptakan hubungan yang sukses dengan pelanggan.

CRM adalah sarana untuk mengatasi persaingan yang semakin ketat, perubahan kondisi ekonomi dan ketergantungan promosi melalui penggunaan pengetahuan pelanggan yang intim. Pengetahuan diperoleh melalui pengembangan hubungan dan program pemasaran masa lalu [17]. 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. CRM telah digunakan untuk mengidentifikasi *Continuous Relationship Marketing*, *Customer Relationship Marketing* dan *Customer Relationship Management*. Setiap istilah mewakili proses yang sama [17].

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. Dalam penggunaan teknologi, sejumlah aplikasi teknologi dapat diidentifikasi yang digunakan pada pengembangan strategi CRM. Tiga komponen utama komponen sistem CRM adalah [18]:

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 pelanggan adalah proses membagi basis pelanggan menjadi kelompok-kelompok *homogen* yang berbeda, untuk mengembangkan strategi pemasaran yang berbeda sesuai dengan karakteristiknya [19]. 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 [20]:

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

*Clustering* adalah studi yang ditujukan untuk mengungkapkan kelompok atau *cluster* data yang serupa [21]. *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 [21]:

- a. Metode Hirarki, yaitu metode yang membentuk *cluster* yang membagi partisi secara berulang-ulang dari atas ke bawah atau sebaliknya. Hasil dari metode hirarki berupa dendogram yang mewakili kelompok objek dan tingkat kesamaan di mana terdapat perubahan pengelompokkan.

- b. 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.

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

#### 2.2.5 *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 meng*cluster* 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.

Algoritma Fuzzy C-Means adalah sebagai berikut [22]:

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

$$Q = \sum_{i=1}^K \mu_{ik}, \quad j = 1,2,3, \dots, m$$

Hitung:

$$\mu_{ik} = \frac{\mu_{ik}}{Q_j}$$

- d. Menghitung Pusat *Cluster* ke- $k$   
Hitung pusat *cluster* ke- $k$ :  $V_{kj}$  dengan  $k = 1,2,3,\dots,c$ ;  $j = 1,2,3,\dots,m$ .

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

- e. Menghitung Fungsi Objectif  
Hitung fungsi objektif pada iteasi ke-t, Pt:

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

- f. Menghitung Perubahan Matriks

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

$$i=1,2,3,\dots,n; k=1,2,3,\dots,c$$

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

### 2.2.6 Model LRFM

RFM merupakan singakatan dari *Recency*, *frequency* dan *Monetary*. *Recency*, *Frequency*, dan *Monetary* (RFM) adalah metode yang sering digunakan untuk menentukan apakah pelanggan tersebut berharga (*valuable customer*) dengan cara melihat pembelian pelanggan yang paling akhir dilakukan (*recency*), pelanggan yang selalu membeli (*frequency*) dan pelanggan yang mengeluarkan uang yang paling tinggi (*monetary*) [23]. Metodologi ini bermanfaat dalam segmentasi pelanggan dengan membagi pelanggan ke dalam beberapa kelompok untuk diberikan layanan personal di masa depan dan untuk

mengidentifikasi pelanggan pelanggan yang berpotensi merespon promosi. Penggalan data berbasis konsep RFM sudah diterapkan ke berbagai area seperti keamanan komputer dan industri elektronik.

Berkembangnya penelitian terkait dengan segmentasi pelanggan, membuat beberapa peneliti memunculkan ide penambahan variable pada RFM. Salah satu modifikasi model RFM adalah LRFM, dengan menambahkan variable *Length* atau interval waktu dari pembelian pertama hingga pembelian terakhir. Pengembangan ini dilakukan karena beberapa hal, salah satunya adalah fakta bahwa model RFM tidak bisa mensegmen pelanggan mana yang memiliki hubungan jangka panjang dan jangka pendek dengan perusahaan. Padahal loyalitas pelanggan bergantung pada durasi hubungan antara pelanggan dan perusahaan, maka dari itu perlu ditambahkan sebuah variabel untuk mengukur tingkat loyalitas yang telah dibangun oleh pelanggan. Itulah mengapa beberapa literatur lebih memilih menggunakan model LRFM untuk melakukan segmentasi terhadap pelanggan suatu perusahaan.

### 2.2.7 Normalisasi *Min-Max*

Normalisasi Min-Max merupakan salah satu metode normalisasi dengan melakukan perubahan linier terhadap data asli [12]. Misalkan  $Min_a$  dan  $Max_a$  adalah nilai minimum dan maksimum dari atribut A. Normalisasi Min-Max memetakan nilai  $v$  dari A ke  $v'$  dalam rentang  $[new\_min(a), new\_max(a)]$ .

Normalisasi Min-Max dihitung dengan normalisasi berikut [24]:

$$v' = \frac{[v - \min(a)] \times [\text{new\_max}(a) - \text{new\_min}(a)]}{[\max(a) - \min(a)]} + \text{new\_min}(a)$$

**Keterangan:**

$v'$	= nilai yang baru hasil normalisasi
$v$	= nilai awal sebelum dinormalisasi
$\min(a)$	= nilai terkecil atribut
$\max(a)$	= nilai terbesar atribut
$\text{new\_min}(a)$	= rentang minimal
$\text{new\_max}(a)$	= rentang maksimal

### 2.2.8 Metode *Elbow*

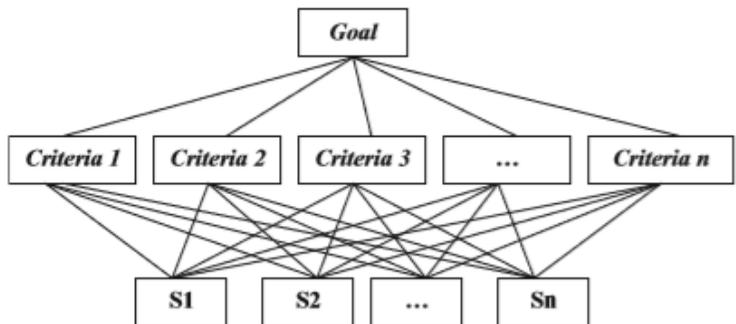
Metode *Elbow* adalah suatu metode untuk menghasilkan informasi dalam menentukan jumlah *cluster* terbaik dengan cara melihat prosentase hasil perbandingan antara jumlah *cluster* yang akan membentuk sudut siku pada suatu titik [25]. Hasil prosentase dari setiap perhitungan ditunjukkan dalam grafik dan 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 [26]:

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.

### 2.2.9 Analytical Hierarchy Process

*Analytic Hierarchy Process* (AHP), yang awalnya digunakan oleh Saaty pada tahun 1970an, adalah salah satu metode pengambilan keputusan multi kriteria yang banyak digunakan, dan telah berhasil diterapkan pada banyak masalah pembuatan keputusan praktis [27]. Dalam AHP, masalah yang rumit dapat didekomposisi menjadi beberapa hierarki sesuai dengan atribut atau kriteria yang terkait. Struktur Hirarki pada AHP dengan tiga tingkat ditunjukkan pada Gambar 1 [27].



*gambar 2 Struktur Hirarki pada AHP dengan tiga tingkat*

Langkah-langkah yang harus dilakukan untuk memperoleh bobot dengan metode AHP adalah sebagai berikut [28]:

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*

<b>Tingkat preferensi</b>	<b>Definisi</b>
1	Kedua kriteria memiliki pengaruh yang sama
3	Penilaian dan pengalaman sedikit memihak pada salah satu kriteria tertentu dibanding kriteria pasangannya
5	Penilaian dan pengalaman memihak pada salah satu kriteria tertentu dibanding kriteria pasangannya
7	Salah satu kriteria lebih diprioritaskan dan relatif lebih penting dibandingkan kriteria pasangannya
9	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. *Consistency check*, yaitu instrument yang menentukan konsistensi dan menunjukkan tingkat kepercayaan prioritas yang diperoleh dari perbandingan. Terdapat beberapa tahapan dalam uji rasio konsistensi, yaitu sebagai berikut:
  - 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 formula berikut:

$$CI = \frac{x - n}{n - 1}$$

**Keterangan:**

CI = indeks konsistensi

$x$  = hasil dari langkah 4b

$n$  = banyaknya kriteria

- d. Menghitung tingkat ketidak konsistensian yang dapat diterima dengan persamaan berikut:

$$\frac{CI}{RI}$$

**Keterangan:**

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 CI/RI lebih kecil dari 0.1.

**2.2.10 Customer Lifetime Value (CLV)**

*Customer lifetime value* (CLV) adalah pengukuran kuantitatif dari arus kas bersih perusahaan yang dihasilkan oleh pelanggannya sepanjang hubungan mereka dengan perusahaan berlangsung [13]. Pengukuran ini sangat penting dan banyak digunakan oleh berbagai perusahaan seperti lembaga keuangan, toko ritel, perusahaan telekomunikasi dan lain-lain untuk menemukan perbedaan antara para pelanggan mereka, sehingga mereka dapat menyesuaikan layanan yang paling sesuai untuk para pelanggan mereka. Perhitungan CLV yang benar dapat memudahkan perusahaan mengklasifikasikan pelanggannya, sehingga strategi pemasaran yang berbeda dapat dikembangkan untuk setiap *cluster*.

Dalam penelitian tugas akhir ini, penggunaan CLV mengacu pada indeks nilai CLV. Indeks CLV ditentukan melalui bobot variabel LRFM

yang sebelumnya telah diperoleh dan nilai normalisasi masing-masing variabel LRFM Adapun formula untuk menghitung nilai CLV adalah sebagai berikut [18]:

$$CLV = NL \times WL + NR \times WR + NF \times WF + NM \times WM$$

**Keterangan:**

CLV	= Nilai CLV yang dicari
NL	= Nilai Normalisasi Variabel <i>length</i>
WL	= Nilai Bobot Variabel <i>length</i>
NR	= Nilai Normalisasi Variabel <i>retency</i>
WR	= Nilai Bobot Variabel <i>retency</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>

### 2.2.11 Uji Performa

Uji performa *clustering* bertujuan untuk mencari kelompok *cluster* yang cocok dengan jumlah *cluster* tanpa informasi kelas. Pada uji performa *clustering* mempunyai dua tipe uji performa *clustering*. Tipe pertama adalah uji performa eksternal, tipe ini menggunakan informasi terdahulu mengenai data [29]. Tipe ini digunakan untuk mengetahui algoritma terbaik untuk proses *clustering* sesuai dengan data set tertentu. Sementara, tipe data yang kedua adalah uji performa internal yang hanya menggunakan data yang ada. Uji performa ini digunakan untuk mengevaluasi seberapa bagus struktur *cluster*. Pada kenyataannya, data eksternal untuk uji performa eksternal tidak ada, sehingga seringkali uji performa internal yang digunakan.

Uji performa internal diukur berdasarkan dua kriteria. Pertama adalah kepadatan yang mengukur seberapa dekat objek yang berhubungan dalam satu *cluster*. Kedua adalah pemisahan yang mengukur seberapa bedanya antara satu *cluster* dengan *cluster* lainnya [30]. Uji performa internal dapat menggunakan SSE (*Sum Square Error*).

*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 [25]:

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

**Keterangan:**

- $K$  = 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- $i$

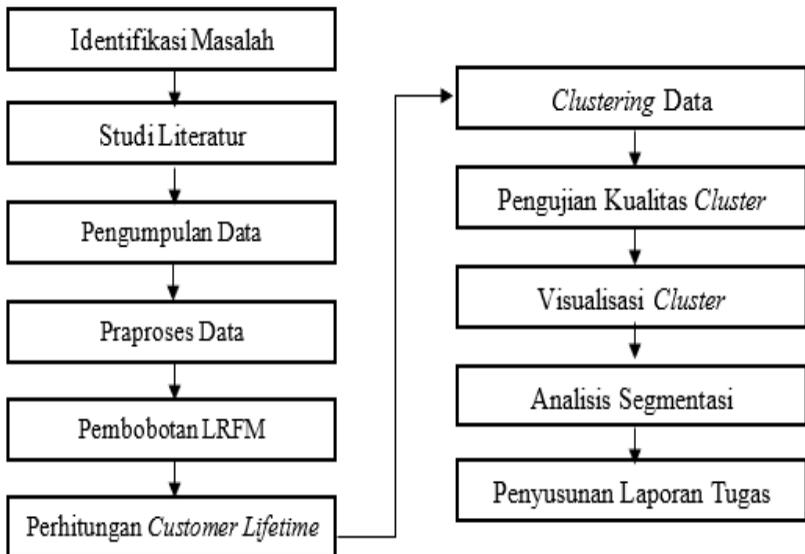
*Halaman ini sengaja dikosongkan*

## 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 *length*, *recency*, *frequency*, dan *monetary* menghasilkan data LRFM. Nilai pada data LRFM selanjutnya akan dinormalisasi untuk menyetarakan skala nilai dan menghasilkan data LRFM ternormalisasi. Kemudian, dilakukan pembobotan pada setiap variabel LRFM 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 LRFM ternormalisasi dikalikan dengan bobot masing-masing dan menghasilkan data LRFM bobot untuk masukan pada tahap *clustering*. Setelah itu, ditentukan terlebih dahulu nilai K dengan metode Elbow dan dilakukan proses *clustering* terhadap data LRFM 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. Alif Duta Persada, dimana permasalahan yang diangkat adalah segmentasi pelanggan. 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 LRFM, dan segmentasi pelanggan yang berhubungan dengan tugas akhir. Studi literatur dilakukan berdasarkan wawancara dengan narasumber, 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 tanggal 1 Desember 2017 sampai 31 Januari 2018, yaitu sebanyak 929 transaksi penjualan

Data transaksi penjualan yang diperoleh memiliki atribut seperti tanggal transaksi, nama pelanggan, nomor invoice, jenis produk yang dibeli, kuantitas produk, dan biaya total. 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 2 bulan dari tanggal 1 Desember 2017 hingga 31 Januari 2018.

### 3.1.4 Praproses Data

Setelah pengumpulan data, tahap selanjutnya adalah melakukan praproses 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 LRFM dan normalisasi.

#### 1.1.1.1. *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.

#### 1.1.1.2. Transformasi LRFM

Transformasi RFM merupakan subproses untuk membentuk variabel *Length*, *Recency*, *Frequency*, dan *Monetary* dari data transaksi yang masih mentah. Adapun pemodelan variabel-variabel tersebut antara lain:

- a. *Length* : variabel ini diperoleh dari selisih waktu antara transaksi yang pertama kali dilakukan pelanggan dengan transaksi terakhirnya. Untuk menghitung nilai dari variabel ini digunakan atribut tanggal transaksi.

- b. *Recency* : variabel *recency* diperoleh dari selisih waktu transaksi yang terakhir kali dilakukan oleh masing-masing pelanggan. Atribut yang berkaitan dengan variabel ini adalah tanggal transaksi.
- c. *Frequency* : variabel *frequency* diperoleh dari jumlah transaksi yang dilakukan oleh masing-masing pelanggan dengan mengagregasikan semua data berdasarkan pelanggan
- d. *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.

#### 1.1.1.3. 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 *length*, *recency* dan *frequency*.

Setelah tiga subproses tersebut dilakukan, akan menghasilkan data transaksi yang teragregasi berdasarkan tiap pelanggan, dengan empat atribut, yaitu *Length*, *Recency*, *Frequency*, dan *Monetary* yang siap untuk *dicluster*

### 3.1.5 Pembobotan LRFM

Pada tahapan ini dilakukan pembobotan terhadap variabel LRFM dengan metode AHP. Penilaian bobot diperoleh memberika kuisioner kepada pimpinan PT. Alif Duta Persada, manajer dan satu karyawan pada bagian marketing PT. Alif Duta Persada, yaitu sebanyak tiga responden. Dari perhitungan AHP tersebut, diperoleh nilai bobot untuk variabel *Length*, *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 LRFM dari masing-masing pelanggan dan bobot masing-masing variabel. Kemudian dilakukan perkalian nilai variabel LRFM dengan bobot masing-masing. Indeks CLV didapatkan dari penjumlahan semua nilai variabel LRFM. 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 pelanggan. Sebelum *clustering* dilakukan, 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 keempat variable (*length*,

*recency*, *frequency*, dan *monetary*). Setelah proses *clustering* selesai, luarannya adalah *cluster* yang berisikan data pelanggan yang memiliki perilaku tertentu. Perilaku yang dimaksud disini adalah terkait keempat variabel yang jadi acuan, yaitu *length*, *recency*, *frequency*, dan *monetary*.

### **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 bertujuan untuk memudahkan dalam pembacaan hasil *clustering* untuk menganalisis segmentasi pelanggan pada setiap *clusternya*.

### **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*. Dalam tahap ini juga dilakukan penjelasan mengenai hasil visualisasi *dashboard* yang telah dibuat, sehingga dapat digunakan oleh PT. Alif Duta Persada untuk menentukan strategi pemasaran dan pemberian promo terhadap para pelanggan mereka.

### **3.1.11 Penyusunan Laporan Tugas Akhir**

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

## BAB IV PERANCANGAN

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

### 4.1 Pengumpulan data

Pengumpulan data dilakukan secara langsung pada departemen penjualan, PT. Alif Duta Persada. Data diperoleh dari salah satu admin yang bertanggung jawab dalam menginputkan transaksi penjualan. Data yang dibutuhkan adalah riwayat transaksi penjualan dari tanggal 1 Desember 2017 sampai 31 Januari 2018 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.

*Tabel 4.1 Data Mentah Transaksi Penjualan*

<b>Variabel</b>	<b>Tipe Data</b>	<b>Keterangan</b>
DistID	Number	Id Distributor
DistName	Text	Nama Distributor
Salesman	Number	Id Salesman
SalesmanNama	Text	Nama Salesman
INVNumber	Number	Nomor Invoice
INVDate	Date	Tanggal Invoice
Outlet	Number	Id Outlet
Outlet Name	Text	Nama Outlet
Outlet Address	Text	Alamat Outlet

SKUCode	Number	Code Unik Produk
ProductName	Text	Nama Produk
Case	Number	Jumlah Penjualan dalam satuan Case (Kardus)
Dozen	Number	Jumlah Penjualan dalam satuan Lusin
Pieces	Number	Jumlah Penjualan dalam satuan Pieces/Eceran
Weight	Number	Berat Produk
Total Quantites (PCS)	Number	Jumlah Penjualan Produk dalam satuan Pieces/Pack
PriceCase	Number	Harga dalam satuan Case
GSV	Number	NA
Discount	Number	Nilai Potongan yang diberikan
Tax	Number	Nilai Pajak yang di Bebaskan
Net	Number	Nilai Total Uang yang diterima

Data transaksi yang diperoleh berjumlah 78262 baris data, terhitung mulai 1 Desember 2017 hingga 31 Januari 2018.

## 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 LRFM, yaitu *length*, *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 LRFM untuk mendapatkan nilai variabel LRFM yang dibutuhkan berdasarkan atribut yang tersedia.
    - a. Variabel *length* diperoleh dari selisih waktu transaksi pertama kali yang dilakukan dengan transaksi yang terakhir kali dilakukan oleh masing-masing pelanggan. Atribut yang berkaitan dengan variabel ini adalah tanggal transaksi.
    - b. Variabel *recency* diperoleh dari selisih waktu transaksi yang terakhir kali dilakukan oleh masing-masing pelanggan per tanggal 1 Februari 2018. Atribut yang berkaitan dengan variabel ini adalah tanggal transaksi.
    - c. Variabel *frequency* diperoleh dari jumlah transaksi yang dilakukan oleh masing-masing pelanggan dengan mengagregasikan semua data berdasarkan setiap pelanggan.
    - d. 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 *length*, *recency* dan *frequency*. Metode normalisasi yang digunakan adalah metode *min-max*.

### 4.3 Pembobotan LRFM

Tahap selanjutnya dalam analisis LRFM adalah menentukan bobot untuk setiap variabel LRFM. 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 PT. Alif Duta Persada. Responden yang mengisi kuesioner ada tiga orang, yaitu direktur PT. Alif Duta Persada, *Sales Manager*, dan *Admin*. Kuesioner yang diberikan berisikan data responden, petunjuk pengisian kuesioner, penjelasan variabel LRFM, 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.

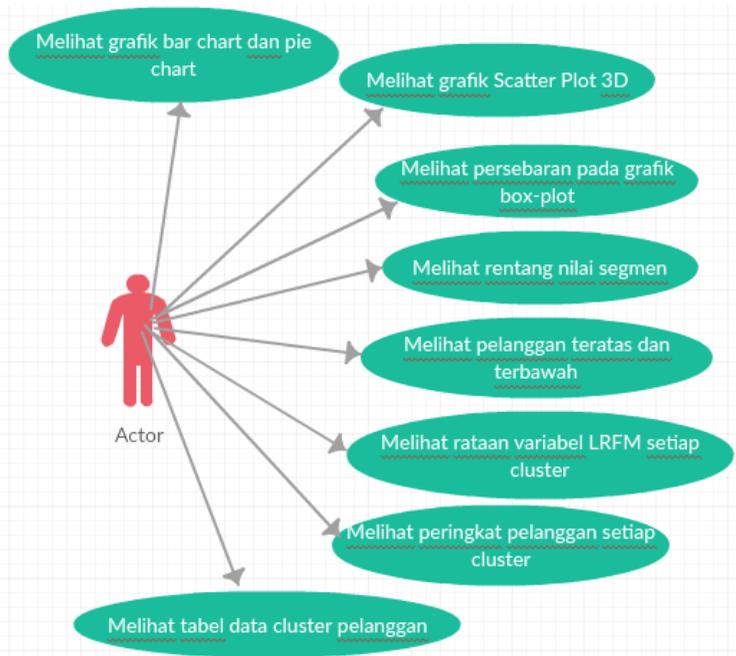
### 4.5 Rancangan visualisasi

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 4.1, 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 rata-rata variabel LRFM 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 4.2 Use case melihat grafik *bar chart* dan *pie chart*

<b>UC-1 Melihat grafik <i>bar chart</i> dan <i>pie chart</i></b>	
<b>Purpose</b>	Mengetahui jumlah dan prosentase pelanggan pada masing-masing <i>cluster</i> dalam grafik <i>bar chart</i> dan <i>pie chart</i>
<b>Overview</b>	Dimulai ketika pengguna mengakses aplikasi
<b>Actors</b>	Pengguna aplikasi

<b>Pre Condition</b>	Data hasil <i>clustering</i> telah dimasukkan pada aplikasi	
<b>Post Condition</b>	Grafik <i>bar chart</i> dan <i>pie chart</i> ditampilkan	
<b>Typical Course Event</b>	<b>Actor</b>	<b>System</b>
	1. Mengklik ikon  <i>slide bar</i> 3. Mengeklik menu “Halaman Utama”	2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan grafik “Jumlah Anggota Segmen” dan “Prosentase Anggota Segmen”, menampilkan grafik jumlah dan prosentase persebaran anggota segmen berdasarkan area.
<b>Alternate Flow Event</b>	-	
<b>Exceptional Flow Event</b>	Aplikasi tertutup karena <i>error</i>	

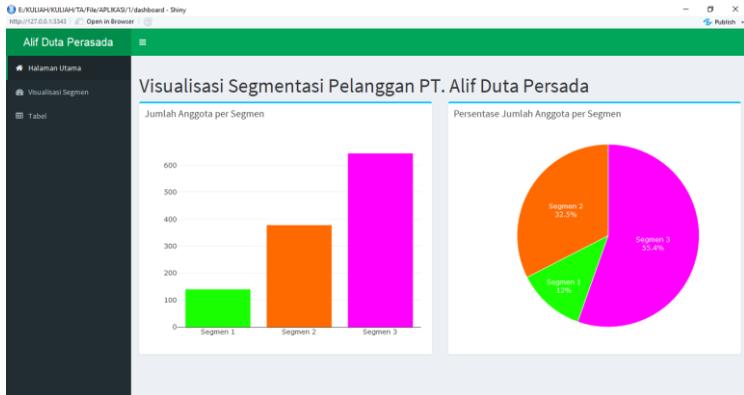
### 4.5.3 Desain Antarmuka

Desain antarmukan 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*

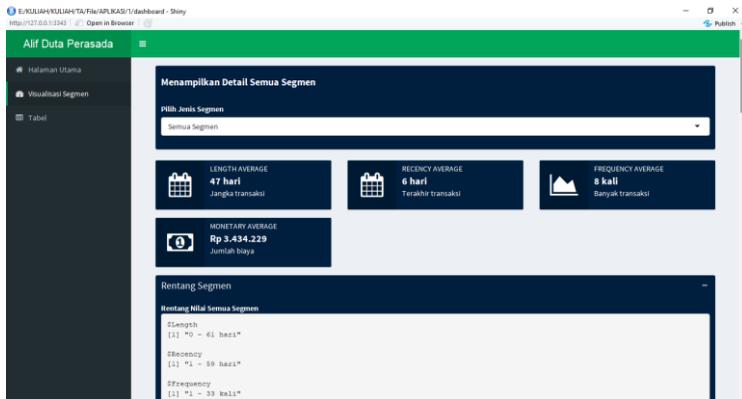
*chart.* Rancangan halaman utama dashboard disajikan pada Gambar 4.2.



*Gambar 4.2 Halaman Utama Dashboard*

*b. Halaman visualisasi segmen*

Halaman visualisasi segmen menampilkan informasi rata-rata variabel LRFM, rentang nilai variabel LRFM, pelanggan teratas dan terbawah, tabel peringkat pelanggan, dan grafik *box plot* variabel LRFM 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.

Customer	Length	Recency	Frequency	Monetary	CLV	Segmen
2 PUTRI	52	2	5	1140490	0.10572242	2
27 WANDA,TK	40	0	7	923455	0.0958747	3
371	58	6	9	2258208	0.12010993	3
3F	49	7	7	978818	0.09712931	3
08	56	6	9	564030	0.12071816	3
9,TK	50	6	0	207538	0.10414821	3
AA JAYA	49	5	6	763952	0.09726723	3
AA HAKHUR	56	1	0	1152093	0.12480958	3
AAN, BU	56	3	15	4377325	0.16189554	1
ABADI	49	5	7	425920	0.10025992	3

Gambar 4.4 Halaman Tabel Data

*Halaman ini sengaja dikosongkan*

## BAB V IMPLEMENTASI

Pada bab ini menjelaskan proses implementasi berdasarkan dari rancangan yang dibuat yaitu pra proses data, pembobotan LRFM, 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 LRFM 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 LRFM. Dari 13 atribut yang ada, atribut yang dipilih adalah INVDate, INVNumber, dan net. Penjelasan mengenai atribut yang dipilih tersaji pada Tabel 5.1.

*Tabel 5.1 Atribut Terpilih*

Atribut	Keterangan
INVDate	Atribut date digunakan untuk memperoleh variabel <i>length</i> dan <i>recency</i> , dengan mencari selisih hari.
INVNumber	Atribut date digunakan untuk memperoleh variabel <i>frequency</i> , dengan menghitung jumlah transaksi per pelanggan.

Net	Atribut net 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. Pencarian baris yang bernilai 0 dan negatif dilakukan dengan fitur *sort* pada *excel*. Setelah diurutkan, didapatkan 3642 baris data yang bernilai negatif dan bernilai 0. 3642 baris data yang tidak valid tersebut kemudian dihapus dan tersisa 75608 baris data. Adapun data yang dihapus ditampilkan ada Tabel 5.2.

*Tabel 5.2 Baris Data Tidak Valid*

<b>INVDate</b>	<b>INVnumber</b>	<b>Net</b>
1/4/2018	18007700003-04	-3128400
12/4/2017	17007700032-04	-2906218
1/20/2018	18000300028-04	-892800
12/29/2017	17007700044-04	-69120f0
12/4/2017	17000100217-02	-606720
1/31/2018	18007700009-04	-508800
12/5/2017	17011000188-02	-442801
12/16/2017	17000700678-02	-412087
1/17/2018	18000100018-04	-398400
12/20/2017	17000300470-02	-3128400

### 5.1.2 Pemodelan LRFM

Pemodelan LRFM dilakukan untuk membentuk nilai variabel *length*, *recency*, *frequency*, dan *monetary* sebagai masukan proses *clustering*. Nilai LRFM didapat dengan mengekstrak

nilai atribut yang sebelumnya dipilih. Pemodelan LRFM dilakukan dengan *query* pada basis data *MySQL*.

Nilai variabel *length* atau L adalah selisih hari sejak transaksi terakhir kali dilakukan oleh setiap pelanggan dengan hari saat transaksi pertama kali dilakukan. Atribut yang dibutuhkan untuk memperoleh nilai L adalah INVNumber dan INVDate. Nilai L didapatkan dari tanggal transaksi yang paling terakhir menggunakan fungsi MAX, dan transaksi yang paling awal menggunakan fungsi MIN berdasarkan tiap pelanggan dengan fungsi GROUP BY, sedangkan untuk mencari selisihnya menggunakan fungsi DATEDIFF.

Nilai variabel *recency* atau R adalah selisih hari sejak transaksi terakhir kali dilakukan oleh setiap pelanggan dengan hari pengerjaan, yaitu tanggal 1 Februari 2018. Atribut yang dibutuhkan untuk memperoleh nilai R adalah INVNumber dan INVDate. 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 INVNumber dan INVDate. Nilai F diperoleh dengan menghitung banyaknya date dengan fungsi COUNT dan GROUP BY berdasarkan INVNumber.

Nilai variabel *monetary* atau M adalah jumlah pemasukan yang diterima perusahaan dari tiap pelanggan. Atribut yang dibutuhkan untuk memperoleh nilai M adalah INVNumber dan Net. Nilai M didapatkan dengan menjumlah atribut Net dengan fungsi COUNT dan fungsi GROUP BY berdasarkan INVNumber.

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

```
SELECT Outlet_Name as Customer,
datediff(max(INVdate),min(INVdate)) as L,
datediff('2018-02-01',max(INVdate)) as R,
COUNT(DISTINCT (INVdate)) as F,
sum(net) as M
from `transaksiclean`
GROUP BY Outlet_Name
```

*Script 5.1 Query Ekstraksi Data*

Setelah *query* dieksekusi, diperoleh data sejumlah 1163 pelanggan yang melakukan transaksi sejak 1 Desember 2017 hingga 31 Januari 2018. Sebagian hasil pemodelan LRFM dapat dilihat pada Tabel 5.3 dan lebih lengkapnya dapat dilihat pada Lampiran C.

*Tabel 5.3 Sebagian Hasil Pemodelan LRFM*

<b>Customer</b>	<b>L</b>	<b>R</b>	<b>F</b>	<b>M</b>
SURYA JAYA (KS)	57	5	7	178,526,088
TAUFIK	42	12	3	96,180,647
BAMBANG KUSNADI (REJEKI)	21	20	3	90,626,503
BASMALAH, TK	61	1	16	81,878,971
PUTRA CAHAYA	51	1	5	63,187,930
SUMBER AGUNG	49	13	6	52,388,211
PP NURUL JADID PUTRI	42	8	5	49,294,618
SIDOGIRI, KOP.	55	1	7	45,866,756
AIR MANCUR / POJOK	56	2	10	43,152,365
HARTONO	56	3	13	37,488,319

### 5.1.3 Normalisasi

Normalisasi data dilakukan untuk menyetarakan skala variabel *monetary* yang terlalu jauh dengan variabel *length*, *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 L, 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>Length</i>	61	0
<i>Recency</i>	59	1
<i>Frequency</i>	33	1
<i>Monetary</i>	178526088	4620

Perhitungan normalisasi *Min-Max* dilakukan berdasarkan rumus yang telah dijelaskan pada Bab II. Setiap nilai dikurangi 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.

```

> adp <- read.csv("E:/KULIAH/KULIAH/TA/File/OLAH D
ATA/1.LRFM/Clean/QueryClean.csv")
> Lnorm <- (adp$L-min(adp$L))/(max(adp$L)-min(adp$
L))
> Rnorm <- (adp$R-min(adp$R))/(max(adp$R)-min(adp$
R))
> Fnorm <- (adp$F-min(adp$F))/(max(adp$F)-min(adp$
F))
> Mnorm <- (adp$M-min(adp$M))/(max(adp$M)-min(adp$
M))
> hasil <- data.frame(adp,Lnorm,Rnorm,Fnorm,Mnorm)
> write.csv(hasil,"E:/KULIAH/KULIAH/TA/File/OLAH D
ATA/1.LRFM/Clean/normalisasiLRFM.csv")

```

*Script 5.2 Normalisasi Data*

Nilai variabel R yang dihasilkan memiliki makna berbeda dengan variabel L, 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 mengurangi nilai R terhadap 1 pada kolom baru bernama R\_Balik. Hasil lengkap nilai LRFM yang ternormalisasi dapat dilihat pada Lampiran D.

## 5.2 Pembobotan LRFM

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

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 LRFM-nya saja. Data yang dibutuhkan adalah data LRFM tiap pelanggan dan bobot untuk setiap variabel LRFM. Setiap nilai LRFM 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 LRFM 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 LRFM 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.

```

> adp <- read.csv("E:/KULIAH/KULIAH/TA/File/OLAH
DATA/1.LRFM/Clean/normalisasiLRFM.csv")

> library(e1071)

> data <- data.frame(adp$Lbot, adp$Rbot, adp$Fb
ot, adp$Mbot)

> k.max <- 10

> elbow <- sapply(2:k.max, function(k)
{cmeans(data, k, 100,verbose = TRUE, method ="c
means", 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 LRFM 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.

```

> adp <- read.csv("E:/KULIAH/KULIAH/TA/File/OLA
H DATA/1.LRFM/Clean/normalisasiLRFM.csv")

library(e1071)

> data <- data.frame(adp$Lbot, adp$Rbot, adp$Fb
ot, adp$Mbot)

> clustering <- cmeans(data, 3, 100, method = "
cmeans", verbose = TRUE, rate.par = 0.5, weight
= 0.5)

> output <- data.frame(adp$Customer, clustering
$cluster, clustering$membership)

> write.csv(output, "E:/KULIAH/KULIAH/TA/File/O
LAH DATA/1.LRFM/Clean/HasilClustering.csv")

```

*Script 5.4 Proses Clustering*

## 5.5 Pengujian Kualitas *Clustering*

Setelah *clustering* dilakukan, dihasilkan *cluster* atau segmen yang membagi semua pelanggan ke dalamnya. 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 LRFM 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.

```

data <- data.frame(adp$Lbot, adp$Rbot, adp$Fbot
, adp$Mbot)

cl<-cmeans(input, 3, 100, 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 menggunakan 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 diintegrasikan 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$clustering.cluster == "1"),
           sum(data$clustering.cluster == "2"),
           sum(data$clustering.cluster == "3"))
#Bar Chart
output$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$clustering.cluster == "1"),
           sum(data$clustering.cluster == "2"),
           sum(data$clustering.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 LRFM, rentang nilai variabel LRFM, pelanggan teratas dan terbawah, tabel peringkat pelanggan, dan grafik *box plot* variabel LRFM setiap segmen yang dipilih. Masing-masing tampilan akan dijelaskan lebih detail pada poin berikut.

### 5.6.2.1 Tampilan rata-rata variabel LRFM

Rata-rata variabel LRFM ditampilkan dalam bentuk kotak-kotak informasi terpisah. Data yang diperlukan untuk menghitung rata-rata variabel LRFM adalah nilai LRFM masing-masing pelanggan yang sudah dikalikan dengan bobot variabel. Proses visualisasi rata-rata variabel LRFM dapat dilihat pada Script 5.8 dan 5.9.

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

#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) })
```

*Script 5.8 Visualisasia rata-rata variabel LRFM (1)*

```
#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 5.9 Visualisasia rata-rata variabel LRFM (2)*

Hasil eksekusi Script 5.8 dan 5.9 ditampilkan pada Gambar 5.1.



*Gambar 5.1 Hasil Eksekusi Visualisasi Rata-Rata Variabel LRFM*

### 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 LRFM 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 LRFM dapat dilihat pada Script 5.10.

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

Script 5.10 Visualisasi Pelanggan Teratas dan Terbawah

### 5.6.2.3 Tampilan Grafik Scatter Plot 3D

Grafik *scatter plot 3D* ditampilkan dengan bentuk kubus 3 dimensi yang mewakili tiga variabel. Sumbu X, sumbu Y, dan sumbu Z. Dengan 4 variabel LRFM, menjadi 4 kemungkinan grafik, yaitu *scatter plot 3D RFM*, *scatter plot 3D LRF*, *scatter*

*plot 3D LRM*, dan *scatter plot 3D LFM*. masing-masing mewakili variabel *length*, *recency*, *frequency*, dan *monetary*. Setiap titik dalam kubus menggambarkan setiap pelanggan yang dipetakan berdasarkan nilai variabel 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 LRFM dikali bobot masing-masing pelanggan dan cluster pelanggan berada. Proses visualisasi grafik *scatter plot 3D* dapat dilihat pada Script 5.11, 5.12, 5.13, dan 5.14

```
output$scatterRFM <- renderRglwidget ({
  dplot <- pilihdata()
  open3d()
  plot3d(x=dplot$R,
        y=dplot$F,
        z=dplot$M,
        col = c("#19ff00", "#ff6a00", "#ff00ff")[a
s.numeric(dplot$clustering.cluster)],
        type = 's', size = 1,
        xlab = "recency",
        ylab = "frequency",
        zlab = "monetary",
        colkey = list(length = 0.5, width = 0.5,
                      cex.clab = 0.75))

  scene2 <- scene3d()
  rgl.close()
  save <- options(rgl.inShiny = TRUE)
  on.exit(options(save))
  rglwidget(scene2) })
```

*Script 5.11 Visualisasi Grafik Scatter Plot 3D RFM*

```

output$scatterLRF <- renderRglwidget ({
  dplot <- pilihdata()
  open3d()
  plot3d(x=dplot$L,
         y=dplot$R,
         z=dplot$F,
         col = c("#19ff00", "#ff6a00", "#ff00ff")[a
s.numeric(dplot$clustering.cluster)],
         type = 's', size = 1,
         xlab = "length",
         ylab = "recency",
         zlab = "frequency",
         colkey = list(length = 0.5, width = 0.5,
                       cex.clab = 0.75))

```

*Script 5.12 Visualisasi Grafik Scatter Plot 3D LRF*

```

output$scatterLRM <- renderRglwidget ({
  dplot <- pilihdata()
  open3d()
  plot3d(x=dplot$L,
         y=dplot$R,
         z=dplot$M,
         col = c("#19ff00", "#ff6a00", "#ff00ff")[a
s.numeric(dplot$clustering.cluster)],
         type = 's', size = 1,
         xlab = "length",
         ylab = "recency",
         zlab = "monetary",
         colkey = list(length = 0.5, width = 0.5,
                       cex.clab = 0.75))

```

*Script 5.13 Visualisasi Grafik Scatter Plot 3D LRM*

```

output$scatterLFM <- renderRglwidget ({
  dplot <- pilihdata()
  open3d()
  plot3d(x=dplot$L,
        y=dplot$F,
        z=dplot$M,
        col = c("#19ff00", "#ff6a00", "#ff00ff")[a
s.numeric(dplot$clustering.cluster)],
        type = 's', size = 1,
        xlab = "length",
        ylab = "frequency",
        zlab = "monetary",
        colkey = list(length = 0.5, width = 0.5,
                      cex.clab = 0.75))

```

*Script 5.14 Visualisasi Grafik Scatter Plot 3D LFM*

#### 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 5.15.

```

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

```

*Script 5.15 Visualisasi Tabel Peringkat Pelanggan*

Hasil eksekusi Script 5.15 ditampilkan pada Gambar 5.2.

Peringkat Pelanggan

Show  entries Search:

Customer	Length	Recency	Frequency	Monetary	CLV
SURYA JAYA (KS)	57	5	7	178526088	0.7148900
BAROKAH	61	1	33	14082097	0.4057185
ARIFIN	58	1	31	13028739	0.3643939
SUMBER REJEKI	56	2	31	11071782	0.3573270
HARAPAN JAYA	57	2	29	23730577	0.3362027
BASMALAH, TK	61	1	16	81878971	0.3100850
ABD HADI	58	1	27	6567071	0.3009879
CANDRA	57	2	25	5363438	0.2695072

Customer  Length  Recency  Frequency  Monetary  CLV

Showing 1 to 8 of 1,162 entries Previous **1** 2 3 4 5 ... 146 Next

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 LRFM dikali bobot tiap pelanggan dan hasil *cluster*. Proses visualisasi grafik *box plot* dapat dilihat pada Script 5.16 dan 5.17.

```

#Data box plot
y1 <- subset(data,data$clustering.cluster==1)
y2 <- subset(data,data$clustering.cluster==2)
y3 <- subset(data,data$clustering.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 5.16 Visualisasi Grafik Box-Plot (1)*

```

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")) %>%

```

*Script 5.17 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 5.18.

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

*Script 5.18 Visualisasi Tabel Data*

Hasil eksekusi Script 5.18 ditampilkan pada Gambar 5.3.

Show  entries Search:

Customer	Length	Recency	Frequency	Monetary	CLV	Segmen
2 PUTRI	52	2	5	1140490	0.10572242	2
27 WANDA, TK	49	8	7	922455	0.09558747	3
371	56	6	9	2258208	0.12080693	3
3F	49	7	7	976618	0.09712031	3
88	56	6	9	594030	0.12071816	3
9, TK	50	6	8	207538	0.10414821	3
AA JAYA	49	5	6	763592	0.09726733	3
AA MAKMUR	56	1	8	1152983	0.12489858	3
AAN , BU	56	3	15	4377325	0.16189554	1
ABADI	49	5	7	425820	0.10025692	3

Customer  Length  Recency  Frequency  Monetary  CLV  Segmen

Showing 1 to 10 of 1,162 entries Previous      ...  Next

*Gambar 5.3 Hasil Eksekusi Visualisasi Tabel Data*

## BAB VI HASIL PEMBAHASAN

Pada bab ini akan menjelaskan hasil dan pembahasan dari proses pembobotan LRFM, 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 LRFM dengan AHP

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

*Tabel 6.1 Hasil Kuesioner Responden 1*

<i>Kriteria</i>	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>	<i>Length</i>
<i>Recency</i>	1	0.17	0.11	1
<i>Frequency</i>	6	1	0.25	7
<i>Monetary</i>	9	4	1	8
<i>Length</i>	1	0.14	0.13	1

*Tabel 6.2 Hasil Kuesioner Responden 2*

<i>Kriteria</i>	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>	<i>Length</i>
<i>Recency</i>	1	0.14	0.11	0.33
<i>Frequency</i>	7	1	0.20	5
<i>Monetary</i>	9	5	1	7
<i>Length</i>	3	0.20	0.14	1

*Tabel 6.3 Hasil Kuesioner Responden 3*

<i>Kriteria</i>	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>	<i>Length</i>
<i>Recency</i>	1	0.17	0.11	0.50
<i>Frequency</i>	6	1	0.33	7
<i>Monetary</i>	9	3	1	7
<i>Length</i>	2	0.14	0.14	1

Data kuesioner kemudian diolah dengan cara merata-rata semua pengisian responden terhadap perbandingan empat kriteria LRFM. Hasil rata-rata ini dijumlahkan setiap kolomnya. Hasil rata-rata dapat dilihat pada Tabel 6.4.

*Tabel 6.4 Hasil Rata-Rata Penilaian Kuesioner*

<b>Kriteria</b>	<b>Recency</b>	<b>Frequency</b>	<b>Monetary</b>	<b>Length</b>
<b>Recency</b>	1	0.16	0.11	0.61
<b>Frequency</b>	6.33	1	0.25	6.33
<b>Monetary</b>	9	4	1	7.33
<b>Lenth</b>	2	0.16	0.14	1
<b>Jumlah</b>	18.33	5.32	1.498	15.278

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 6.5. dari tabel ini didapatkan nilai bobot untuk setiap kriteria.

*Tabel 6.5 Normal Comparisons Matrix*

<b>Kriteria</b>	<b>Recency</b>	<b>Frequency</b>	<b>Monetary</b>	<b>Length</b>	<b>Bobot</b>
<b>Recency</b>	1	0.16	0.11	0.61	0.205
<b>Frequency</b>	6.33	1	0.25	6.33	1.212
<b>Monetary</b>	9	4	1	7.33	2.702
<b>Length</b>	2	0.16	0.14	1	0.3

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 6.6.

Tabel 6.6 Overall Score Bobot

<b>Kriteria</b>	<b>Recency</b>	<b>Frequency</b>	<b>Monetary</b>	<b>Length</b>	<b>Overall Score</b>
<b>Recency</b>	0.055	0.030	0.074	0.040	0.050
<b>Frequency</b>	0.345	0.188	0.167	0.415	0.279
<b>Monetary</b>	0.491	0.752	0.668	0.480	0.598
<b>Length</b>	0.109	0.030	0.091	0.065	0.074

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

$$Length = \frac{0.205}{0.074} = 4.05$$

$$Recency = \frac{1.212}{0.05} = 4.14$$

$$Frequency = \frac{2.702}{0.279} = 4.35$$

$$Monetary = \frac{0.3}{0.598} = 4.52$$

$$x = \frac{4.05 + 4.14 + 4.35 + 4.52}{4} = 4.27$$

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

$$CI = \frac{4.27 - 4}{4 - 1} = 0.089$$

Nilai CI yang diperoleh adalah 0.089, berdasarkan hukum uji konsistensi, karena nilai CI tidak sama dengan nol, maka dibutuhkan uji ketidakkonsistensian. Mengacu pada 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  $n = 4$ , maka nilai RI adalah 0.9. Sehingga nilai  $CI/RI$  yang diperoleh adalah :

$$\frac{0.089}{0.9} = 0.0985$$

Karena nilai  $CI/RI$  adalah 0.0985 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 0.074, 0.05, 0.279 dan 0.598 berturut-turut untuk *length*, *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 6.7 dan hasil lengkapnya pada Lampiran E.

Tabel 6.7 Sebagian Hasil Perhitungan CLV Pelanggan

<i>No</i>	<i>Customer</i>	<i>L x Bobot</i>	<i>R x Bobot</i>	<i>F x Bobot</i>	<i>M x Bobot</i>	<i>CLV</i>
1	2 PUTRI	0.06223	0.049138	0.034875	0.003805	0.105722
2	27 WANDA, TK	0.058639	0.043966	0.052313	0.003075	0.095587
3	371	0.067016	0.04569	0.06975	0.007549	0.120807
4	3F	0.058639	0.044828	0.052313	0.003256	0.09712
5	88	0.067016	0.04569	0.06975	0.001974	0.120718
6	9, TK	0.059836	0.04569	0.061031	0.00068	0.104148

7	AA JAYA	0.058639	0.046552	0.043594	0.002542	0.097267
8	AA MAKMUR	0.067016	0.05	0.061031	0.003847	0.124899
9	AAN , BU	0.067016	0.048276	0.122063	0.014647	0.161896
10	ABADI	0.058639	0.046552	0.052313	0.001411	0.100257

### 6.3 Proses *Clustering*

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

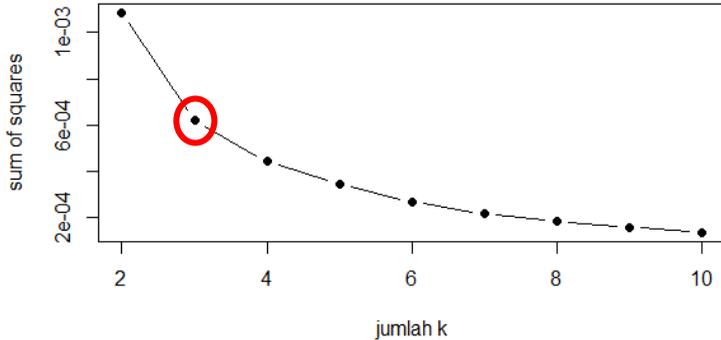
#### 6.3.1 Penentuan Nilai K

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. Nilai K ditentukan dengan perbandingan penurunan nilai SSE yang terbesar. Dapat dilihat pada table 6.8 bahwa nilai rasio terbesar adalah 2,567. Yang terletak pada nilai K=3, sehingga nilai 3 dipilih sebagai nilai K untuk proses *clustering*.

Tabel 6.8 Selisih Penurunan Nilai Error

<b>K</b>	<b>SSE</b>	<b>Selisih Error</b>	<b>Rasio</b>
2	0.0010839864	0	0
3	0.0006224164	0.0004615700	<b>2.56788</b>
4	0.0004426689	0.0001797475	1.857475
5	0.0003458991	0.0000967698	1.25205
6	0.0002686100	0.0000772891	1.532264

7	0.0002181689	0.0000504411	1.483754
8	0.0001841733	0.0000339956	1.336867
9	0.0001587440	0.0000254293	1.192587
10	0.0001374212	0.0000213228	<i>Undefined</i>



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  $L=0,3-0,6$ ;  $W=0,5$ ;  $K=3$  dapat dilihat pada Tabel 6.9 dan hasil lebih lengkapnya dapat dilihat pada Lampiran F.

Tabel 6.9 Percobaan Terhadap Perubahan Nilai Parameter

<b>Max Iter</b>	<b>Learning Rate</b>	<b>Weights</b>	<b>SSE Training</b>	<b>SSE Testing</b>
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

Percobaan dilakukan dengan mengganti nilai *Learning rate* dan *Weights*, dengan menggunakan *range* nilai 0,1 sampai 1 untuk masing-masing parameter, sehingga dilakukan seratus kali percobaan. Dari tabel tersebut dapat diketahui nilai *SSE testing* yang paling minimal atau terkecil, berada saat nilai *learning rate* 0,5 dan *weight* 0,5. Nilai *SSE* terkecil terbentuk ketika parameter *max iter* mencapai nilai 20. Sehingga parameter yang digunakan untuk *clustering* 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 6.9.

### Fuzzy c-means clustering with 3 clusters

Cluster centers:

	adp.Lbot	adp.Rbot	adp.Fbot	adp.Mbot
1	0.06704010	0.04866019	0.12961159	0.030094139
2	0.04380278	0.04282909	0.02625646	0.004385931
3	0.06266787	0.04734945	0.06067577	0.007675126

Memberships:

	1	2	3
[1,]	0.027390319	0.5846274283	0.3879822532
[2,]	0.015206462	0.1145951762	0.8701983615
[3,]	0.023756239	0.0397807316	0.9364630291
[4,]	0.014491387	0.1086591469	0.8768494660
[5,]	0.028652184	0.0513717569	0.9199760595
[6,]	0.010115392	0.0382601584	0.9516244500
[7,]	0.024461454	0.3743679575	0.6011705885
[8,]	0.007327716	0.0219602424	0.9707120421
[9,]	0.903262914	0.0271124934	0.0696245924
[10,]	0.015877680	0.1183896603	0.8657326600
[11,]	0.877677417	0.0342102362	0.0881123472
[12,]	0.968463193	0.0100819071	0.0214549003
[13,]	0.636142584	0.1466739386	0.2171834771

Closest hard clustering:

[1]	2	3	3	3	3	3	3	1	3	1	1	1	3	3	3	3	3	3	3	3	3	2	2	3	3	2	2	2	3	3	3	3	3	1	2			
[40]	1	2	1	1	1	2	3	3	2	2	1	2	2	1	3	3	3	3	3	3	3	3	2	3	3	3	2	3	3	2	3	2	3	2	1	2	3	2
[79]	2	3	3	2	2	3	2	3	3	2	3	3	3	1	3	3	3	3	2	2	1	2	3	3	3	2	3	1	1	3	2	3	3	3	2	2		
[118]	3	3	3	2	2	2	3	3	3	3	3	3	3	3	2	2	3	3	3	3	1	3	2	1	2	2	2	3	2	2	3	3	3	2	3	3		
[157]	3	2	3	3	3	1	3	3	3	1	3	3	2	2	2	3	1	2	3	2	1	3	3	3	2	3	2	2	3	2	1	3	3	3	2	3	1	
[196]	2	3	3	1	3	2	3	2	3	2	1	2	3	2	3	3	2	2	3	2	2	1	2	2	3	3	2	2	2	2	2	3	2	3	3	3	3	

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  $L = 0.23$   $R = 0.082$ ,  $F = 0.35$ ,  $M = 0.062$ , 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 :

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

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 :

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

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

$$\begin{aligned} u_{11} &= \frac{[0.203635331^2]^{-\frac{1}{0.5-1}}}{\sum_{k=1}^3 [\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{0.5-1}}} \\ u_{11} &= 0,065556 \end{aligned}$$

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

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

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

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

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

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

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 6.10 dan lebih lengkapnya dapat dilihat pada Lampiran G. Dari keseluruhan hasil *clustering*, 1162 objek telah valid atau berada pada cluster dengan nilai keanggotaan terbesar.

Tabel 6.10 Sebagian Hasil Validasi Derajat Keanggotaan

Customer	X1	X2	X3	Max	Cluster	Validasi
2 PUTRI	0.0273	0.5865	0.3862	0.5865	2	Valid
27 WANDA, TK	0.0153	0.1155	0.8693	0.8693	3	Valid
371	0.0236	0.0397	0.9367	0.9367	3	Valid
3F	0.0145	0.1095	0.8759	0.8759	3	Valid
88	0.0285	0.0513	0.9202	0.9202	3	Valid
9, TK	0.0101	0.0384	0.9515	0.9515	3	Valid
AA JAYA	0.0245	0.3763	0.5992	0.5992	3	Valid
AA MAKMU R	0.0073	0.0220	0.9707	0.9707	3	Valid
AAN , BU	0.9030	0.0272	0.0698	0.9030	1	Valid
ABADI	0.0159	0.1193	0.8648	0.8648	3	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 0.0006224164. Hasil pengujian SSE dapat dilihat pada Gambar 6.3.

## 6.5 Analisis Hasil Clustering

Proses *clustering* menghasilkan tiga segmen dengan 1162 pelanggan yang menjadi anggotanya. Masing-masing segmen memiliki karakteristik yang berbeda. Karakteristik setiap segmen dipengaruhi oleh variable *length*, *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: 1, Error: 0.0007598042
Iteration: 2, Error: 0.0007322963
Iteration: 3, Error: 0.0006959211
Iteration: 4, Error: 0.0006590157
Iteration: 5, Error: 0.0006362332
Iteration: 6, Error: 0.0006267122
Iteration: 7, Error: 0.0006236278
Iteration: 8, Error: 0.0006227463
Iteration: 9, Error: 0.0006225061
Iteration: 10, Error: 0.0006224413
Iteration: 11, Error: 0.0006224237
Iteration: 12, Error: 0.0006224188
Iteration: 13, Error: 0.0006224173
Iteration: 14, Error: 0.0006224168
Iteration: 15, Error: 0.0006224166
Iteration: 16, Error: 0.0006224165
Iteration: 17, Error: 0.0006224165
Iteration: 18, Error: 0.0006224164
Iteration: 19, Error: 0.0006224164
Iteration: 20 converged, Error: 0.0006224164
>
> print(cl$withinerror)
[1] 0.0006224164
```

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 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.

<i>Segmen</i>	<i>CLV</i>	<i>Peringkat Segmen</i>
1	0.187	1
2	0.07	3
3	0.114	2

*Tabel 6.11 Peringkat Segmen Berdasarkan CLV*

### **6.5.2 Analisis Karakteristik Segmen 1**

Segmen 1 memiliki anggota 140 pelanggan. Rata-rata nilai CLV segmen 1 adalah 0.187 atau yang paling tinggi dari dua segmen lainnya, sehingga segmen 1 merupakan segmen terbaik. Rentang jumlah uang yang dikeluarkan oleh pelanggan antara Rp. 647.690 hingga Rp. 178.526.088, dengan rata-rata Rp. 12.602.196. Hal ini menunjukkan bahwa pelanggan mengeluarkan uang yang relatif besar pada setiap transaksi.

Rentang jumlah transaksi yang dilakukan pelanggan antara 3 hingga 33 kali, dengan rata-rata 16 transaksi per pelanggan. Hal ini menunjukkan bahwa pelanggan cukup sering melakukan transaksi pembelian. Rentang waktu transaksi terakhir kali dilakukan adalah antara 1 hingga 20 hari. Dengan transaksi rata-rata 3 hari yang lalu. Selisih waktu transaksi terakhir dengan transaksi paling pertama kali yang dilakukan oleh pelanggan adalah 21 hingga 61 hari, dengan rata-rata selisih transaksi 56 hari. Karakteristik segmen 1 lebih lengkap dapat dilihat pada Tabel 6.12.

*Tabel 6.12 Karakteristik Pelanggan pada Segmen 1*

<b>Segmen</b>	1
<b>Peringkat Segmen</b>	1

<b>Jumlah Anggota</b>		140
<b>Karakteristik</b>	<b>Length</b>	21 – 61 hari (rata-rata : 56 hari)
	<b>Recency</b>	1 - 20 hari (rata-rata : 3 hari)
	<b>Frequency</b>	3 - 33 kali (rata-rata : 16 kali)
	<b>Monetary</b>	Rp 647.690 - Rp 178.526.088 (rata-rata : Rp. 12.602.196,-)

### 6.5.3 Analisis Karakteristik Segmen 2

Segmen 2 memiliki anggota 378 pelanggan. Rata-rata nilai CLV segmen 2 adalah 0.07 atau yang paling rendah dari dua segmen lainnya, sehingga segmen 2 merupakan segmen terburuk. Rentang jumlah uang yang dikeluarkan oleh pelanggan antara Rp. 4.620 hingga Rp. 35.207.012, dengan rata-rata Rp. 991.610. 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 4 transaksi per pelanggan. Hal ini menunjukkan bahwa pelanggan jarang melakukan transaksi pembelian. Rentang waktu transaksi terakhir kali dilakukan adalah antara 1 hingga 59 hari. Dengan transaksi rata-rata 9 hari yang lalu. Selisih waktu transaksi terakhir dengan transaksi paling pertama kali yang dilakukan oleh pelanggan adalah 0 hingga 57 hari, dengan rata-rata selisih transaksi 35 hari. Karakteristik segmen 2 lebih lengkap dapat dilihat pada Tabel 6.13.

Tabel 6.13 Karakteristik Pelanggan pada Segmen 2

<b>Segmen</b>		2
<b>Peringkat Segmen</b>		3
<b>Jumlah Anggota</b>		378
<b>Karakteristik</b>	<b>Lenth</b>	0 - 57 hari (rata-rata : 35 hari)
	<b>Recency</b>	1 - 59 hari (rata-rata : 9 hari)
	<b>Frequency</b>	1 - 6 kali (rata-rata : 4 kali)
	<b>Monetary</b>	Rp 4.620 - Rp 35.207.012 (rata-rata : Rp. 991.610)

#### 6.5.4 Analisis Karakteristik Segmen 3

Segmen 3 memiliki anggota 644 pelanggan. Rata-rata nilai CLV segmen 3 adalah 0.114 atau tertinggi kedua, sehingga segmen 3 dikategorikan sebagai segmen menengah. CLV yang sedang menggambarkan bahwa pelanggan pada segmen 3 cukup loyal terhadap perusahaan.

Rentang jumlah uang yang dikeluarkan oleh pelanggan antara Rp. 110.596 hingga Rp. 52.388.211, dengan rata-rata Rp. 2.874.903. Hal ini menunjukkan bahwa pelanggan mengeluarkan uang yang relatif tinggi pada setiap transaksi.

Rentang jumlah transaksi yang dilakukan pelanggan antara 5 hingga 12 kali, dengan rata-rata 8 transaksi per pelanggan. Hal ini menunjukkan bahwa pelanggan cukup sering melakukan transaksi pada perusahaan. Rentang waktu transaksi terakhir kali dilakukan adalah antara 1 hingga 20 hari. Dengan transaksi rata-rata 4 hari yang lalu. Selisih waktu transaksi terakhir dengan transaksi paling pertama kali yang dilakukan oleh pelanggan adalah 37 hingga 61 hari, dengan rata-rata selisih

transaksi 53 hari. Karakteristik segmen 3 lebih lengkap dapat dilihat pada Tabel 6.14.

Tabel 6.14 Karakteristik Pelanggan pada Segmen 3

<b>Segmen</b>		3
<b>Peringkat Segmen</b>		2
<b>Jumlah Anggota</b>		644
<b>Karakteristik</b>	<b>Length</b>	37 - 61 hari (rata-rata : 53 hari)
	<b>Recency</b>	1 - 20 hari (rata-rata : 4 hari)
	<b>Frequency</b>	5 - 12 kali (rata-rata : 8 kali)
	<b>Monetary</b>	Rp 110.596 - Rp 52.388.211 (rata-rata : Rp. 2.874.903)

### 6.5.5 Perbandingan Karakteristik Segmen

Perbandingan karakteristik ketiga segmen dapat dilihat pada Tabel 6.15. 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 tinggi, segmen 2 paling rendah, dan segmen 3 menengah.

Tabel 6.15 Perbandingan Karakteristik 3 Segmen

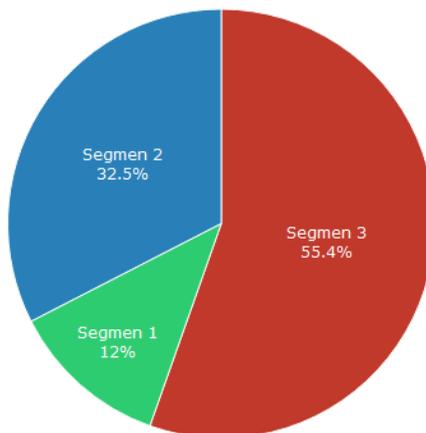
	Segmen 1	Segmen 2	Segmen 3
<b>Peringkat</b>	1	3	2
<b>Jumlah Anggota</b>	140	378	644
<b>Length</b>	21 – 61 hari	0 - 57 hari	37 - 61 hari
<b>Recency</b>	1 - 20 hari	1 - 59 hari	1 - 20 hari
<b>Frequency</b>	3 - 33 kali	1 - 6 kali	5 - 12 kali
<b>Monetary</b>	Rp 647.690 – Rp 178.526.088	Rp 4.620 – Rp 35.207.012	Rp 110.596 – Rp 52.388.211

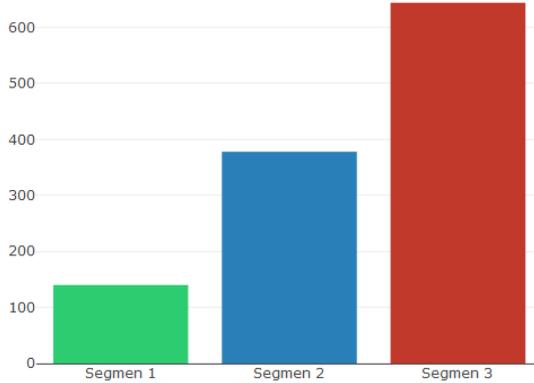
## 6.6 Analisis Visualisasi Clustering

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

### 6.6.1 Grafik *Bar Chart* dan *Pie Chart*

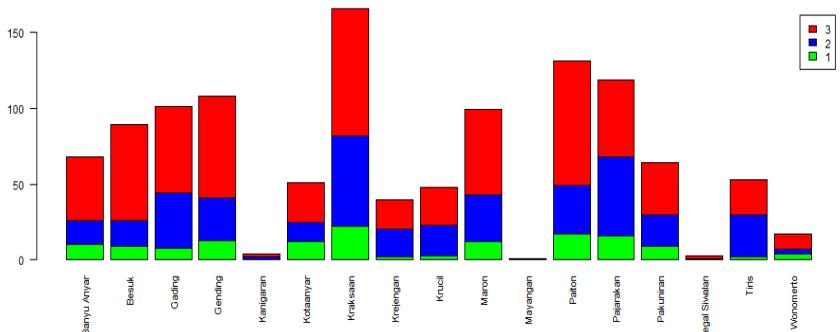
Grafik *bar chart* dan *pie chart* ditampilkan untuk mengetahui perbandingan jumlah anggota setiap segmen.

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 140 pelanggan atau sebanyak 12%. Segmen 2 memiliki anggota terbanyak, dengan 378 pelanggan atau 32.5%. Segmen 3 memiliki anggota paling sedikit, dengan 644 pelanggan atau 55.4%.



Gambar 6.6 Grafik Stacked Bar Chart Persebaran 3 Segmen Pada Semua wilayah

Dari gambar 6.6 diketahui persebaran wilayah dari ketiga segmen, dapat diketahui segmen mana saja yang terlihat dominan per wilayah, sekaligus dapat membandingkan persebaran segmen secara langsung. Untuk gambaran yang lebih jelas dijelaskan dalam tabel 6.16.

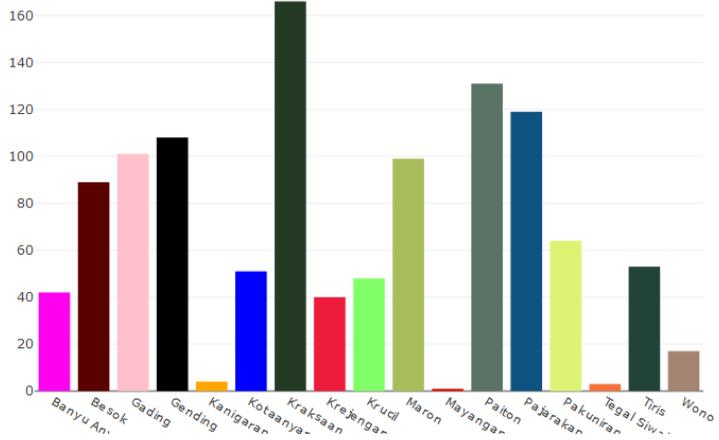
*Tabel 6.16 Perbandingan Persebaran Wilayah 3 Segmen*

<b>Wilayah</b>	<b>Segmen 1</b>	<b>Segmen 2</b>	<b>Segmen 3</b>
Banyu Anyar	10	16	42
Besuk	9	17	63
Gading	8	36	57
Gending	13	28	67
Kanigaran	0	2	2
Kotaanyar	12	13	26
Kraksaan	22	60	84
Krejengan	2	18	20
Krucil	3	20	25
Maron	12	31	56
Mayangan	1	0	0
Paiton	17	32	82
Pajarakan	16	52	51
Pakuniran	9	21	34
Tegal Siwalan	0	1	2
Tiris	2	28	23
Wonomerto	4	3	10

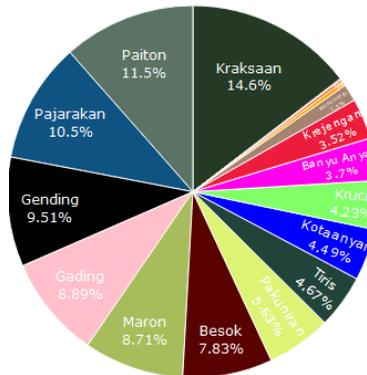
Pada table 6.16 diketahui persebaran wilayah dari ketiga segmen, dan *highlight* lima besar jumlah pelanggan setiap segmen. Dapat diketahui wilayah mana saja yang berpotensi untuk setiap segmennya.

Grafik *bar chart* dan *pie chart* juga digunakan untuk menampilkan persebaran jumlah pelanggan berdasarkan area atau wilayah mereka pada setiap segmen.

Pada gambar 6.7 dan 6.8 dapat dilihat persebaran pelanggan pada semua segmen. Kecamatan kraksaan memiliki jumlah pelanggan terbanyak, dengan 116 pelanggan atau 14,6%. Dan kecamatan mayangan memiliki jumlah pelanggan paling sedikit, yaitu sebanyak 1 pelanggan atau 0.088%.

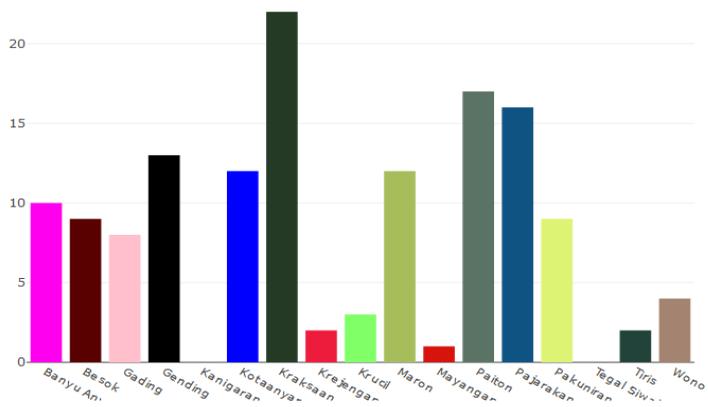


Gambar 6.7 Grafik Bar Chart Persebaran Wilayah Cluster Pada Semua Segmen



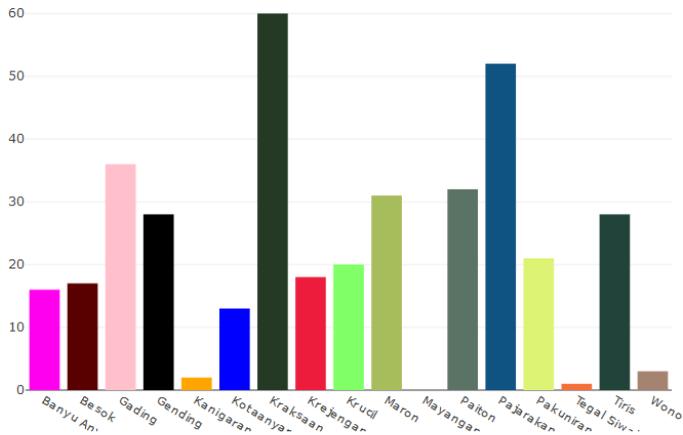
*Gambar 6.8 Grafik Pie Chart Persebaran Wilayah Cluster Pada Semua Segmen*

Pada gambar 6.9 dan 6.10 dapat dilihat persebaran pelanggan pada segmen 1. Kecamatan kraksaan memiliki jumlah pelanggan terbanyak, dengan 22 pelanggan atau 15,7%. Dan kecamatan kanigaran dan tegal siwalan tidak memiliki pelanggan pada segmen 1.

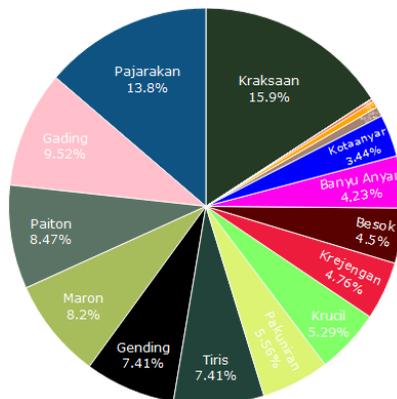


*Gambar 6.9 Grafik Bar Chart Persebaran Wilayah Cluster Pada Segmen 1*





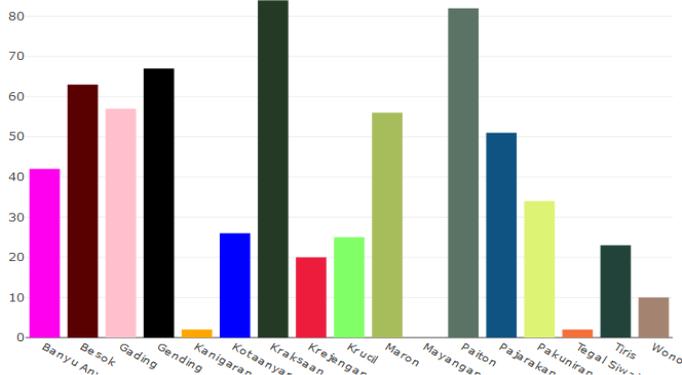
*Gambar 6.11 Grafik Bar Chart Persebaran Wilayah Cluster Pada Segmen 2*



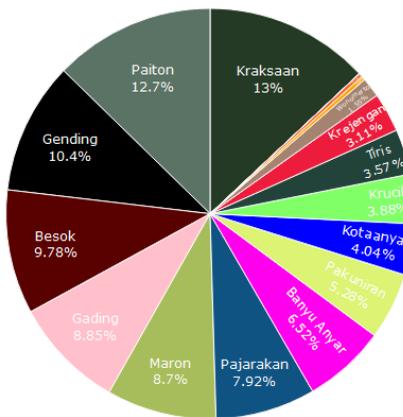
*Gambar 6.12 Grafik Pie Chart Persebaran Wilayah Cluster Pada Segmen 2*

Pada gambar 6.13 dan 6.14 dapat dilihat persebaran pelanggan pada segmen 3. Kecamatan kraksaan memiliki jumlah pelanggan terbanyak, dengan 84 pelanggan atau 13%. Dan

kecamatan Mayangan tidak memiliki pelanggan pada segmen 3.



Gambar 6.13 Grafik Bar Chart Persebaran Wilayah Cluster Pada Segmen 3



Gambar 6.14 Grafik Pie Chart Persebaran Wilayah Cluster Pada Segmen 3

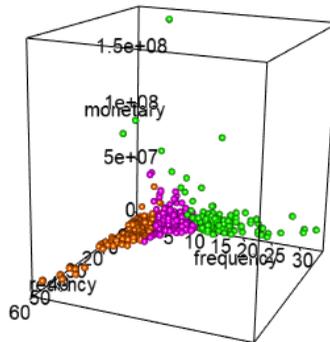
Pada Gambar 6.7 sampai gambar 6.14, dapat diketahui persebaran wilayah pelanggan dari semua segmen, untuk persebaran wilayah *cluster* pada segmen 1, segmen 2, dan segmen 3 menunjukkan hasil bahwa kecamatan Kraksaan mendominasi semua peringkat tertinggi. Hal tersebut disebabkan karena kecamatan Kraksaan adalah ibu kota kabupaten Probolinggo, sehingga roda perekonomian lebih maju dari kecamatan lainnya, dan menyebabkan jumlah toko atau minimarket lebih banyak dari wilayah lainnya.

### **6.6.2 Grafik Scatter Plot 3D**

Grafik *scatter plot 3D* digunakan untuk melihat persebaran supplier dan bagaimana supplier tersebut berkelompok. *Scatter plot 3D* terdiri dari tiga sumbu atau dimensi.

#### **6.6.2.1 Grafik Scatter Plot 3D RFM**

Gambar 6.15 adalah grafik Scatter Plot 3D yang 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 hijau, segmen 2 digambarkan dengan warna oranye, sedangkan segmen 3 digambarkan dengan warna ungu.

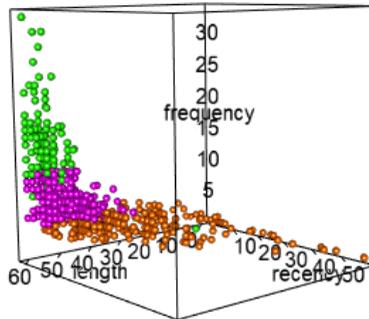


Gambar 6.15 Grafik Scatter Plot 3D RFM

Dari Gambar 6.15 dapat diketahui bahwa pelanggan pada segmen 1 cenderung berkumpul pada variabel *frequency* dan mengarah pada nilai *frequency* dan *monetary* yang semakin tinggi. Pelanggan pada segmen 2 cenderung berkumpul pada variabel *recency* dan mengarah pada nilai yang semakin tinggi. Pelanggan pada segmen 3 cenderung berkumpul titik 0.

#### 6.6.2.2 Grafik Scatter Plot 3D LRF

Gambar 6.16 adalah grafik Scatter Plot 3D yang mendeskripsikan persebaran pelanggan pada ruang tiga dimensi berdasarkan variabel LRF. Sumbu X mewakili variabel *length*, sumbu Y mewakili variabel *recency*, dan sumbu Z mewakili variabel *frequency*. Segmen 1 digambarkan dengan warna hijau, segmen 2 digambarkan dengan warna oranye, sedangkan segmen 3 digambarkan dengan warna ungu.

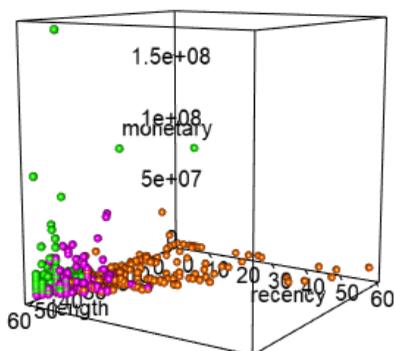


Gambar 6.16 Grafik Scatter Plot 3D LRF

Dari Gambar 6.16 dapat diketahui bahwa pelanggan pada segmen 1 cenderung berkumpul pada variabel *length* dan mengarah pada nilai *frequency* dan *length* yang semakin tinggi. Pelanggan pada segmen 2 cenderung merata pada variabel *recency*, *length*, dan titik 0. Pelanggan pada segmen 3 cenderung berkumpul pada variabel *length* dan mengarah pada nilai *length* dan *frequency* yang semakin tinggi.

### 6.6.2.3 Grafik Scatter Plot 3D LRM

Gambar 6.17 adalah grafik *Scatter Plot 3D* yang mendeskripsikan persebaran pelanggan pada ruang tiga dimensi berdasarkan variabel LRM. Sumbu X mewakili variabel *length*, sumbu Y mewakili variabel *recency*, dan sumbu Z mewakili variabel *monetary*. Segmen 1 digambarkan dengan warna hijau, segmen 2 digambarkan dengan warna oranye, sedangkan segmen 3 digambarkan dengan warna ungu.

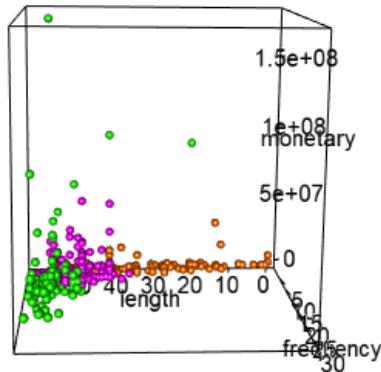


Gambar 6.17 Grafik Scatter Plot 3D LRM

Dari Gambar 6.17 dapat diketahui bahwa pelanggan pada segmen 1 cenderung berkumpul pada variabel *length* dan mengarah pada nilai *monetary* dan *length* yang semakin tinggi. Pelanggan pada segmen 2 cenderung merata pada variable *monetary*, *length*, dan titik 0. Pelanggan pada segmen 3 cenderung berkumpul pada variable *length* dan mengarah pada nilai *length* dan *monetary* yang semakin tinggi.

#### 6.6.2.4 Grafik Scatter Plot 3D LFM

Gambar 6.18 adalah grafik *Scatter Plot 3D* yang mendeskripsikan persebaran pelanggan pada ruang tiga dimensi berdasarkan variabel LRM. Sumbu X mewakili variabel *length*, sumbu Y mewakili variabel *frequency*, dan sumbu Z mewakili variabel *monetary*. Segmen 1 digambarkan dengan warna hijau, segmen 2 digambarkan dengan warna oranye, sedangkan segmen 3 digambarkan dengan warna ungu.



Gambar 6.18 Grafik Scatter Plot 3D LFM

Dari Gambar 6.18 dapat diketahui bahwa pelanggan pada segmen 1, 2, dan 3 cenderung berkumpul pada variable *length* dan mengarah pada nilai *length* dan *monetary* yang semakin tinggi.

### 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.

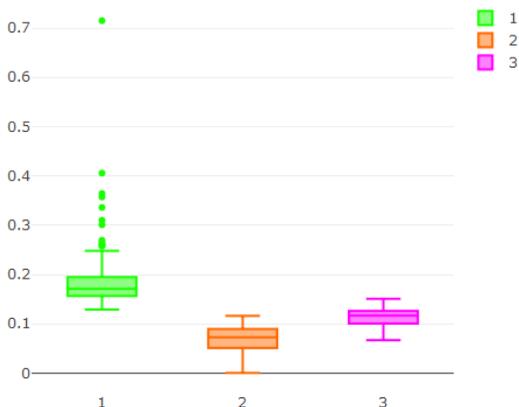
#### 6.6.3.1 Grafik box plot CLV

Gambar 6.19 adalah grafik *box plot* untuk nilai CLV. Segmen 3 memiliki ukuran kotak yang paling pendek dibanding dengan dua segmen lain. Hal ini menunjukkan bahwa sebaran data pada segmen 3 juga kecil. Nilai median segmen 3 tidak terletak di tengah kotak dan Panjang garis *whisker* bagian bawah lebih

panjang, yang berarti segmen 3 tidak bersifat simetris dan cenderung ke *negative skewness*. Garis *whisker* bagian bawah beririsan dengan segmen 2, sehingga menyebabkan *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 bawah lebih panjang. Hal ini menunjukkan bahwa segmen 2 tidak bersifat simetris dan cenderung ke *negative skewness*. Garis *whisker* bagian atas beririsan dengan segmen 1 dan 3, sehingga menyebabkan *overlap*.

Segmen 1 memiliki kotak yang paling panjang, sehingga persebaran data menjadi lebar. Nilai median segmen 1 tidak terletak di tengah kotak dan panjang garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan bahwa segmen 1 tidak simetris dan cenderung ke *positive skewness*. Selain itu terdapat data *outlier* di bagian atas kotak.

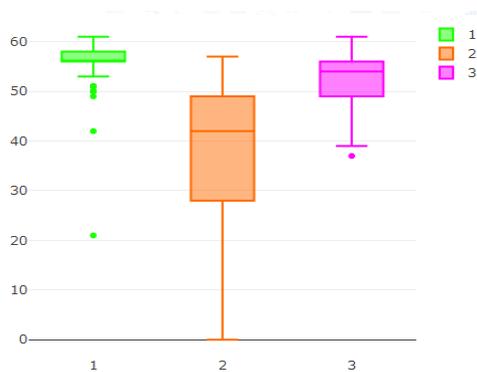


Gambar 6.19 Grafik Box Plot CLV Semua Segmen

### 6.6.3.2 Grafik box plot length

Gambar 6.20 adalah grafik *box plot* untuk nilai *recency*. Segmen 1 memiliki ukuran kotak yang paling pendek dibanding dengan dua segmen lain, sehingga persebaran data menjadi lebih kecil. Segmen 1 tidak memiliki nilai median dan garis *whisker* bagian bawah lebih panjang. Hal ini menunjukkan menunjukkan bahwa segmen 1 tidak bersifat simetris dan cenderung ke *negative skewness*. Garis *whisker* bagian bawah sedikit beririsan dengan segmen 2 sehingga menyebabkan sedikit *overlap*. Dan juga terdapat data *outlier* di bawah kotak.

Segmen 2 memiliki kotak yang paling Panjang dibanding dua segmen lainnya, sehingga persebaran data lebih besar dibanding segmen lainnya. Nilai median segmen 2 tidak terletak di tengah kotak dan panjang garis *whisker* bagian bawah lebih panjang. Hal ini menunjukkan bahwa segmen 2 tidak bersifat simetris dan cenderung ke *negative skewness*. Sebagian kotak segmen 2 beririsan dengan segmen 3, sehingga menyebabkan banyak *overlap*.



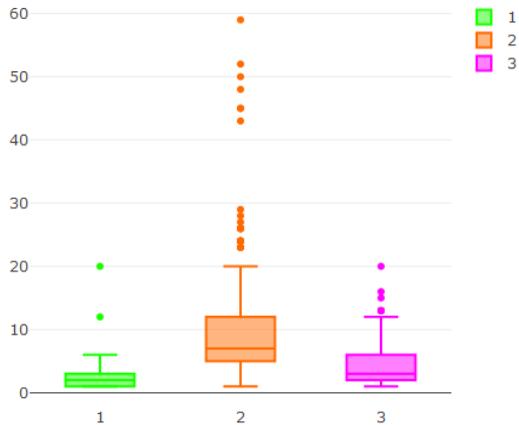
Gambar 6.20 Grafik Box Plot Recency Semua Segmen

Segmen 3 memiliki kotak yang lebih pendek, sehingga persebaran data menjadi lebih kecil. 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*. Selain itu terdapat data *outlier* di bagian bawah kotak.

### 6.6.3.3 Grafik box plot recency

Gambar 6.21 adalah grafik *box plot* untuk nilai *recency*. Segmen 1 memiliki ukuran kotak yang paling pendek dibanding dengan dua segmen lain, sehingga persebaran data menjadi lebih kecil. Nilai median segmen 1 berada di tengah kotak dan garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan menunjukkan bahwa segmen 1 bersifat simetris dan cenderung ke *positive skewness*. Garis *whisker* bagian atas sedikit berurusan dengan segmen 2 sehingga menyebabkan sedikit *overlap*, selain itu juga terdapat *outlier* dibagian atas kotak.

Segmen 2 memiliki kotak yang lebih Panjang dari dua segmen lainnya, sehingga persebaran data lebih besar dibanding dua segmen lainnya. 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 berurusan dengan segmen 3, sehingga menyebabkan banyak *overlap*, dan memiliki banyak data *outlier* dibagian atas kotak.

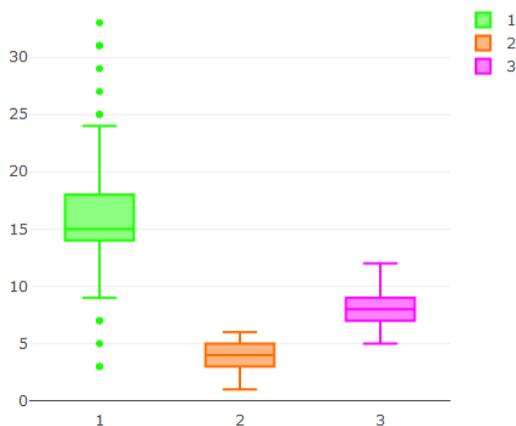


Gambar 6.21 Grafik Box Plot Recency Semua Segmen

Segmen 3 memiliki kotak yang lebih pendek, sehingga persebaran data menjadi lebih kecil disbanding segmen 2. 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.

#### 6.6.3.4 Grafik box plot frequency

Gambar 6.22 adalah grafik *box plot* untuk nilai *frequency*. Segmen 1 memiliki ukuran kotak yang paling panjang dibanding dengan dua segmen lain. Hal ini menunjukkan bahwa sebaran data pada segmen 1 juga paling besar. Nilai median segmen 1 tidak terletak di tengah kotak, dan garis *whisker* bagian atas lebih panjang. Selain itu terdapat data *outlier* dibagian atas dan bawah kotak.



Gambar 6.22 Grafik Box Plot Frequency Semua Segmen

Segmen 2 memiliki kotak yang lebih pendek, sehingga sebaran data pada segmen 2 lebih kecil. Nilai median segmen 2 terletak di tengah kotak, tetapi garis *whisker* bagian bawah lebih panjang. Sebagian kotak beririsan dengan garis *whisker* segmen 3, sehingga terjadi sedikit *overlap*.

Nilai median segmen 3 terletak di tengah kotak dan panjang garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan bahwa segmen 3 simetris dan cenderung ke *positive skewness*.

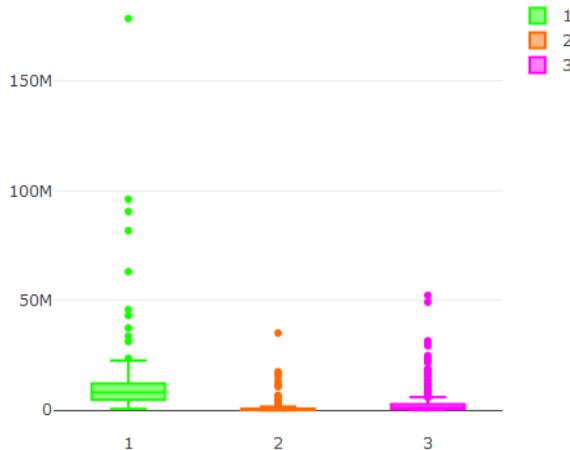
#### 6.6.3.5 Grafik box plot monetary

Gambar 6.23 adalah grafik *box plot* untuk nilai *monetary*. Segmen 1 memiliki ukuran kotak yang paling panjang dibanding dengan dua segmen lain. Hal ini menunjukkan bahwa sebaran data pada segmen 1 juga paling besar. Nilai median segmen 1 terletak di tengah kotak dan garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan menunjukkan bahwa segmen 1 bersifat simetris dan cenderung ke *positive skewness*.

Selain itu, terdapat beberapa data yang *outlier* di bagian atas kotak.

Segmen 2 memiliki kotak yang tidak begitu jelas, sehingga tidak dapat menentukan nilai median dan kecenderungan garis *whiskernya*. Sebagian kotak beririsan dengan kotak segmen 3, sehingga terjadi banyak *overlap*. Dan juga memiliki beberapa data *outlier* diatas kotak.

Segmen 3 memiliki kotak yang hamper sama dengan segmen 2, sehingga tidak dapat menentukan nilai mediannya. Tetapi panjang garis *whisker* bagian atas lebih panjang. Hal ini menunjukkan bahwa segmen 3 cenderung ke *positive skewness*. Sebagian kotak beririsan dengan kotak segmen 2 sehingga terjadi banyak *overlap*. Dan memiliki beberapa data *outlier* diatas kotak.



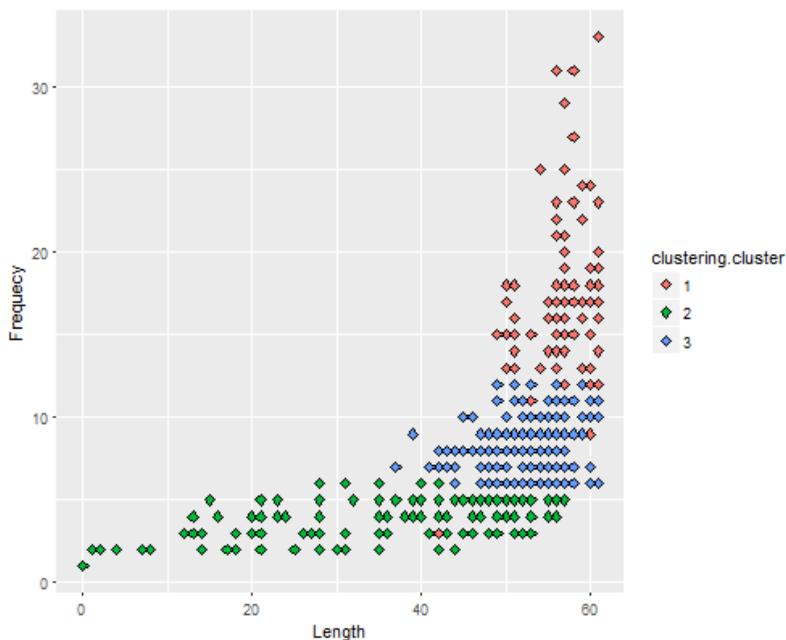
Gambar 6.23 Grafik Box Plot Monetary Semua Segmen

### 6.6.4 Grafik *Scatter Plot 2D*

Grafik *Scatter plot 2D* digunakan untuk melihat pola persebaran supplier dan bagaimana supplier tersebut berkelompok diantara dua variabel. *Scatter plot 2D* terdiri dua sumbu atau dimensi.

#### 6.6.4.1 Grafik *Scatter Plot 2D LF*

Gambar 6.24 adalah grafik *Scatter Plot 2D* yang mendeskripsikan persebaran pelanggan pada gambar dua dimensi berdasarkan variabel LF. Sumbu X mewakili variabel *length*, sumbu Y mewakili variabel *frequency*.



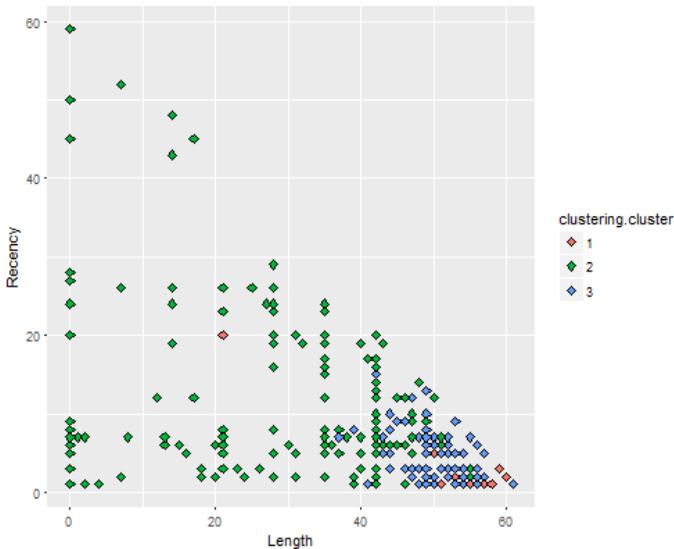
Gambar 6.24 Grafik *Scatter Plot 2D LF*

Dari Gambar 6.24 dapat diketahui bahwa pelanggan pada segmen 1 cenderung berkumpul pada variabel *length* dan

*frequency* yang mengarah pada nilai yang semakin tinggi, yang berarti bahwa pelanggan sering melakukan transaksi pembelian. Pelanggan pada segmen 2 cenderung merata pada variabel *length* dari nilai rendah dan mengarah pada nilai yang tinggi, sedangkan pada variabel *frequency* cenderung pada nilai yang rendah, yang mengidentifikasi bahwa pelanggan jarang melakukan transaksi pembelian. Pelanggan pada segmen 3 cenderung berkumpul pada variabel *length* dan *frequency* yang mengarah pada nilai yang cukup tinggi yang menandakan bahwa pelanggan cukup sering melakukan transaksi pembelian.

#### 6.6.4.2 Grafik Scatter Plot 2D LR

Gambar 6.25 adalah grafik *Scatter Plot 2D* yang mendeskripsikan persebaran pelanggan pada gambar dua dimensi berdasarkan variabel LR. Sumbu X mewakili variabel *length*, sumbu Y mewakili variabel *recency*.

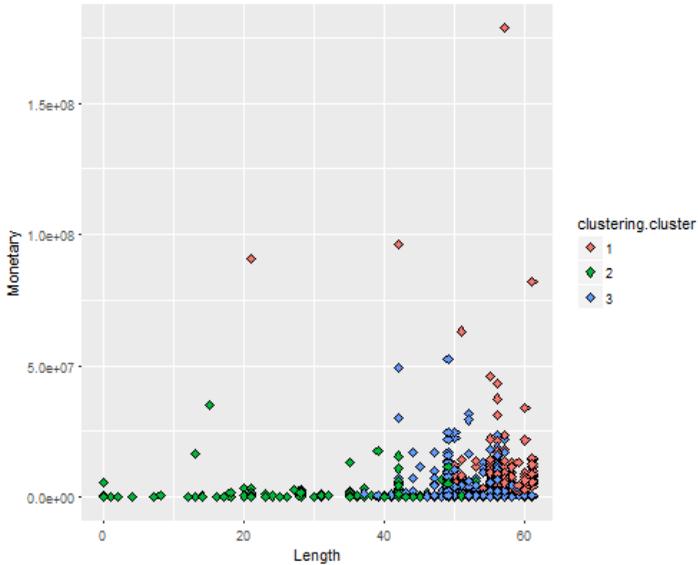


Gambar 6.25 Grafik Scatter Plot 2D LR

Dari Gambar 6.25 dapat diketahui bahwa pelanggan pada segmen 1 cenderung berkumpul pada variabel *length* yang memiliki nilai tinggi dan nilai *recency* yang rendah, yang menandakan bahwa pelanggan baru melakukan transaksi penjualan. Pelanggan pada segmen 2 cenderung merata pada variabel *length* dan memiliki nilai *recency* yang tinggi, membuktikan bahwa pelanggan cukup lama tidak melakukan transaksi pembelian. Pelanggan pada segmen 3 cenderung berkumpul pada variabel *length* yang tinggi dan beberapa pelanggan yang mengarah pada nilai *recency* yang cenderung rendah, hal ini menandakan bahwa pelanggan baru saja melakukan transaksi pembelian.

#### 6.6.4.3 Grafik Scatter Plot 2D LM

Gambar 6.26 adalah grafik *Scatter Plot 2D* yang mendeskripsikan persebaran pelanggan pada gambar dua dimensi berdasarkan variabel LM. Sumbu X mewakili variabel *length*, sumbu Y mewakili variabel *monetary*.

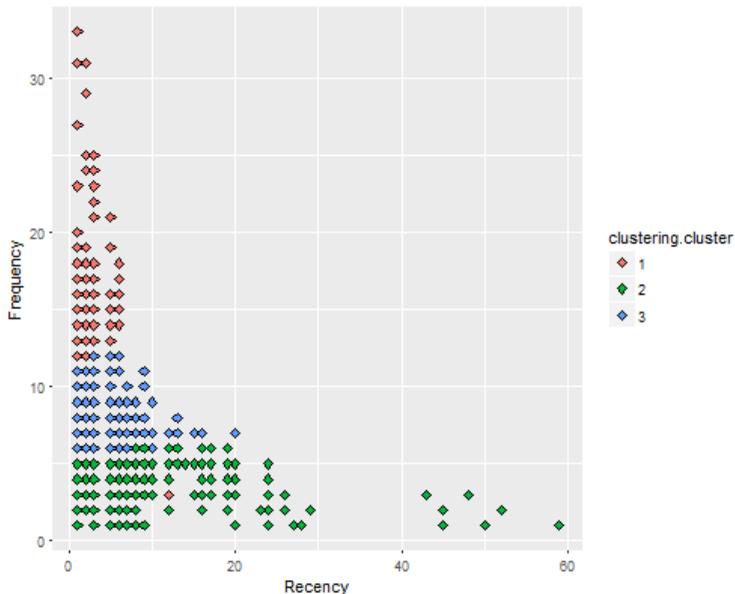


Gambar 6.26 Grafik Scatter Plot 2D LM

Dari Gambar 6.26 dapat diketahui bahwa pelanggan pada segmen 1 cenderung berkumpul pada variabel *length* yang memiliki nilai tinggi dan nilai *monetary* yang tinggi, menunjukkan bahwa pelanggan memiliki nilai transaksi pembelian yang tinggi. Pelanggan pada segmen 2 cenderung merata pada variable *length* dan *monetary* yang rendah, dan membuktikan bahwa pelanggan memiliki nilai transaksi yang rendah. Pelanggan pada segmen 3 cenderung berkumpul pada variable *length* yang tinggi dan beberapa pelanggan yang mengarah pada nilai *monetary* yang cukup tinggi, dan menunjukkan bahwa pelanggan memiliki nilai transaksi yang cukup tinggi.

#### 6.6.4.4 Grafik Scatter Plot 2D RF

Gambar 6.27 adalah grafik *Scatter Plot 2D* yang mendeskripsikan persebaran pelanggan pada gambar dua dimensi berdasarkan variabel RF. Sumbu X mewakili variabel *recency*, sumbu Y mewakili variabel *frequency*.



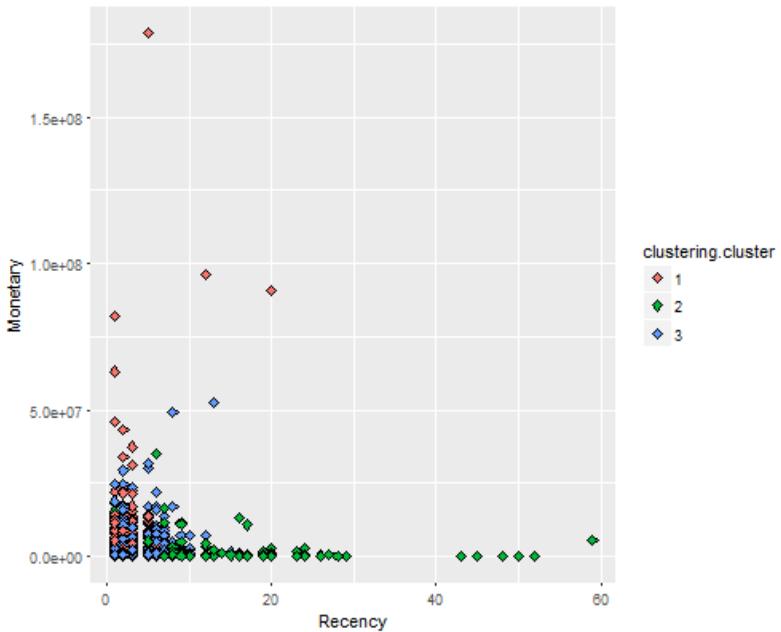
Gambar 6.27 Grafik Scatter Plot 2D RF

Dari Gambar 6.27 dapat diketahui bahwa pelanggan pada segmen 1 cenderung berkumpul pada variabel *recency* yang memiliki nilai rendah dan nilai *frequency* yang tinggi, mengidentifikasi bahwa pelanggan sering melakukan transaksi pembelian. Pelanggan pada segmen 2 cenderung merata pada variabel *recency* dan nilai *frequency* yang cenderung rendah. Pelanggan pada segmen 3 cenderung

berkumpul pada variable *recency* yang rendah frequency mengarah pada nilai yang cukup tinggi.

#### 6.6.4.5 Grafik Scatter Plot 2D RM

Gambar 6.28 adalah grafik *Scatter Plot 2D* yang mendeskripsikan persebaran pelanggan pada gambar dua dimensi berdasarkan variabel RM. Sumbu X mewakili variabel *recency*, sumbu Y mewakili variabel *monetary*.



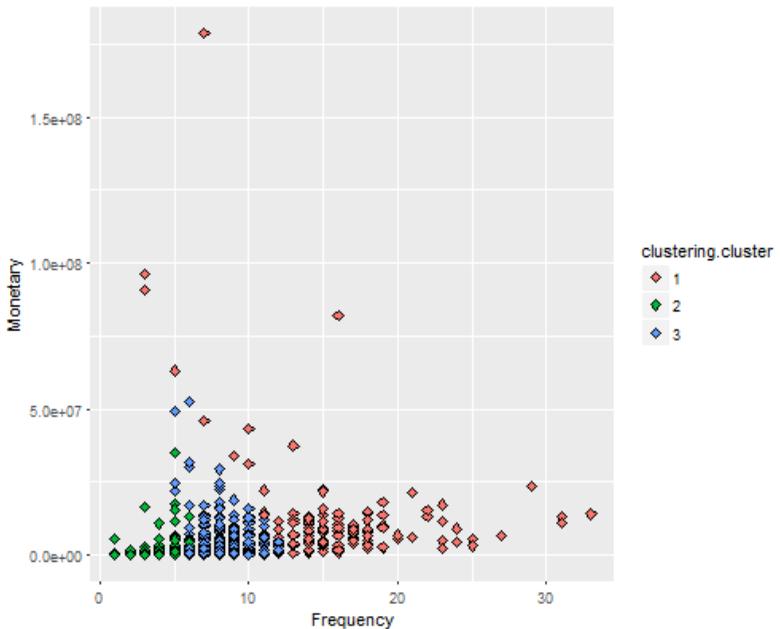
Gambar 6.28 Grafik Scatter Plot 2D RM

Dari Gambar 6.28 dapat diketahui bahwa pelanggan pada segmen 1 cenderung berkumpul pada variabel *recency* yang memiliki nilai rendah dan nilai *monetary* yang tinggi, menunjukkan bahwa pelanggan memiliki nilai transaksi pembelian yang tinggi. Pelanggan pada segmen 2 cenderung

merata pada variable *recency* dan nilai *monetary* yang rendah. Pelanggan pada segmen 3 cenderung berkumpul pada variable *recency* yang rendah dan mengarah pada nilai *monetary* yang cukup tinggi.

#### 6.6.4.6 Grafik Scatter Plot 2D FM

Gambar 6.29 adalah grafik *Scatter Plot 2D* yang mendeskripsikan persebaran pelanggan pada gambar dua dimensi berdasarkan variabel FM. Sumbu X mewakili variabel *frequency*, sumbu Y mewakili variabel *monetary*.



Gambar 6.29 Grafik Scatter Plot 2D FM

Dari Gambar 6.29 dapat diketahui bahwa pelanggan pada segmen 1 cenderung berkumpul pada variabel *frequency* yang memiliki nilai tinggi dan nilai *monetary* yang tinggi,

membuktikan bahwa pelanggan sering melakukan transaksi penjualan dan memiliki nilai transaksi pembelian yang tinggi. Pelanggan pada segmen 2 cenderung rendah pada variable *frequency* dan *monetary*, yang menunjukkan bahwa pelanggan jarang melakukan transaksi dan memiliki nilai transaksi pembelian yang rendah. Pelanggan pada segmen 3 cenderung berkumpul pada variable *frequency* yang cukup tinggi dan beberapa pelanggan yang mengarah pada nilai *monetary* yang cukup tinggi, mengidentifikasi bahwa pelanggan cukup sering melakukan transaksi pembelian dan memiliki nilai transaksi pembelian yang cukup tinggi.

## 6.7 Percobaan

Percobaan dilakukan untuk menentukan variasi dari hasil *cluster*. Percobaan yang dilakukan adalah dengan mengganti jumlah K menjadi empat, dan lima. Dan tidak menggunakan bobot untuk masing-masing variable LRFM ketika melakukan *clustering*.

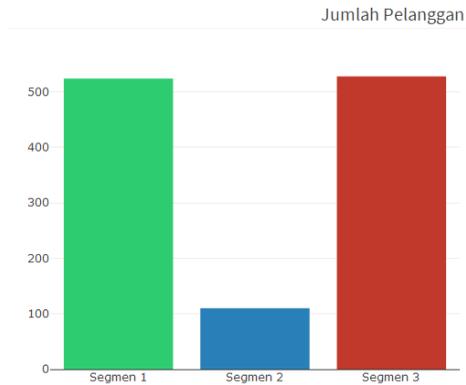
### 6.7.1 Percobaan 1 (tanpa bobot)

Pada percobaan pertama dalam menentukan nilai K dan *clustering*, dengan menggunakan nilai LRFM hasil normalisasi (belum diberi bobot). Dihasilkan nilai K=3 dan karakteristik tiga segmen pada tabel 6.16.

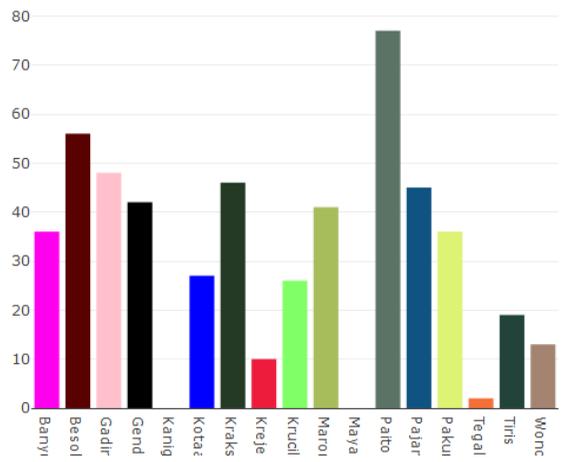
Tabel 6.167 Perbandingan Karakteristik 3 Segmen (tanpa bobot)

	Segmen 1	Segmen 2	Segmen 3
<i>Peringkat</i>	1	3	2
<i>Jumlah Anggota</i>	524	110	528
<i>Length</i>	48 – 61 hari	0 - 31 hari	31 - 56 hari
<i>Recency</i>	1 - 9 hari	1 - 59 hari	1 - 24 hari
<i>Frequency</i>	5 - 33 kali	1 - 6 kali	2 - 10 kali

<b>Monetary</b>	Rp 113.072 – Rp 178.526.088	Rp 4.620 – Rp 90.626.503	Rp 8.120 – Rp 96.180.674
<b>CLV</b>	0.14089	0.04301	0.09052



Gambar 6.30 Grafik Barchart 3 segmen (tanpa bobot)



Gambar 6.31 Grafik Barchart persebaran wilayah pada segmen 1 (tanpa bobot)

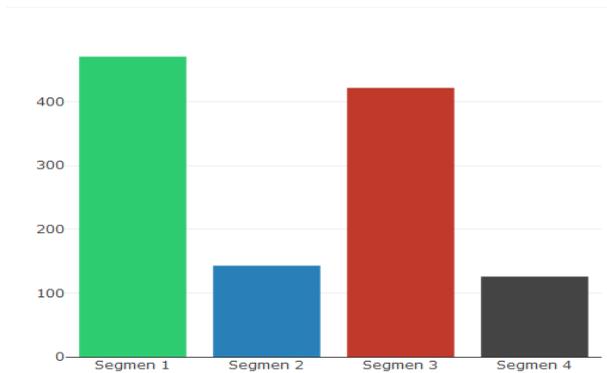
Dari tabel 6.16 dapat dilihat bahwa hasil percobaan dengan tidak menggunakan bobot, memiliki hasil yang berbeda dengan hasil yang menggunakan bobot. Dapat dilihat pada gambar 6.30 hasil persebaran wilayah pada segmen 1. Paiton memiliki jumlah pelanggan yang paling banyak, berbeda dengan hasil *clustering* menggunakan bobot, yang menunjukkan bahwa kecamatan Kraksaan selalu mendominasi jumlah wilayah pada semua *cluster*.

### 6.7.2 Percobaan 2 (K=4)

Pada percobaan kedua dilakukan dengan merubah nilai K=4, saat melakukan clustering. Didapatkan hasil *cluster* empat segmen dengan karakteristik pada table 6.17.

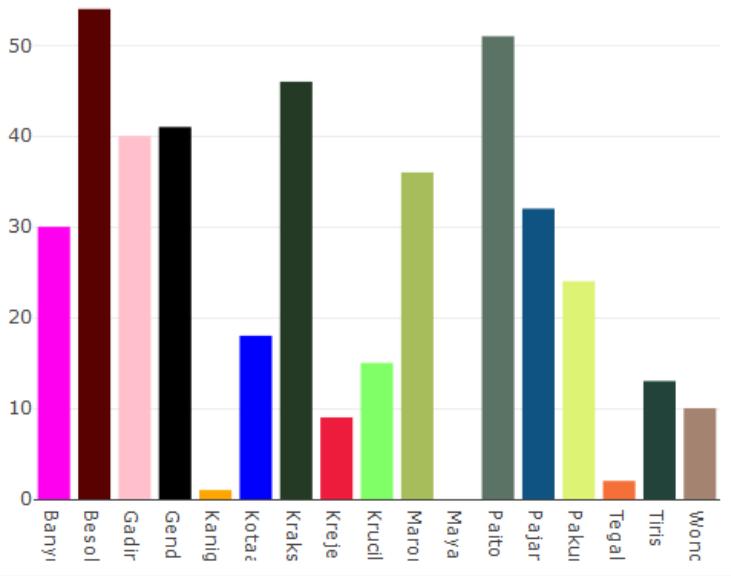
Tabel 6.178 Perbandingan Karakteristik 4 Segmen (K=4)

	Segmen 1	Segmen 2	Segmen 3	Segmen 4
<b>Peringkat</b>	3	4	2	1
<b>Jumlah Anggota</b>	471	143	422	126
<b>Length</b>	28 – 61 hari	0 - 44 hari	39 - 61 hari	42 – 61 hari
<b>Recency</b>	1 - 24 hari	1 - 59 hari	1 - 13 hari	1 – 12 hari
<b>Frequency</b>	3 - 7 kali	1 - 5 kali	5 - 13 kali	3 – 33 kali
<b>Monetary</b>	Rp 13.620 – Rp 49.294.618	Rp 4.620 – Rp 90.626.503	Rp 110.596 – Rp 63.187.930	Rp 971.390 – Rp 178.526.088
<b>CLV</b>	0.09398	0.04618	0.12174	0.19132



Gambar 6.32 Grafik Barchart 4 segmen ( $K=4$ )

Dari tabel 6.17 dapat dilihat bahwa hasil percobaan dengan tidak menggunakan bobot, memiliki hasil yang berbeda dengan hasil yang menggunakan bobot. Dapat dilihat pada gambar 6.33 hasil persebaran wilayah pada segmen 3, Besok memiliki jumlah pelanggan yang paling banyak, berbeda dengan hasil *clustering* menggunakan bobot, yang menunjukkan bahwa kecamatan Kraksaan selalu mendominasi jumlah wilayah pada semua *cluster*, dan menunjukkan hasil yang cenderung merata.



Gambar 6.33 Grafik Barchart Persebaran wilayah pada segmen 3(K=4)

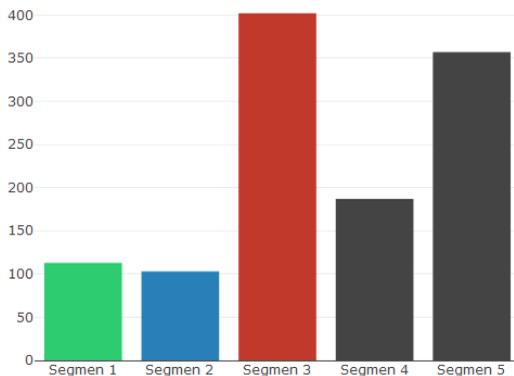
### 6.7.3 Percobaan 3 (K=5)

Pada percobaan kedua dilakukan dengan merubah nilai K=5, saat melakukan clustering. Didapatkan hasil *cluster* empat segmen dengan karakteristik pada table 6.18.

Tabel 6.189 Perbandingan Karakteristik 5 Segmen (K=5)

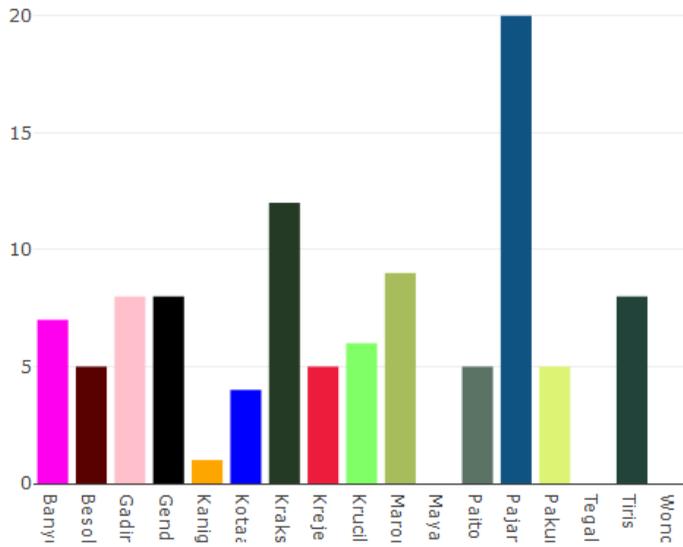
	Segmen 1	Segmen 2	Segmen 3	Segmen 4	Segmen 5
<i>Peringkat</i>	1	5	3	2	4
<i>Jumlah Anggota</i>	113	103	402	187	357
<i>Length</i>	49 – 61 hari	0 - 35 hari	37 - 61 hari	21 – 61 hari	23 – 57 hari
<i>Recency</i>	1 - 6 hari	1 - 59 hari	1 - 20 hari	1 – 20 hari	1 – 24 hari
<i>Frequency</i>	7- 33 kali	1 - 5 kali	6 - 9 kali	3 – 14 kali	2 – 6 kali

<b>Monetary</b>	Rp 1.029.783 – Rp 178.526.088	Rp 4.620 – Rp 35.207.012	Rp 110.596 – Rp 17.264.473	Rp 221.730 – Rp 96.180.647	Rp 8.120 – Rp 24.764.091
<b>CLV</b>	0.19564	0.04094	0.11188	0.13181	0.08516



Gambar 6.34 Grafik Barchart 5 segmen ( $K=5$ )

Dari tabel 6.18 dapat dilihat bahwa hasil percobaan dengan tidak menggunakan bobot, memiliki hasil yang berbeda dengan hasil yang menggunakan bobot. Dapat dilihat pada gambar 6.35 hasil persebaran wilayah pada segmen 2, Pajajaran memiliki jumlah pelanggan yang paling banyak, berbeda dengan hasil *clustering* menggunakan bobot, yang menunjukkan bahwa kecamatan Kraksaan selalu mendominasi jumlah wilayah pada semua *cluster*, dan menunjukkan hasil yang cenderung merata.



Gambar 6.35 Grafik Bar chart Persebaran Wilayah pada segmen 2 ( $K=5$ )

## 6.8 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 0.598. Hal ini berarti bahwa variabel *monetary* menjadi pertimbangan paling penting dalam menentukan peringkat masing-masing *cluster* dan pelanggan. Sedangkan variabel *length*, *recency* dan *frequency* masing-masing memiliki bobot 0.073, 0.05 dan 0.279.
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 1162 pelanggan yang di-*cluster*, 140 pelanggan tergabung pada *cluster* 1, 378 pelanggan tergabung pada *cluster* 2, dan 644 pelanggan tergabung pada *cluster* 3.
4. *Cluster* 1 mempunyai rata-rata nilai CLV paling tinggi. 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* 1 adalah *cluster* terbaik. Dengan karakteristik tersebut, perusahaan dapat menerapkan strategi promosi yang bertujuan untuk menjaga mempertahankan loyalitas pelanggan, misalkan memberikan diskon terhadap produk dalam periode tertentu.
5. *Cluster* 2 memiliki rata-rata 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* 2 adalah *cluster* terburuk. Dengan karakteristik tersebut, perusahaan dapat menerapkan strategi promosi yang bertujuan untuk menarik kembali pelanggan, misalkan dengan melakukan rutin memberikan penawaran-penawaran menarik terbaru.
6. *Cluster* 3 memiliki rata-rata CLV sedang. begitu juga dengan nilai ketiga variabel. Sehingga dapat disimpulkan bahwa *cluster* 3 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*.
7. Persebaran wilayah pelanggan dari semua segmen dengan menggunakan bobot dan  $K=3$ , untuk persebaran wilayah *cluster* pada segmen 1, segmen 2, dan segmen 3

menunjukkan hasil bahwa kecamatan kraksaan mendominasi peringkat tertinggi. Hal tersebut disebabkan karena kecamatan kraksaan adalah ibu kota kabupaten Probolinggo, sehingga roda perekonomian lebih maju dari kecamatan lainnya, dan menyebabkan jumlah toko atau minimarket lebih banyak dari wilayah lainnya.

8. Pada percobaan 1, *clustering* tanpa menggunakan bobot dengan  $K=3$ , persebaran wilayah *cluster* menjadi lebih merata. Kecamatan Kraksaan tidak lagi menjadi wilayah yang mendominasi pada semua *cluster*. Segmen 1 memiliki nilai *monetary*, *frequency*, dan *length* yang tinggi, sedangkan nilai *recency* yang tinggi. Segmen 2 memiliki nilai *recency* yang rendah dan nilai *monetary*, *frequency*, dan *length* yang rendah. Segmen 3 memiliki nilai *monetary*, *frequency*, dan *length* yang sedang, sedangkan nilai *recency* yang rendah.
9. Pada percobaan 2 dan 3, *clustering* menggunakan bobot dengan  $K=4$  dan  $K=5$  pada masing-masing percobaan. Menghasilkan persebaran wilayah *cluster* yang lebih merata.

*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, yaitu:

1. Penggalan data dengan *clustering* dapat membantu pihak perusahaan dalam mengetahui segmentasi pelanggan yang dimiliki. *Clustering* mampu mengelompokkan pelanggan-pelanggan dalam segmen-segmen 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 selisih transaksi awal dan transaksi terakhir, rentang waktu terakhir transaksi, jumlah transaksi, dan total pemasukan. Dengan pembagian segmen-segmen 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 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, yaitu :

1. Penelitian ini dapat dikembangkan dengan memperhatikan atribut lain selain atribut LRFM untuk memberikan variasi hasil *clustering*, misalkan dengan menambahkan variabel lokasi pelanggan.
2. Rentang data transaksi sebagai masukan lebih baik diperbanyak lebih dari 2 bulan, menjadi 6 bulan atau 1 tahun transaksi agar hasil *clustering* lebih akurat.
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] P. A. Julianto, “Sri Mulyani: Sumber Kekayaan Bukan Lagi Tambang Emas, tapi Data,” *kompas.com*, 2018. [Online]. Available: <https://ekonomi.kompas.com/read/2018/03/02/201029426/sri-mulyani-sumber-kekayaan-bukan-lagi-tambang-emas-tapi-data>. [Accessed: 05-Mar-2018].
- [2] G. Gumelar, “BPS Sebut Masyarakat Masih Akan Tahan Konsumsi di Awal 2018,” *cnnindonesia.com*, 2018. [Online]. Available: <https://www.cnnindonesia.com/ekonomi/20180205131521-92-273869/bps-sebut-masyarakat-masih-akan-tahan-konsumsi-di-awal-2018>. [Accessed: 05-Mar-2018].
- [3] Lavinda, “Konsumsi Rumah Tangga Lesu, Omzet Unilever Cuma Naik 2,9%,” *cnnindonesia.com*, 2018. [Online]. Available: <https://www.cnnindonesia.com/ekonomi/20180202133313-92-273390/konsumsi-rumah-tangga-lesu-omzet-unilever-cuma-naik-29>. [Accessed: 05-Mar-2018].
- [4] M. J. a. Berry and G. S. Linoff, *Data mining techniques: for marketing, sales, and customer relationship management*. 2004.

- [5] P.-N. Tan, M. Steinbach, A. Karpatne, and V. Kumar, "Introduction to Data Mining," 2nd ed., 2005.
- [6] M. Diana, J. David, and R. Cote, "Exploring the benefits of a traveller clustering approach based on multimodality attitudes and behaviours," *Transp. Res. Procedia*, vol. 25, pp. 2556–2569, 2017.
- [7] M. Alvandi, S. Fazli, and F. S. Abdoli, "K-Mean Clustering Method For Analysis Customer Lifetime Value With LRFM Relationship Model In Banking Services," vol. 3, no. 11, pp. 2294–2302, 2012.
- [8] W. J. Reinartz and V. Kumar, "On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing," *J. Mark.*, vol. 64, no. 4, pp. 17–35, 2000.
- [9] N. Y. Ratna Ekawati, "Klasifikasi Usaha Kecil Dan Menengah ( Ukm ) Clustering Wilayah Kota Cilegon," *Klasifikasi Usaha Kecil Dan Menengah ( Ukm ) Clust. Wil. Kota Cilegon*, pp. 1–7, 2013.
- [10] I. I. Sari, "SEGMENTASI PELANGGAN MENGGUNAKAN ALGORITMA FUZZY C-MEANS PADA CV . SWALAYAN MOTOR," 2017.
- [11] A. Ramadhan, Z. Efendi, and Mustakim, "Perbandingan K-Means dan Fuzzy C-Means untuk Pengelompokan Data User Knowledge Modeling," *Semin. Nas. Teknol.*

*Informasi, Komun. dan Ind.* 9, *Fak. Sains dan Teknol. UIN Sultan Syarif Kasim Riau*, pp. 219–226, 2017.

- [12] D. C. Li, W. L. Dai, and W. T. Tseng, “A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business,” *Expert Syst. Appl.*, vol. 38, no. 6, pp. 7186–7191, 2011.
- [13] A. Hiziroglu and S. Sengul, “Investigating Two Customer Lifetime Value Models from Segmentation Perspective,” *Procedia - Soc. Behav. Sci.*, vol. 62, pp. 766–774, 2012.
- [14] D. B. Saputra, “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 ) IMPLEMENTASI FUZZY C-MEANS DAN MODEL RFM UNTUK SEGMENTASI PELA,” 2018.
- [15] L. Tari, C. Baral, and S. Kim, “Fuzzy c-means clustering with prior biological knowledge,” *J. Biomed. Inform.*, vol. 42, no. 1, pp. 74–81, 2009.
- [16] K. Anderson and C. Kerr, *Customer Relationship*, vol. 5, no. 5. 2011.

- [17] R. Baran, C. Zerres, and M. Zerres, “Customer Relationship Management.”
- [18] M. Viljoen, J. A. Bennett, A. D. Berndt, and C. R. Van Zyl, “The use of technology in customer relationship management (CRM),” *Acta Commer.*, vol. 5, no. 1, pp. 106–116, 2005.
- [19] K. Tsipstsis, *Data Mining Tehniques in CRM: Inside Customer Segmentation*. 2010.
- [20] M. J. A. Berry and G. S. Linoff, *Mastering Data Mining - The Art and Science of Customer Relationship Management*. 1999.
- [21] B. G. Mirkin, *Clustering for data mining: a data recovery approach*. 2012.
- [22] P. K. Sri, *Aplikasi Logika Fuzzy untuk Pendukung Keputusan*. Yogyakarta: Graha Ilmu, 2004.
- [23] D. Birant, “Data Mining Using RFM Analysis,” *Knowledge-Oriented Appl. Data Min.*, no. iii, pp. 91–108, 2011.
- [24] L. Al Shalabi, Z. Shaaban, and B. Kasasbeh, “Data Mining: A Preprocessing Engine,” *J. Comput. Sci.*, vol. 2, no. 9, pp. 735–739, 2006.
- [25] R. L. Hale, “Cluster analysis in school psychology: An example,” *J. Sch. Psychol.*, vol. 19, no. 1, pp. 51–56, 1981.

- [26] T. S. Madhulatha, “an Overview on Clustering Methods,” *IOSR J. Eng.*, vol. 2, no. 4, pp. 719–725, 2012.
- [27] M. Information, *Quantitative Management Research*. 2010.
- [28] S. Square, J. Ale, and A. Highway, “An integrated AHP-RFM method to banking customer segmentation Seyed Mahdi Rezaeinia Abbas Keramati \* Amir Albadvi,” *Electron. Cust. Relatsh. Manag.*, vol. 6, no. 2, pp. 153–168, 2012.
- [29] E. Rendón, I. Abundez, A. Arizmendi, and E. M. Quiroz, “Internal versus External cluster validation indexes,” *Int. J. Comput. Commun.*, vol. 5, no. 1, pp. 27–34, 2011.
- [30] Y. Liu, Z. Li, H. Xiong, X. Gao, J. Wu, and S. Wu, “Understanding and enhancement of internal clustering validation measures,” *IEEE Trans. Cybern.*, vol. 43, no. 3, pp. 982–994, 2013.

## BIODATA PENULIS



Penulis bernama lengkap Nolan Firdaus B, lahir di Probolinggo, 20 Februari 1996. Penulis merupakan anak pertama dari dua bersaudara. Penulis mengenyam pendidikan formal Sekolah Dasar di SD Negeri Pajarakan Kulon 1 dan melanjutkan ke jenjang Sekolah Menengah Pertama di SMP Negeri 1 Kraksaan. Lulus tahun 2011, penulis melanjutkan di SMA Negeri 1 Probolinggo. Tiga tahun berselang, penulis mendaftarkan kuliah di Departemen Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi, Institut Teknologi Sepuluh Nopember Surabaya melalui jalur SNMPTN. Selama menjadi mahasiswa, penulis aktif berorganisasi dan mengikuti kepanitiaan di tingkat fakultas, maupun institut. Penulis terdaftar sebagai staf *Competition Test Maker, Information System Expo 2015*; staf *Internal Affairs BEM FTIf 2015/2016*; dan Anggota Dewan Perwakilan Mahasiswa 2016/2017. Dalam pengerjaan Tugas Akhir, penulis memilih laboratorium Rekayasa Data dan Intelektual Bisnis, dengan bimbingan Bapak Edwin Riksakomara, S.Kom., M.T., dengan topik mengenai penggalan data. Penulis dapat dihubungi melalui email [firdaus.nolan@gmail.com](mailto:firdaus.nolan@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 dibanding 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. 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
  - d. *Length* : Jarak waktu (hari) dari transaksi pertama dengan transaksi terakhir dalam periode data yang telah ditentukan

**CONTOH PENGISIAN KUESIONER**

Kriteria																			Kriteria
<i>Recency</i>	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<i>Frequency</i>	
													√						

Artinya :

Pada pernyataan tersebut, tanda centang (√) di kolom jawaban sebelah kiri menunjukkan bahwa: *recency* 1/5 kali lebih penting dibandingkan dengan *frequency* dalam hal peringkat 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 kriteria pada kolom sebelah kiri dengan kriteria pada kolom sebelah kanan?”

Kriteria																			Kriteria
<i>Recency</i>	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<i>Frequency</i>	
<i>Frequency</i>	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<i>Monetary</i>	
<i>Recency</i>	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<i>Monetary</i>	
<i>Length</i>	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<i>Frequency</i>	
<i>Monetary</i>	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<i>Length</i>	
<i>length</i>	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<i>Recency</i>	

***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*

<b>UC-2 Melihat Grafik Scatter Plot 3D</b>		
<b>Purpose</b>	Mengetahui persebaran pelanggan dalam grafik <i>scatter plot</i> 3 dimensi, berdasarkan RFM, LRM, LRF, dan LFM	
<b>Overview</b>	Dimulai ketika pengguna mengakses aplikasi	
<b>Actors</b>	Pengguna aplikasi	
<b>Pre Condition</b>	Pengguna telah berada pada halaman utama	
<b>Post Condition</b>	Grafik <i>scatter plot</i> 3D ditampilkan	
<b>Typical Course Event</b>	<b>Actor</b>	<b>System</b>
	1. Mengklik ikon <i>slide bar</i>  3. Mengklik 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	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
<b>Alternate Flow Event</b>	Langkah 5 : jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan grafik semua segmen secara default	

<b>Exceptional Flow Event</b>	Aplikasi tertutup karena <i>error</i>
-------------------------------	---------------------------------------

Tabel B.2 Use Case Melihat Grafik Box-Plot

<b>UC-3 Melihat Persebaran Pada Grafik Box-Plot</b>		
<b>Purpose</b>	Mengetahui persebaran nilai variabel CLV, <i>length</i> , <i>recency</i> , <i>frequency</i> , dan <i>monetary</i> dalam grafik <i>box-plot</i>	
<b>Overview</b>	Dimulai ketika pengguna mengakses aplikasi	
<b>Actors</b>	Pengguna aplikasi	
<b>Pre Condition</b>	Pengguna telah berada pada halaman utama	
<b>Post Condition</b>	Grafik <i>box-plot</i> ditampilkan	
<b>Typical Course Event</b>	<b>Actor</b>	<b>System</b>
	1. Mengklik ikon <i>slide bar</i>  3. Mengklik menu “Visualisasi Segmen” 5. Mengklik salah satu pilihan segmen pada menu <i>dropdown</i> 7. Mengklik ikon  pada kotak dengan judul “CLV”, “Length”, “Recency”, “Frequency”, atau “Monetary”	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> CLV, <i>length</i> , <i>recency</i> , <i>frequency</i> , atau <i>monetary</i>
<b>Alternate Flow Event</b>	Langkah 5 : jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan grafik semua segmen secara default	
<b>Exceptional Flow Event</b>	Aplikasi tertutup karena <i>error</i>	

Tabel B.3 Use Case Melihat Rentang Nilai Segmen

<b>UC-4 Melihat rentang nilai segmen</b>	
<b>Purpose</b>	Mengetahui rentang nilai variabel LRFM pada setiap segmen
<b>Overview</b>	Dimulai ketika pengguna mengakses aplikasi
<b>Actors</b>	Pengguna aplikasi
<b>Pre Condition</b>	Pengguna telah berada pada halaman utama
<b>Post Condition</b>	Rentang nilai ditampilkan
<b>Typical Course Event</b>	<b>Actor</b>
	<b>System</b>
	<ol style="list-style-type: none"> <li>1. Mengklik ikon  <i>slide bar</i></li> <li>3. Mengklik menu “Visualisasi Segmen”</li> <li>5. Mengklik salah satu pilihan segmen pada menu <i>dropdown</i></li> <li>7. Mengklik ikon  pada kotak dengan judul “Rentang Segmen”</li> </ol>
	<ol style="list-style-type: none"> <li>2. Menampilkan pilihan menu pada <i>slide bar</i></li> <li>4. Menampilkan halaman “Visualisasi Segmen”</li> <li>6. Menampilkan visualisasi sesuai pilihan</li> <li>8. Menampilkan rentang nilai keempat variabel pada segmen</li> </ol>
<b>Alternate Flow Event</b>	Langkah 5 : jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan pilihan semua segmen secara default
<b>Exceptional Flow Event</b>	Aplikasi tertutup karena <i>error</i>

B-4

*Halaman ini sengaja dikosongkan*

## LAMPIRAN C

*Tabel C.1 Pemodelan RFM Keseluruhan Pelanggan*

<b>Customer</b>	<b>L</b>	<b>R</b>	<b>F</b>	<b>M</b>
2 PUTRI	52	2	5	1140490
27 WANDA, TK	49	8	7	922455
371	56	6	9	2258208
3F	49	7	7	976618
88	56	6	9	594030
9, TK	50	6	8	207538
AA JAYA	49	5	6	763592
AA MAKMUR	56	1	8	1152983
AAN , BU	56	3	15	4377325
ABADI	49	5	7	425820
ABADI JAYA II	56	3	15	3606031
ABC	57	2	17	6997048
ABD HADI	58	1	27	6567071
ABD KAMAL	49	6	6	187920
ABD KARIM , HAJI	49	3	6	1652731
ABD RAHMAN	49	7	7	302563
ABD ROHIM/MU ADI	55	1	8	9254238
ABD. JALIL	50	6	7	946295
ABDUL HALIM	56	6	8	253235
ABDUL HANAN	56	1	9	3409433
ABDUL MUIS	50	1	17	5935943
ABDULAH / POJOK , HAJI	44	3	6	1403266
ABDULLAH , HAJI	56	2	9	277440
ABDUS SAHAM	42	5	6	2553390
ABI	47	7	5	338365
ABI QODIR	52	5	9	385955
ABIL, TK	54	5	8	13163642
ABNA	42	9	4	2609195
ADA	42	7	3	460100
ADAM BP	50	6	4	354090

## C-2

<b>Customer</b>	<b>L</b>	<b>R</b>	<b>F</b>	<b>M</b>
ADI	56	2	8	2580744
ADI MULYA	47	5	5	359750
AF	49	7	6	1959016
AFLAH	56	2	9	2953982
AGIL	42	2	7	999155
AGUN BP/HALIL BP	49	10	6	253189
AGUS	45	9	10	11540373
AGUS , BU	61	1	14	11559587
AGUS , HAJI	48	14	5	1193050
AGUSTIN	56	6	17	9880411
AHBAB	36	6	3	328330
AHMAD	59	3	17	3809370
AHMAD , BU	55	3	14	971390
AHMAD , HAJI	56	1	13	14264708
AHMAD HAJI	0	1	1	14040
AHMAD, BP	56	1	9	1438795
AHMADI	56	2	9	1483540
AIFA	23	3	5	1078128
AIR MANCUR	0	5	1	111370
AIR MANCUR / POJOK	56	2	10	43152365
AIR MANCUR II	46	3	4	1477650
AIR MANCUR, TK	28	29	2	37910
AISYAH	57	1	19	18069776
AISYAH, BU	42	8	7	352050
AIZAH	59	3	12	706754
AJAIB	49	3	7	1704852
AKBAR	49	6	8	3279970
AL - AMIN	56	6	7	758735
AL - HIDAYAH	51	6	12	1014368
AL HAMDULILLAH, TK	56	2	9	3786914
AL HIDAYAH	47	3	6	144016
AL HIKMAH	35	20	4	602660
AL-FAN	56	3	5	5499832
ALDI	56	1	10	1187772
ALEX	49	3	7	1142990
ALI, HAJI	42	8	7	4854497

<b>Customer</b>	<b>L</b>	<b>R</b>	<b>F</b>	<b>M</b>
ALIF	0	1	1	44208
ALIF BAROKAH	57	2	10	1108549
ALIFIA	49	6	7	1099390
ALIMUDIN	46	1	4	140760
ALVIN	56	6	9	1556856
AM 89	35	15	5	457485
AMANAHA	53	3	11	1952449
AMIN , BU	35	19	6	489844
AMIN BP	56	3	10	31335198
AMIN JAYA	42	3	3	27240
AMIN MADINAH	49	3	7	3689103
AMINAH	28	8	3	196300
AMINAH , BU	49	7	5	450120
AMINAH, BU	56	3	8	336508
AMIRUL BP	49	7	6	708568
AMPERA/FAUSI	56	2	8	1633879
AN , BU	14	24	2	50895
AN, BU	49	13	7	181740
ANA	18	2	3	326179
ANAM	56	1	9	4348257
ANAS , BU	53	1	3	175890
ANDALUS	49	5	7	5962796
ANDIKA	56	1	8	725180
ANDIN	42	9	3	87654
ANDRI	56	3	6	750270
ANDRI BP	49	6	8	2542666
ANGGA	49	10	6	561110
ANI	56	3	15	9332608
ANI , BU	56	3	8	189040
ANIDA	53	9	11	1527871
ANIK	49	5	7	1996261
ANIK , BU	51	3	12	647235
ANIS , BU	48	1	8	573922
ANIS, TK	49	9	5	520848
ANISA, TK	49	7	4	2579440
ANUGERAH	59	3	13	695855
ANUGRAH	0	8	1	4620
ANUGRAH JAYA	56	1	8	2527744

*Halaman ini sengaja dikosongkan*

## LAMPIRAN D

*Tabel 0.1 Hasil Keseluruhan Normalisasi Variabel RFM*

<b>Customer</b>	<b>Lnorm</b>	<b>Rbalik</b>	<b>Fnorm</b>	<b>Mnorm</b>
2 PUTRI	0.852459	0.982759	0.125	0.006363
27 WANDA, TK	0.803279	0.87931	0.1875	0.005141
371	0.918033	0.913793	0.25	0.012624
3F	0.803279	0.896552	0.1875	0.005445
88	0.918033	0.913793	0.25	0.003302
9, TK	0.819672	0.913793	0.21875	0.001137
AA JAYA	0.803279	0.931034	0.15625	0.004251
AA MAKMUR	0.918033	1	0.21875	0.006433
AAN , BU	0.918033	0.965517	0.4375	0.024494
ABADI	0.803279	0.931034	0.1875	0.002359
ABADI JAYA II	0.918033	0.965517	0.4375	0.020174
ABC	0.934426	0.982759	0.5	0.039169
ABD HADI	0.95082	1	0.8125	0.03676
ABD KAMAL	0.803279	0.913793	0.15625	0.001027
ABD KARIM , HAJI	0.803279	0.965517	0.15625	0.009232
ABD RAHMAN	0.803279	0.896552	0.1875	0.001669
ABD ROHIM/MU ADI	0.901639	1	0.21875	0.051812
ABD. JALIL	0.819672	0.913793	0.1875	0.005275
ABDUL HALIM	0.918033	0.913793	0.21875	0.001393
ABDUL HANAN	0.918033	1	0.25	0.019072
ABDUL MUIS	0.819672	1	0.5	0.033225
ABDULAH / POJOK , HAJI	0.721311	0.965517	0.15625	0.007835
ABDULLAH , HAJI	0.918033	0.982759	0.25	0.001528
ABDUS SAHAM	0.688525	0.931034	0.15625	0.014277
ABI	0.770492	0.896552	0.125	0.001869
ABI QODIR	0.852459	0.931034	0.25	0.002136
ABIL, TK	0.885246	0.931034	0.21875	0.073711
ABNA	0.688525	0.862069	0.09375	0.01459
ADA	0.688525	0.896552	0.0625	0.002551
ADAM BP	0.819672	0.913793	0.09375	0.001958

Customer	Lnorm	Rbalik	Fnorm	Mnorm
ADI	0.918033	0.982759	0.21875	0.01443
ADI MULYA	0.770492	0.931034	0.125	0.001989
AF	0.803279	0.896552	0.15625	0.010948
AFLAH	0.918033	0.982759	0.25	0.016521
AGIL	0.688525	0.982759	0.1875	0.005571
AGUN BP/HALIL BP	0.803279	0.844828	0.15625	0.001392
AGUS	0.737705	0.862069	0.28125	0.064618
AGUS , BU	1	1	0.40625	0.064726
AGUS , HAJI	0.786885	0.775862	0.125	0.006657
AGUSTIN	0.918033	0.913793	0.5	0.05532
AHBAB	0.590164	0.913793	0.0625	0.001813
AHMAD	0.967213	0.965517	0.5	0.021313
AHMAD , BU	0.901639	0.965517	0.40625	0.005415
AHMAD , HAJI	0.918033	1	0.375	0.079879
AHMAD HAJI	0	1	0	5.28E-05
AHMAD, BP	0.918033	1	0.25	0.008034
AHMADI	0.918033	0.982759	0.25	0.008284
AIFA	0.377049	0.965517	0.125	0.006013
AIR MANCUR	0	0.931034	0	0.000598
AIR MANCUR / POJOK	0.918033	0.982759	0.28125	0.241695
AIR MANCUR II	0.754098	0.965517	0.09375	0.008251
AIR MANCUR, TK	0.459016	0.517241	0.03125	0.000186
AISYAH	0.934426	1	0.5625	0.101193
AISYAH, BU	0.688525	0.87931	0.1875	0.001946
AIZAH	0.967213	0.965517	0.34375	0.003933
AJAIB	0.803279	0.965517	0.1875	0.009524
AKBAR	0.803279	0.913793	0.21875	0.018347
AL - AMIN	0.918033	0.913793	0.1875	0.004224
AL - HIDAYAH	0.836066	0.913793	0.34375	0.005656
AL HAMDULILLAH, TK	0.918033	0.982759	0.25	0.021187
AL HIDAYAH	0.770492	0.965517	0.15625	0.000781
AL HIKMAH	0.57377	0.672414	0.09375	0.00335
AL-FAN	0.918033	0.965517	0.125	0.030782
ALDI	0.918033	1	0.28125	0.006628
ALEX	0.803279	0.965517	0.1875	0.006377
ALI, HAJI	0.688525	0.87931	0.1875	0.027167

Customer	Lnorm	Rbalik	Fnorm	Mnorm
ALIF	0	1	0	0.000222
ALIF BAROKAH	0.934426	0.982759	0.28125	0.006184
ALIFIA	0.803279	0.913793	0.1875	0.006132
ALIMUDIN	0.754098	1	0.09375	0.000763
ALVIN	0.918033	0.913793	0.25	0.008695
AM 89	0.57377	0.758621	0.125	0.002537
AMANAHA	0.868852	0.965517	0.3125	0.010911
AMIN , BU	0.57377	0.689655	0.15625	0.002718
AMIN BP	0.918033	0.965517	0.28125	0.1755
AMIN JAYA	0.688525	0.965517	0.0625	0.000127
AMIN MADINAH	0.803279	0.965517	0.1875	0.020639
AMINAH	0.459016	0.87931	0.0625	0.001074
AMINAH , BU	0.803279	0.896552	0.125	0.002495
AMINAH, BU	0.918033	0.965517	0.21875	0.001859
AMIRUL BP	0.803279	0.896552	0.15625	0.003943
AMPERA/FAUSI	0.918033	0.982759	0.21875	0.009126
AN , BU	0.229508	0.603448	0.03125	0.000259
AN, BU	0.803279	0.793103	0.1875	0.000992
ANA	0.295082	0.982759	0.0625	0.001801
ANAM	0.918033	1	0.25	0.024331
ANAS , BU	0.868852	1	0.0625	0.000959
ANDALUS	0.803279	0.931034	0.1875	0.033375
ANDIKA	0.918033	1	0.21875	0.004036
ANDIN	0.688525	0.862069	0.0625	0.000465
ANDRI	0.918033	0.965517	0.15625	0.004177
ANDRI BP	0.803279	0.913793	0.21875	0.014217
ANGGA	0.803279	0.844828	0.15625	0.003117
ANI	0.918033	0.965517	0.4375	0.052251
ANI , BU	0.918033	0.965517	0.21875	0.001033
ANIDA	0.868852	0.862069	0.3125	0.008533
ANIK	0.803279	0.931034	0.1875	0.011156
ANIK , BU	0.836066	0.965517	0.34375	0.0036
ANIS , BU	0.786885	1	0.21875	0.003189
ANIS, TK	0.803279	0.862069	0.125	0.002892
ANISA, TK	0.803279	0.896552	0.09375	0.014423
ANUGERAH	0.967213	0.965517	0.375	0.003872
ANUGRAH	0	0.87931	0	0
ANUGRAH JAYA	0.918033	1	0.21875	0.014133



*Halaman ini sengaja dikosongkan*

## LAMPIRAN E

*Tabel 0.1 Hasil Keseluruhan Perhitungan CLV dan Clustering*

Customer	Lbobot	Rbobot	Fbobot	Mbobot	CLV	Cluster
2 PUTRI	0.06223	0.04914	0.03488	0.00380	0.10572	2
27 WANDA, TK	0.05864	0.04397	0.05231	0.00307	0.09559	3
371	0.06702	0.04569	0.06975	0.00755	0.12081	3
3F	0.05864	0.04483	0.05231	0.00326	0.09712	3
88	0.06702	0.04569	0.06975	0.00197	0.12072	3
9, TK	0.05984	0.04569	0.06103	0.00068	0.10415	3
AA JAYA	0.05864	0.04655	0.04359	0.00254	0.09727	3
AA MAKMUR	0.06702	0.05000	0.06103	0.00385	0.12490	3
AAN , BU	0.06702	0.04828	0.12206	0.01465	0.16190	1
ABADI	0.05864	0.04655	0.05231	0.00141	0.10026	3
ABADI JAYA II	0.06702	0.04828	0.12206	0.01206	0.16178	1
ABC	0.06821	0.04914	0.13950	0.02342	0.18270	1
ABD HADI	0.06941	0.05000	0.22669	0.02198	0.30099	1
ABD KAMAL	0.05864	0.04569	0.04359	0.00061	0.09567	3
ABD KARIM , HAJI	0.05864	0.04828	0.04359	0.00552	0.10058	3
ABD RAHMAN	0.05864	0.04483	0.05231	0.00100	0.09710	3
ABD ROHIM/MU ADI	0.06582	0.05000	0.06103	0.03098	0.12430	3
ABD. JALIL	0.05984	0.04569	0.05231	0.00315	0.10062	3
ABDUL HALIM	0.06702	0.04569	0.06103	0.00083	0.11663	3
ABDUL HANAN	0.06702	0.05000	0.06975	0.01141	0.12918	3
ABDUL MUIS	0.05984	0.05000	0.13950	0.01987	0.16946	1
ABDULAH / POJOK , HAJI	0.05266	0.04828	0.04359	0.00469	0.09144	3
ABDULLAH , HAJI	0.06702	0.04914	0.06975	0.00091	0.12725	3
ABDUS SAHAM	0.05026	0.04655	0.04359	0.00854	0.08488	2
ABI	0.05625	0.04483	0.03488	0.00112	0.08789	2
ABI QODIR	0.06223	0.04655	0.06975	0.00128	0.11383	3
ABIL, TK	0.06462	0.04655	0.06103	0.04408	0.11715	3
ABNA	0.05026	0.04310	0.02616	0.00872	0.07434	2
ADA	0.05026	0.04483	0.01744	0.00153	0.07589	2
ADAM BP	0.05984	0.04569	0.02616	0.00117	0.09325	2
ADI	0.06702	0.04914	0.06103	0.00863	0.12329	3
ADI MULYA	0.05625	0.04655	0.03488	0.00119	0.09104	2
AF	0.05864	0.04483	0.04359	0.00655	0.09418	3
AFLAH	0.06702	0.04914	0.06975	0.00988	0.12741	3
AGIL	0.05026	0.04914	0.05231	0.00333	0.09272	3
AGUN BP/HALIL BP	0.05864	0.04224	0.04359	0.00083	0.08960	3
AGUS	0.05385	0.04310	0.07847	0.03864	0.10145	3
AGUS , BU	0.07300	0.05000	0.11334	0.03871	0.17155	1
AGUS , HAJI	0.05744	0.03879	0.03488	0.00398	0.07968	2
AGUSTIN	0.06702	0.04569	0.13950	0.03308	0.17485	1
AHBAB	0.04308	0.04569	0.01744	0.00108	0.06827	2
AHMAD	0.07061	0.04828	0.13950	0.01274	0.18492	1

Customer	Lbobot	Rbobot	Fbobot	Mbobot	CLV	Cluster
AHMAD , BU	0.06582	0.04828	0.11334	0.00324	0.15202	1
AHMAD , HAJI	0.06702	0.05000	0.10463	0.04777	0.15457	1
AHMAD HAJI	0.00000	0.05000	0.00000	0.00003	0.05000	2
AHMAD, BP	0.06702	0.05000	0.06975	0.00480	0.12900	3
AHMADI	0.06702	0.04914	0.06975	0.00495	0.12729	3
AIFA	0.02752	0.04828	0.03488	0.00360	0.06137	2
AIR MANCUR	0.00000	0.04655	0.00000	0.00036	0.04334	2
AIR MANCUR / POJOK	0.06702	0.04914	0.07847	0.14453	0.16682	1
AIR MANCUR II	0.05505	0.04828	0.02616	0.00493	0.09062	2
AIR MANCUR, TK	0.03351	0.02586	0.00872	0.00011	0.02903	2
AISYAH	0.06821	0.05000	0.15694	0.06051	0.20814	1
AISYAH, BU	0.05026	0.04397	0.05231	0.00116	0.08308	3
AIZAH	0.07061	0.04828	0.09591	0.00235	0.14788	3
AJAIB	0.05864	0.04828	0.05231	0.00570	0.10358	3
AKBAR	0.05864	0.04569	0.06103	0.01097	0.10241	3
AL - AMIN	0.06702	0.04569	0.05231	0.00253	0.11309	3
AL - HIDAYAH	0.06103	0.04569	0.09591	0.00338	0.12577	3
AL HAMDULILLAH, TK	0.06702	0.04914	0.06975	0.01267	0.12752	3
AL HIDAYAH	0.05625	0.04828	0.04359	0.00047	0.09676	3
AL HIKMAH	0.04189	0.03362	0.02616	0.00200	0.04910	2
AL-FAN	0.06702	0.04828	0.03488	0.01841	0.11306	3
ALDI	0.06702	0.05000	0.07847	0.00396	0.13362	3
ALEX	0.05864	0.04828	0.05231	0.00381	0.10355	3
ALL, HAJI	0.05026	0.04397	0.05231	0.01625	0.08352	3
ALIF	0.00000	0.05000	0.00000	0.00013	0.05000	2
ALIF BAROKAH	0.06821	0.04914	0.07847	0.00370	0.13412	3
ALIFIA	0.05864	0.04569	0.05231	0.00367	0.09869	3
ALIMUDIN	0.05505	0.05000	0.02616	0.00046	0.09396	2
ALVIN	0.06702	0.04569	0.06975	0.00520	0.12076	3
AM 89	0.04189	0.03793	0.03488	0.00152	0.05717	2
AMANAHA	0.06343	0.04828	0.08719	0.00652	0.12904	3
AMIN , BU	0.04189	0.03448	0.04359	0.00163	0.05463	2
AMIN BP	0.06702	0.04828	0.07847	0.10495	0.14862	1
AMIN JAYA	0.05026	0.04828	0.01744	0.00008	0.08231	2
AMIN MADINAH	0.05864	0.04828	0.05231	0.01234	0.10378	3
AMINAH	0.03351	0.04397	0.01744	0.00064	0.05513	2
AMINAH , BU	0.05864	0.04483	0.03488	0.00149	0.09166	2
AMINAH, BU	0.06702	0.04828	0.06103	0.00111	0.12149	3
AMIRUL BP	0.05864	0.04483	0.04359	0.00236	0.09411	3
AMPERA/FAUSI	0.06702	0.04914	0.06103	0.00546	0.12321	3
AN , BU	0.01675	0.03017	0.00872	0.00016	0.02233	2
AN, BU	0.05864	0.03966	0.05231	0.00059	0.08836	3
ANA	0.02154	0.04914	0.01744	0.00108	0.05574	2
ANAM	0.06702	0.05000	0.06975	0.01455	0.12931	3
ANAS , BU	0.06343	0.05000	0.01744	0.00057	0.10620	2
ANDALUS	0.05864	0.04655	0.05231	0.01996	0.10092	3
ANDIKA	0.06702	0.05000	0.06103	0.00241	0.12488	3

## E-4

Customer	Lbobot	Rbobot	Fbobot	Mbobot	CLV	Cluster
ANDIN	0.05026	0.04310	0.01744	0.00028	0.07285	2
ANDRI	0.06702	0.04828	0.04359	0.00250	0.11496	3
ANDRI BP	0.05864	0.04569	0.06103	0.00850	0.10233	3
ANGGA	0.05864	0.04224	0.04359	0.00186	0.08961	3
ANI	0.06702	0.04828	0.12206	0.03125	0.16317	1
ANI , BU	0.06702	0.04828	0.06103	0.00062	0.12149	3
ANIDA	0.06343	0.04310	0.08719	0.00510	0.11956	3
ANIK	0.05864	0.04655	0.05231	0.00667	0.10033	3
ANIK , BU	0.06103	0.04828	0.09591	0.00215	0.13061	3
ANIS , BU	0.05744	0.05000	0.06103	0.00191	0.10856	3
ANIS, TK	0.05864	0.04310	0.03488	0.00173	0.08863	2
ANISA, TK	0.05864	0.04483	0.02616	0.00862	0.08987	2
ANUGERAH	0.07061	0.04828	0.10463	0.00232	0.15415	1
ANUGRAH	0.00000	0.04397	0.00000	0.00000	0.03866	2
ANUGRAH JAYA	0.06702	0.05000	0.06103	0.00845	0.12499	3
ANWAR , HAJI	0.05864	0.04483	0.05231	0.00491	0.09714	3
ANWARI	0.07180	0.04914	0.08719	0.00139	0.14617	3
APOLO	0.05864	0.04483	0.05231	0.00741	0.09719	3
APOTIK RAJAWALI	0.02393	0.04914	0.02616	0.01145	0.05881	2
ARIEF	0.06223	0.04569	0.06103	0.05551	0.11330	3
ARIFAH CELL	0.07300	0.05000	0.11334	0.02209	0.16986	1
ARIFIN	0.06941	0.05000	0.26156	0.04363	0.36439	1
ARININGSIH	0.05864	0.04483	0.05231	0.00151	0.09711	3
ARIP BP	0.04189	0.04483	0.02616	0.00771	0.06677	2
AS , BU	0.05864	0.04655	0.04359	0.00190	0.09726	3
ASAD	0.06702	0.05000	0.06975	0.00253	0.12897	3
ASBAH	0.06582	0.05000	0.07847	0.00190	0.13142	3
ASDO	0.05026	0.04828	0.01744	0.00009	0.08231	2
ASHARI	0.05864	0.04483	0.05231	0.04617	0.10067	3
ASHARI , HAJI	0.06702	0.04828	0.04359	0.00062	0.11495	3
ASIA	0.05864	0.04828	0.04359	0.00405	0.10055	3
ASIA ABADI	0.06462	0.05000	0.06103	0.00232	0.12057	3
ASIFA'	0.03351	0.03362	0.03488	0.01032	0.04253	2
ASILVA	0.06702	0.04914	0.06103	0.00164	0.12317	3
ASMAWADI , BU	0.04189	0.04483	0.03488	0.00428	0.06861	2
ASMI'ATI	0.06702	0.04914	0.05231	0.00363	0.11964	3
ASORI , BU	0.03351	0.04397	0.02616	0.00036	0.05649	2
ATIM	0.05864	0.04483	0.06103	0.01322	0.10094	3
ATINA, BU	0.06702	0.04914	0.06103	0.00115	0.12317	3
ATNAZI	0.06462	0.04655	0.07847	0.03694	0.12490	3
AVI, TK	0.06941	0.05000	0.06975	0.00142	0.13344	3
AZIZ , BU	0.05864	0.04483	0.05231	0.00433	0.09713	3
AZIZAH HAJAH	0.06702	0.04569	0.05231	0.01213	0.11333	3
AZIZAH, TK	0.06702	0.04914	0.06975	0.00250	0.12726	3
AZRIL	0.05864	0.04655	0.05231	0.00434	0.10029	3
BABU RAHMAN, TK	0.05026	0.03362	0.05231	0.00453	0.06706	3
BABURISQI, TK	0.06343	0.04655	0.04359	0.00122	0.10526	3
BADRIYAH , BU	0.00000	0.04569	0.00000	0.00000	0.04175	2

*Halaman ini sengaja dikosongkan*



## LAMPIRAN F

*Tabel D.1 Keseluruhan Hasil Percobaan Perubahan Nilai Parameter*

<i>Max Iter</i>	<i>Learning Rate</i>	<i>Weights</i>	<i>SSE Training</i>	<i>SSE Testing</i>
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

<i>Max Iter</i>	<i>Learning Rate</i>	<i>Weights</i>	<i>SSE Training</i>	<i>SSE Testing</i>
18	0,7	0,1	0,0017450201	0,0015782471
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

<i>Max Iter</i>	<i>Learning Rate</i>	<i>Weights</i>	<i>SSE Training</i>	<i>SSE Testing</i>
17	0,5	0,2	0,0017450209	0,0015786277
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



*Halaman ini sengaja dikosongkan*

## LAMPIRAN G

Tabel E.1 Keseluruhan Validasi Derajat Keanggotaan Pelanggan

Customer	X1	X2	X3	Max	Cluster	Validasi
2 PUTRI	0.0273	0.5865	0.3862	0.5865	2	Valid
27 WANDA, TK	0.0153	0.1155	0.8693	0.8693	3	Valid
371	0.0236	0.0397	0.9367	0.9367	3	Valid
3F	0.0145	0.1095	0.8759	0.8759	3	Valid
88	0.0285	0.0513	0.9202	0.9202	3	Valid
9, TK	0.0101	0.0384	0.9515	0.9515	3	Valid
AA JAYA	0.0245	0.3763	0.5992	0.5992	3	Valid
AA MAKMUR	0.0073	0.0220	0.9707	0.9707	3	Valid
AAN , BU	0.9030	0.0272	0.0698	0.9030	1	Valid
ABADI	0.0159	0.1193	0.8648	0.8648	3	Valid
ABADI JAYA II	0.8774	0.0343	0.0883	0.8774	1	Valid
ABC	0.9685	0.0101	0.0214	0.9685	1	Valid
ABD HADI	0.6362	0.1467	0.2172	0.6362	1	Valid
ABD KAMAL	0.0253	0.3909	0.5838	0.5838	3	Valid
ABD KARIM , HAJI	0.0242	0.3555	0.6203	0.6203	3	Valid
ABD RAHMAN	0.0170	0.1291	0.8539	0.8539	3	Valid
ABD ROHIM/MU ADI	0.0882	0.1697	0.7420	0.7420	3	Valid
ABD. JALIL	0.0134	0.0962	0.8904	0.8904	3	Valid
ABDUL HALIM	0.0117	0.0370	0.9513	0.9513	3	Valid
ABDUL HANAN	0.0286	0.0446	0.9268	0.9268	3	Valid
ABDUL MUIS	0.9440	0.0181	0.0379	0.9440	1	Valid
ABDULAH / POJOK , HAJI	0.0239	0.4859	0.4902	0.4902	3	Valid
ABDULLAH , HAJI	0.0308	0.0552	0.9140	0.9140	3	Valid
ABDUS SAHAM	0.0243	0.5335	0.4422	0.5335	2	Valid
ABI	0.0181	0.7443	0.2376	0.7443	2	Valid
ABI QODIR	0.0257	0.0508	0.9236	0.9236	3	Valid
ABIL, TK	0.1608	0.2448	0.5944	0.5944	3	Valid
ABNA	0.0050	0.9532	0.0419	0.9532	2	Valid
ADA	0.0090	0.9318	0.0593	0.9318	2	Valid
ADAM BP	0.0190	0.8042	0.1768	0.8042	2	Valid
ADI	0.0044	0.0126	0.9830	0.9830	3	Valid
ADI MULYA	0.0186	0.7354	0.2460	0.7354	2	Valid
AF	0.0240	0.3665	0.6095	0.6095	3	Valid
AFLAH	0.0254	0.0407	0.9339	0.9339	3	Valid

Customer	X1	X2	X3	Max	Cluster	Validasi
AGIL	0.0260	0.2389	0.7352	0.7352	3	Valid
AGUN BP/HALIL BP	0.0259	0.4076	0.5665	0.5665	3	Valid
AGUS	0.2605	0.1888	0.5507	0.5507	3	Valid
AGUS , BU	0.8795	0.0343	0.0861	0.8795	1	Valid
AGUS , HAJI	0.0202	0.7248	0.2550	0.7248	2	Valid
AGUSTIN	0.9758	0.0079	0.0163	0.9758	1	Valid
AHBAB	0.0067	0.9528	0.0406	0.9528	2	Valid
AHMAD	0.9128	0.0276	0.0596	0.9128	1	Valid
AHMAD , BU	0.6779	0.0828	0.2393	0.6779	1	Valid
AHMAD , HAJI	0.7284	0.0795	0.1921	0.7284	1	Valid
AHMAD HAJI	0.0822	0.6800	0.2377	0.6800	2	Valid
AHMAD, BP	0.0255	0.0436	0.9309	0.9309	3	Valid
AHMADI	0.0246	0.0421	0.9333	0.9333	3	Valid
AIFA	0.0269	0.8158	0.1573	0.8158	2	Valid
AIR MANCUR	0.0814	0.6830	0.2356	0.6830	2	Valid
AIR MANCUR / POJOK	0.3986	0.2730	0.3285	0.3986	1	Valid
AIR MANCUR II	0.0119	0.8797	0.1084	0.8797	2	Valid
AIR MANCUR, TK	0.0345	0.8201	0.1454	0.8201	2	Valid
AISYAH	0.8206	0.0658	0.1136	0.8206	1	Valid
AISYAH, BU	0.0275	0.2685	0.7039	0.7039	3	Valid
AIZAH	0.3589	0.1232	0.5179	0.5179	3	Valid
AJAIB	0.0124	0.0887	0.8989	0.8989	3	Valid
AKBAR	0.0057	0.0198	0.9746	0.9746	3	Valid
AL - AMIN	0.0158	0.0869	0.8973	0.8973	3	Valid
AL - HIDAYAH	0.3484	0.1283	0.5232	0.5232	3	Valid
AL HAMDULILLA H, TK	0.0305	0.0469	0.9226	0.9226	3	Valid
AL HIDAYAH	0.0253	0.4263	0.5484	0.5484	3	Valid
AL HIKMAH	0.0072	0.9444	0.0483	0.9444	2	Valid
AL-FAN	0.0430	0.4684	0.4886	0.4886	3	Valid
ALDI	0.0885	0.0883	0.8232	0.8232	3	Valid
ALEX	0.0135	0.0984	0.8881	0.8881	3	Valid
ALI, HAJI	0.0339	0.2555	0.7105	0.7105	3	Valid
ALIF	0.0822	0.6801	0.2377	0.6801	2	Valid
ALIF BAROKAH	0.0903	0.0893	0.8204	0.8204	3	Valid
ALIFIA	0.0138	0.1031	0.8831	0.8831	3	Valid
ALIMUDIN	0.0141	0.8599	0.1260	0.8599	2	Valid
ALVIN	0.0243	0.0420	0.9337	0.9337	3	Valid
AM 89	0.0095	0.9091	0.0814	0.9091	2	Valid

Customer	X1	X2	X3	Max	Cluster	Validasi
AMANAHA	0.2023	0.1163	0.6814	0.6814	3	Valid
AMIN , BU	0.0290	0.6884	0.2826	0.6884	2	Valid
AMIN BP	0.4077	0.2501	0.3421	0.4077	1	Valid
AMIN JAYA	0.0111	0.9152	0.0736	0.9152	2	Valid
AMIN MADINAH	0.0152	0.0980	0.8867	0.8867	3	Valid
AMINAH	0.0127	0.9210	0.0664	0.9210	2	Valid
AMINAH , BU	0.0213	0.6900	0.2887	0.6900	2	Valid
AMINAH, BU	0.0108	0.0337	0.9554	0.9554	3	Valid
AMIRUL BP	0.0246	0.3855	0.5899	0.5899	3	Valid
AMPERA/FAUSI	0.0050	0.0149	0.9801	0.9801	3	Valid
AN , BU	0.0510	0.7675	0.1816	0.7675	2	Valid
AN, BU	0.0226	0.1718	0.8057	0.8057	3	Valid
ANA	0.0333	0.8237	0.1430	0.8237	2	Valid
ANAM	0.0367	0.0545	0.9087	0.9087	3	Valid
ANAS , BU	0.0298	0.7620	0.2082	0.7620	2	Valid
ANDALUS	0.0311	0.1660	0.8030	0.8030	3	Valid
ANDIKA	0.0094	0.0287	0.9618	0.9618	3	Valid
ANDIN	0.0092	0.9303	0.0605	0.9303	2	Valid
ANDRI	0.0290	0.2727	0.6983	0.6983	3	Valid
ANDRI BP	0.0037	0.0134	0.9829	0.9829	3	Valid
ANGGA	0.0254	0.4024	0.5722	0.5722	3	Valid
ANI	0.9812	0.0055	0.0133	0.9812	1	Valid
ANI , BU	0.0119	0.0371	0.9510	0.9510	3	Valid
ANIDA	0.1998	0.1207	0.6795	0.6795	3	Valid
ANIK	0.0121	0.0870	0.9009	0.9009	3	Valid
ANIK , BU	0.3432	0.1297	0.5271	0.5271	3	Valid
ANIS , BU	0.0114	0.0442	0.9444	0.9444	3	Valid
ANIS, TK	0.0211	0.6960	0.2829	0.6960	2	Valid
ANISA, TK	0.0175	0.8203	0.1622	0.8203	2	Valid
ANUGERAH	0.5256	0.1077	0.3667	0.5256	1	Valid
ANUGRAH	0.0811	0.6843	0.2346	0.6843	2	Valid
ANUGRAH JAYA	0.0050	0.0144	0.9806	0.9806	3	Valid
ANWAR , HAJI	0.0133	0.0994	0.8873	0.8873	3	Valid
ANWARI	0.2090	0.1220	0.6691	0.6691	3	Valid
APOLO	0.0127	0.0918	0.8955	0.8955	3	Valid
APOTIK RAJAWALI	0.0309	0.8217	0.1473	0.8217	2	Valid

*Halaman ini sengaja dikosongkan*

## 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", i
con = icon("home")),
                  menuItem("Visualisasi Segmen", tabName = "de
tail", icon = icon("dashboard")),
                  menuItem("Tabel", tabName = "tabel", icon = i
con("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")
                                )
                            )
                        ),
                  )
                ),
                column(width = 6,
                    box(title = "Persentase Jum
lah Anggota per Segmen",
                        status = "info",
                        width = 50,
                        plotlyOutput("pie")
                    )
                )
              )
            ),
            )

```

```

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")
      )
    )
  )
),
),
),

```

*Keseluruhan Script UI.R*

```

library(shiny)
library(leaflet)
library(plot3Drgl)
library(sqldf)
library(shinydashboard)
library(plotly)
function(input, output, session) {

  #ALL
  alldata <- HasilRFMBobot
  data <- HasilKlaster
  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"))
}

```