

TUGAS AKHIR - KS 141501

**SEGMENTASI RETAILER OPERATOR
TELEKOMUNIKASI MENGGUNAKAN METODE
K-MEANS DAN MODEL LENGTH, RECENCY,
FREQUENCY, MONETARY (LRFM),
(STUDI KASUS: PT. XYZ).**

***RETAILER SEGMENTATION OF
TELECOMMUNICATION OPERATOR USING K-
MEANS METHOD AND LENGTH, RECENCY,
FREQUENCY, MONETARY (LRFM) MODEL,
(CASE STUDY: PT. XYZ)***

**SYAHRUL SEPTA PERDANA
NRP 052114 40000 170**

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

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



TUGAS AKHIR - KS 141501

**SEGMENTASI RETAILER OPERATOR
TELEKOMUNIKASI MENGGUNAKAN METODE
K-MEANS DAN MODEL LENGTH, REGENCY,
FREQUENCY, MONETARY (LRFM),
(STUDI KASUS: PT. XYZ).**

SYAHRUL SEPTA PERDANA
NRP 052114 40000 170

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

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



ITS
Institut
Teknologi
Sepuluh Nopember

FINAL PROJECT - KS 141501

***RETAILER SEGMENTATION OF
TELECOMMUNICATION OPERATOR USING K-
MEANS METHOD AND LENGTH, REGENCY,
FREQUENCY, MONETARY (LRFM) MODEL,
(CASE STUDY: PT. XYZ)***

SYAHRUL SEPTA PERDANA
NRP 052114 40000 170

Supervisors
Edwin Riksakomara, S.Kom., M.T.

INFORMATION SYSTEMS DEPARTMENT
Faculty of Information Technology and Communication
Institut Teknologi Sepuluh Nopember
Surabaya 2018

LEMBAR PENGESAHAN

SEGMENTASI RETAILER OPERATOR TELEKOMUNIKASI MENGGUNAKAN METODE K-MEANS DAN MODEL LENGTH, RECENCY, FREQUENCY, MONETARY (LRFM), (STUDI KASUS: PT. XYZ).

TUGAS AKHIR

Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

pada

Departemen Sistem Informasi
Fakultas Teknologi Informasi dan Komunikasi
Institut Teknologi Sepuluh Nopember

Oleh :

SYAHRUL SEPTA PERDANA

NRP 052114 40000 170

Surabaya, 16 Juli 2018

KETUA

DEPARTEMEN SISTEM INFORMASI



Ir. Aris Tjahyanto, M.Kom
196503101991021001

LEMBAR PERSETUJUAN

SEGMENTASI RETAILER OPERATOR TELEKOMUNIKASI MENGGUNAKAN METODE K-MEANS DAN MODEL LENGTH, RECENCY, FREQUENCY, MONETARY (LRFM), (STUDI KASUS: PT. XYZ).

TUGAS AKHIR

Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

pada

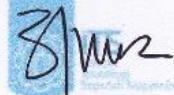
Departemen Sistem Informasi
Fakultas Teknologi Informasi dan Komunikasi
Institut Teknologi Sepuluh Nopember

Oleh :

SYAHRUL SEPTA PERDANA
NRP 052114 40000 170

Disetujui Tim Penguji Tanggal Ujian : 05 Juli 2018
Periode Wisuda : September 2018

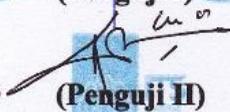
Edwin Riksakomara, S.Kom., M.T.


(Pembimbing I)

Ahmad Mukhlason, S.Kom., M.Sc., Ph.D


(Penguji I)

Faizal Mahananto, S.Kom., M.Eng., Ph.D


(Penguji II)

**SEGMENTASI RETAILER OPERATOR
TELEKOMUNIKASI MENGGUNAKAN METODE
K-MEANS DAN MODEL LENGTH, RECENCY,
FREQUENCY, MONETARY (LRFM),
(STUDI KASUS: PT. XYZ).**

Nama Mahasiswa : Syahrul Septa Perdana
NRP : 05214 40000 170
Departemen : Sistem Informasi, FTIK-ITS
Dosen Pembimbing : Edwin Riksakomara, S.Kom., M.T

ABSTRAK

Perkembangan industri di bidang telekomunikasi di Indonesia pada saat ini sangat berkembang pesat hal ini didukung oleh adanya penggunaan telepon seluler ataupun smartphone yang terus meningkat. Pertumbuhan penggunaan telepon seluler diikuti dengan pertumbuhan pengguna operator seluler mengakibatkan persaingan yang ketat antar operator seluler dan menghasilkan berbagai cara strategi pemasaran. PT XYZ memiliki beberapa program untuk menarik perhatian para pelanggan, namun program tersebut diimplementasikan kepada semua pelanggan tanpa adanya segmentasi sehingga program yang dijalankan tidak efektif dan efisien.

Untuk menyelesaikan permasalahan tersebut, salah satu solusi yang dapat digunakan adalah melakukan pendekatan pelanggan dengan cara menentukan segmentasi pelanggan. Segmentasi yang dilakukan menggunakan metode clustering K-Means dengan menerapkan model LRFM yang menggambarkan perilaku pelanggan dalam melakukan transaksi yang dikelompokkan sesuai segmen.

Hasil dari proses clustering dengan menggunakan metode K-Means dan model LRFM adalah tiga segmen retailer. Segmen 1 adalah segmen terburuk dengan rata-rata nilai CLV 0.12

dengan jumlah retailer pada segmen ini sebanyak 93 retailer. Segmen 2 adalah segmen terbaik dengan nilai rata-rata CLV 0.32 dengan jumlah 133 retailer dan segmen 3 merupakan segmen menengah dengan rata-rata CLV 0.19 dengan jumlah retailer terbanyak yakni 398 retailer. Penelitian ini juga menghasilkan visualisasi berbasis web yang menyajikan grafik yang mempermudah analisa hasil clustering.

Kata Kunci: Clustering, K-Means, Model LRFM, Segmentasi Pelanggan

**RETAILER SEGMENTATION OF
TELECOMMUNICATION OPERATOR USING K-
MEANS METHOD AND LENGTH, RECENCY,
FREQUENCY, MONETARY (LRFM) MODEL,
(CASE STUDY: PT. XYZ)**

Name : Syahrul Septa Perdana
NRP : 05214 40000 170
Department : Information System, FTIK-ITS
Supervisor : Edwin Riksakomara, S.Kom., M.T

ABSTRACT

The development of telecommunication industry in Indonesia at this time is very rapidly developed this is supported by the use of mobile phone or smartphone that continues to increase. The growth in cellular usage followed by the growth of mobile operator users resulted in intense competition among mobile operators and resulted in various ways of marketing strategy. PT XYZ has several programs to attract the attention of customers, but the program is implemented to all customers without any segmentation so that the program is run ineffective and efficient.

To solve the problem, one solution that can be used is to approach the customer by determining customer segmentation. Segmentation is done using K-Means clustering method by applying LRFM model that describes customer behavior in performing transaction which is grouped according to segment.

The result of clustering process using K-Means method and LRFM model are three segment of retailer. Segment 1 is the worst segment with an average CLV of 0.12 with the number of retailers in this segment of 93 retailers. Segment 2 is the best segment with an average CLV value of 0.32 with a total of 133 retailers and segment 3 is a segment with an average CLV of

0.19 with the largest number of retailers of 398 retailers. The research also produces web-based visualizations that provide clustering graphs.

Keywords: Clustering, K-Means, LRFM Model, Customer Segmentation

KATA PENGANTAR

Alhamdulillah penulis panjatkan rasa syukur kepada Allah SWT atas rahmat, nikmat, barakah, dan kemudahan yang telah diberikan sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul: **SEGMENTASI RETAILER OPERATOR TELEKOMUNIKASI MENGGUNAKAN METODE K-MEANS DAN MODEL *LENGTH, RECENCY, FREQUENCY, MONETARY* (LRFM), (STUDI KASUS: PT. XYZ)**, sebagai salah satu syarat kelulusan di Departemen Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi, Institut Teknologi Sepuluh Nopember, Surabaya.

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

1. Allah SWT yang telah menganugerahkan rahmat dan berkahnya 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. XYZ Gresik, 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 Rully Agus Hendrawan, S.Kom, M.Eng 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 pejuang wisuda 118 laboratorium RDIB dan Admin Laboratorium RDIB Mas Ricky yang telah memberikan bantuan dan semangat.
9. Teman-teman OSIRIS yang menjadi keluarga tempat berbagi susah dan senang selama penulis menempuh perkuliahan.
10. Formas Sampang yang telah menjadi keluarga serta rumah bagi penulis selama menjalani perantauan mencari ilmu.
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 | i |
| ABSTRACT | iii |
| KATA PENGANTAR | v |
| DAFTAR ISI | vii |
| DAFTAR GAMBAR | xi |
| DAFTAR TABEL | xiii |
| DAFTAR SCRIPT | xv |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 4 |
| 1.3 Batasan Masalah..... | 5 |
| 1.4 Tujuan Penelitian | 5 |
| 1.5 Manfaat Penelitian..... | 5 |
| 1.6 Relevansi | 6 |
| 1.7 Sistematika Penulisan..... | 6 |
| BAB II TINJAUAN PUSTAKA..... | 9 |
| 2.1 Penelitian Sebelumnya | 9 |
| 2.2 Dasar Teori..... | 12 |
| 2.2.1 PT. XYZ Cabang Gresik | 12 |
| 2.2.2 Customer Relationship Management | 13 |
| 2.2.3 Model LRFM | 15 |
| 2.2.4 Analytical Hierarchy Process | 16 |
| 2.2.5 Customer Lifetime Value (CLV) | 19 |
| 2.2.6 Normalisasi Min-Max | 20 |
| 2.2.7 Clustering | 21 |
| 2.2.8 K-Means..... | 22 |
| 2.2.9 Metode Elbow | 23 |
| 2.2.10 Performa Cluster | 24 |
| BAB III METODOLOGI PENELITIAN..... | 25 |
| 3.1 Tahapan Pelaksanaan Tugas Akhir | 25 |
| 3.2 Uraian Metodologi..... | 26 |
| 3.2.1 Identifikasi Masalah | 26 |
| 3.2.2 Studi Literatur | 26 |
| 3.2.3 Pengumpulan Data | 26 |

| | | |
|---------------------------------|---|-----------|
| 3.2.4 | Praproses Data..... | 27 |
| 3.2.5 | Pembobotan LRFM..... | 29 |
| 3.2.6 | Perhitungan Nilai CLV..... | 29 |
| 3.2.7 | Clustering Data..... | 29 |
| 3.2.8 | Denomarlisasi Data..... | 30 |
| 3.2.9 | Visualisasi <i>Cluster</i> | 30 |
| 3.2.10 | Analisa Hasil Cluster..... | 31 |
| 3.2.11 | Penarikan Kesimpulan dan Saran..... | 31 |
| 3.2.12 | Pembuatan Laporan Tugas Akhir..... | 31 |
| BAB IV PERANCANGAN | | 33 |
| 4.1 | Pengumpulan Data | 33 |
| 4.2 | Praproses Data..... | 34 |
| 4.3 | Pemodelan LRFM..... | 34 |
| 4.4 | Normalisasi Data..... | 35 |
| 4.5 | Pembobotan LRFM..... | 35 |
| 4.6 | Penentuan Jumlah Cluster..... | 35 |
| 4.7 | Rancangan Visualisasi | 36 |
| 4.7.1 | Use Case Diagram..... | 36 |
| 4.7.2 | Deskripsi Use Case..... | 36 |
| 4.7.3 | Desain Antarmuka Aplikasi | 38 |
| BAB V IMPLEMENTASI | | 41 |
| 5.1 | Praproses Data..... | 41 |
| 5.1.1 | Seleksi Data..... | 41 |
| 5.1.2 | Data <i>Cleaning</i> | 42 |
| 5.2 | Pemodelan LRFM..... | 42 |
| 5.3 | Normalisasi Data..... | 44 |
| 5.4 | Pembobotan LRFM..... | 46 |
| 5.5 | Perhitungan CLV | 46 |
| 5.6 | Clustering Data | 47 |
| 5.6.1 | Menentukan Jumlah K dengan Menggunakan Metode Elbow | 47 |
| 5.6.2 | Proses Clustering Menggunakan Metode K-Means | 48 |
| 5.6.3 | Uji Performa Cluster | 49 |
| 5.7 | Denormalisasi Data..... | 50 |
| 5.8 | Pembuatan Halaman Visualisasi | 51 |
| 5.8.1 | Halaman Utama Aplikasi | 51 |

| | | |
|------------------------|--|------------|
| 5.8.2 | Halaman Visualisi Segmen | 53 |
| 5.8.3 | Halaman Tabel Data..... | 63 |
| BAB VI | HASIL DAN PEMBAHASAN | 65 |
| 6.1 | Pembobotan LRFM dengan AHP | 65 |
| 6.2 | Perhitungan Nilai CLV | 68 |
| 6.3 | Clustering Data | 68 |
| 6.3.1 | Menentukan Nilai K Menggunakan Metode Elbow. | 69 |
| 6.3.2 | Proses Clustering Menggunakan Metode K-Means. | 70 |
| 6.3.3 | Uji Performa Cluster | 70 |
| 6.4 | Analisa Hasil Clustering | 71 |
| 6.4.1 | Pemeringkatan Segmen | 71 |
| 6.4.2 | Analisa Karakteristik Segmen 1 | 72 |
| 6.4.3 | Analisa Karakteristik Segmen 2..... | 73 |
| 6.4.4 | Analisa Karakteristik Segmen 3..... | 74 |
| 6.4.5 | Perbandingan Karakteristik Segmen | 75 |
| 6.5 | Analisa Hasil Visualisasi | 76 |
| 6.5.1 | Bar Chart dan Pie Chart | 77 |
| 6.5.2 | Scatter 3D Plot | 85 |
| 6.5.3 | Box and Whisker..... | 88 |
| 6.5.4 | Scatter 2D Plot | 93 |
| 6.6 | Kesimpulan Hasil Analisa..... | 98 |
| 6.7 | Rekomendasi Strategi Marketing..... | 100 |
| 6.8 | Perbandingan Nilai K..... | 102 |
| 6.8.1 | Pembobotan dengan Nilai K=3 | 102 |
| 6.8.2 | Tanpa Pembobotan dengan Nilai K=4 | 103 |
| 6.8.3 | Pembobotan dengan Nilai K=4 | 103 |
| 6.8.4 | Tanpa Pembobotan dengan Nilai K=5 | 104 |
| 6.8.5 | Pembobotan dengan Nilai K=5 | 105 |
| BAB VII | KESIMPULAN DAN SARAN | 107 |
| 7.1 | Kesimpulan | 107 |
| 7.2 | Saran | 108 |
| DAFTAR PUSTAKA | | 109 |
| BIODATA PENULIS | | 113 |
| LAMPIRAN A | | A-1 |
| LAMPIRAN B | | B-1 |

| | |
|------------------|-----|
| LAMPIRAN C | C-1 |
| LAMPIRAN D | D-1 |
| LAMPIRAN E..... | E-1 |
| LAMPIRAN F..... | F-1 |

DAFTAR GAMBAR

| | |
|--|----|
| Gambar 2.1 Metode Elbow | 23 |
| Gambar 3.1 Metodologi Penelitian | 25 |
| Gambar 4.1 Use Case Diagram | 37 |
| Gambar 4.2 Halaman Utama Aplikasi | 39 |
| Gambar 4.3 Halaman Visualisasi Segmen | 40 |
| Gambar 4.4 Halaman Tabel Data | 40 |
| Gambar 5.1 Hasil Visual Nilai Rataan LRFM | 55 |
| Gambar 5.2 Hasil Visual Rentang Nilai LRFM | 56 |
| Gambar 5.3 Hasil Visual Retailer Teratas & Terbawah | 58 |
| Gambar 5.4 Hasil Visual Tabel Peringkat Retailer | 59 |
| Gambar 5.5 Hasil Visual Box-Plot | 59 |
| Gambar 5.6 Hasil Visual Scatter Plot 3D | 61 |
| Gambar 5.7 Hasil Visual Halaman Tabel Data | 63 |
| Gambar 6.1 Metode Elbow | 69 |
| Gambar 6.2 Grafik Bar Chart dan Pie Chart | 77 |
| Gambar 6.3 Persebaran Jumlah Retailer | 78 |
| Gambar 6.4 Persentase Persebaran Retailer | 79 |
| Gambar 6.5 Presentase Persebaran Retailer Area Bangkalan | 81 |
| Gambar 6.6 Persentase Persebaran Retailer Area Gresik | 82 |
| Gambar 6.7 Persentase Persebaran Retailer Area Lamongan | 82 |
| Gambar 6.8 Persentase Persebaran Retailer Area Pamekasan | 83 |
| Gambar 6.9 Persentase Persebaran Retailer Area Sampang | 83 |
| Gambar 6.10 Persentase Persebaran Retailer Area Sumenep | 84 |
| Gambar 6.11 Grafik Scatter 3D Plot Keseluruhan Segmen | 85 |
| Gambar 6.12 Grafik Scatter 3D Plot Segmen 1 | 86 |
| Gambar 6.13 Grafik Scatter 3D Plot Segmen 2 | 87 |
| Gambar 6.14 Grafik Scatter 3D Plot Segmen 3 | 88 |
| Gambar 6.15 Grafik Box Plot Variabel Length | 89 |
| Gambar 6.16 Grafik Box Plot Variabel Recency | 90 |
| Gambar 6.17 Grafik Box Plot Variabel Frequency | 92 |
| Gambar 6.18 Grafik Box Plot Variabel Monetary | 93 |
| Gambar 6.19 Scatter 2D Plot Variabel Length, Recency | 94 |
| Gambar 6.20 Scatter 2D Plot Variabel Length, Frequency | 95 |
| Gambar 6.21 Scatter 2D Plot Variabel Length, Monetary | 96 |

| | |
|--|-----|
| Gambar 6.22 Scatter 2D Plot Variabel Rencency, Frequency | 96 |
| Gambar 6.23 Scatter 2D Plot Variabel Recency, Monetary | 97 |
| Gambar 6.24 Scatter 2D Plot Variabel Frequency, Monetary | 98 |
| Gambar 6.25 Hasil Segmentasi K = 3 Bobot | 102 |
| Gambar 6.26 Segmen 2, K=3 Bobot | 102 |
| Gambar 6.27 Hasil Segmentasi K=4 Tanpa Bobot | 103 |
| Gambar 6.28 Hasil Segmentasi K=4 Bobot | 103 |
| Gambar 6.29 Segmen 3, K=4 Bobot | 104 |
| Gambar 6.30 Hasil Segmentasi K=5 Tanpa Bobot | 104 |
| Gambar 6.31 Hasil Segmentasi K=5 Bobot | 105 |
| Gambar 6.32 Segmen 5, K=5 Bobot | 105 |

DAFTAR TABEL

| | |
|--|----|
| Tabel 2.1 Penelitian Sebelumnya 1 | 9 |
| Tabel 2.2 Penelitian Sebelumnya 2 | 10 |
| Tabel 2.3 Penelitian Sebelumnya 3 | 10 |
| Tabel 2.4 Penelitian Sebelumnya 4 | 11 |
| Tabel 2.5 Skala Preferensi AHP..... | 17 |
| Tabel 2.6 Tabel Random Index | 19 |
| Tabel 2.7 Algoritma K-Means | 22 |
| Tabel 4.1 Data Mentah Tabel Trasanksi | 33 |
| Tabel 4.2 Data Mentah Tabel Retailer | 34 |
| Tabel 4.3 Use case melihat grafik bar chart dan pie chart..... | 38 |
| Tabel 5.1 Atribut Terpilih | 41 |
| Tabel 5.2 Baris Data Tidak Valid..... | 42 |
| Tabel 5.3 Sebagian Hasil Pemodelan LRFM | 44 |
| Tabel 5.4 Rentang Nilai Atribut LRFM | 45 |
| Tabel 6.1 Hasil Kuesioner Responden 1 | 65 |
| Tabel 6.2 Hasil Kuesioner Responden 2 | 65 |
| Tabel 6.3 Hasil Kuesioner Responden 3 | 65 |
| Tabel 6.4 Rata-rata Pengisian Kuesioner | 66 |
| Tabel 6.5 Normalitation Comparisons Matrix | 66 |
| Tabel 6.6 Perkalian Matrix Uji Konsistensi AHP | 67 |
| Tabel 6.7 Sebagian Hasil Perhitungan CLV Retailer..... | 68 |
| Tabel 6.8 Rasio Penurunan SSE..... | 69 |
| Tabel 6.9 Sebagian Hasil Clustering | 70 |
| Tabel 6.10 Hasil Uji SSE | 71 |
| Tabel 6.11 Peringkat Segmen Berdasarkan Nilai CLV | 72 |
| Tabel 6.12 Karakteristik Retailer Segmen 1 | 73 |
| Tabel 6.13 Karakteristik Retailer Segmen 2 | 74 |
| Tabel 6.14 Karakteristik Retailer Segmen 3 | 75 |
| Tabel 6.15 Perbandingan Karakteristik 3 Segmen | 76 |
| Tabel 6.16 Persebaran Retailer..... | 84 |
| Tabel 6.17 Range Variabel Length | 88 |
| Tabel 6.18 Range Variabel Recency | 90 |
| Tabel 6.19 Range Variabel Frequency | 91 |
| Tabel 6.20 Range Variabel Monetary | 92 |

Tabel 6.21 Usulan Strategi CRM101

Tabel B.1 Use Case Melihat Grafik Box PlotB-1

Tabel B.2 Use Case Melihat Grafik Scatter 3D PlotB-2

Tabel B.3 Use Case Melihat Retailer Teratas & Terbawah .B-3

Tabel B.4 Use Case Melihat Rentang Nilai Segmen.....B-4

Tabel B.5 Use Case Melihat Rataan Variabel LRFMB-5

Tabel B.6 Use Case Melihat Ranking Retailer Setiap Segmen
.....B-6

Tabel B.7 Use Case Melihat Tabel Data Segmen Retailer...B-7

DAFTAR SCRIPT

| | |
|---|-----|
| Script 5.1 Query Ekstraksi Data..... | 44 |
| Script 5.2 Normalisasi Data | 45 |
| Script 5.3 Menghitung Nilai CLV..... | 47 |
| Script 5.4 Metode Elbow..... | 48 |
| Script 5.5 Proses Clustering | 49 |
| Script 5.6 Uji Performa Cluster..... | 50 |
| Script 5.7 Denormalisasi Data..... | 51 |
| Script 5.8 Rancang Bangun Visual Bar Chart..... | 52 |
| Script 5.9 Rancang Bangun Visual Pie Chart..... | 53 |
| Script 5.10 Rancangan Bangun Visual Nilai Rataan LRFM.. | 55 |
| Script 5.11 Rancang Bangun Visual Rentang Nilai LRFM ... | 56 |
| Script 5.12 Rancang Bangun Visual Retailer Teratas & Terbawah..... | 57 |
| Script 5.13 Rancang Bangun Tabel Peringkat Retailer | 58 |
| Script 5.14 Rancang Bangun Visual Blox-Plot..... | 61 |
| Script 5.15 Rancang Bangun Visual Scatter Plot 3D | 63 |
| Script 5.16 Rancang Bangun Halaman Tabel Data..... | 63 |
| Script F.1 Keseluruhan Script UI.R | F-1 |

BAB I

PENDAHULUAN

Bab ini menjelaskan tentang gambaran secara umum tugas akhir yang disusun. Gambaran tersebut meliputi latar belakang masalah mengapa tugas akhir ini disusun, perumusan masalah yang akan diselesaikan, batasan tugas akhir, tujuan tugas akhir, serta manfaat yang ditimbulkan oleh tugas akhir yang disusun. Selain itu akan dijelaskan relevansi tugas akhir dengan bidang keilmuan serta pengaplikasiannya dalam kehidupan sehari-hari.

1.1 Latar Belakang

Perkembangan industri di bidang telekomunikasi di Indonesia pada saat ini sangat berkembang pesat, hal ini didukung oleh adanya penggunaan telepon seluler ataupun *smartphone* yang terus meningkat. Diketahui pada tahun 2016 penggunaan *smartphone* di Indonesia menyentuh angka penggunaan sebesar 25% total penduduk Indonesia atau sekitar 65 juta pengguna dengan prediksi penggunaan yang terus meningkat tiap tahunnya [1]. Dengan adanya peningkatan penggunaan telepon seluler ini membuat persaingan penyedia layanan telekomunikasi semakin ketat, hal ini tak lepas dari kebijakan pemerintah terkait Undang-undang No. 36 Tentang Telekomunikasi [2] yang menghapus dominasi pasar telekomunikasi seluler oleh perusahaan BUMN sehingga terbuka peluang kebebasan para operator penyedia layanan telekomunikasi masuk ke dalam pasar industri telekomunikasi seluler.

PT. XYZ merupakan cabang perusahaan yang bergerak di bidang industri telekomunikasi. PT. XYZ menangani penjualan dan pemasaran produk telekomunikasi seluler salah satu operator seluler di Indonesia di beberapa wilayah Jawa Timur. Berbagai produk yang dipasarkan diantaranya adalah penjualan saldo pulsa, kartu perdana telekomunikasi seluler, kartu paket

internet dan voucher pulsa. Pertumbuhan penggunaan telepon seluler diikuti dengan pertumbuhan pengguna operator seluler [3] mengakibatkan persaingan yang ketat antar operator seluler dan menghasilkan berbagai cara strategi pemasaran. Dalam menghadapi persaingan dan agar mampu mempertahankan eksistensinya di era pemasaran, PT. XYZ memberikan beberapa program promosi untuk menarik minat transaksi pelanggan, namun pemberian program tersebut sebagian diberlakukan pada setiap pelanggan dan hanya terdapat sebuah program yang dikhususkan terhadap pelanggan tertentu. Penerapan strategi yang sama terhadap sebagian pelanggan dapat menimbulkan beberapa dampak, yaitu strategi program yang tidak tepat sasaran sehingga strategi tersebut bermanfaat terhadap pelanggan tertentu saja. Selain itu, waktu dan dana yang digunakan untuk melaksanakan program promosi menjadi kurang efektif dan efisien.

Untuk menyelesaikan masalah pelanggan tersebut, perusahaan harus memfokuskan proses bisnis perusahaan terhadap pelanggan. CRM adalah salah satu strategi bisnis untuk meningkatkan keuntungan, penghasilan dan kepuasan pelanggan dengan cara mengatur segmentasi pelanggan [4]. Segmentasi pelanggan adalah proses menghasilkan segmen-segmen pelanggan berdasarkan karakteristik tertentu sebagai dasar pengelompokkan. Setiap pelanggan yang masuk ke dalam segmen tertentu, memiliki kemiripan karakteristik dengan pelanggan dalam satu segmen serta memiliki karakteristik yang berbeda dengan pelanggan pada segmen lainnya. Dari setiap segmen yang dihasilkan dapat diketahui perilaku pelanggan yang dominan karakteristiknya [5]. Dengan hasil segmentasi ini, pihak manajemen dapat menganalisis mengenai strategi marketing yang tepat terhadap setiap segmen berdasarkan kecenderungan perilaku pelanggan di dalamnya.

Banyak metode yang dapat digunakan dalam melakukan segmentasi pelanggan, salah satunya adalah penggalian data terhadap riwayat transaksi pembelian pelanggan pada periode

tertentu. Data transaksi tersebut kemudian ditentukan variabel-variabel yang berpengaruh dan dapat dijadikan kriteria untuk menentukan segmen pelanggan. Beberapa variabel yang umumnya digunakan antara lain :

- a. Interval transaksi awal dan akhir pelanggan
- b. Selisih jarak dari waktu transaksi terakhir kali dilakukan dengan waktu periode penelitian
- c. Total jumlah transaksi yang dilakukan selama periode tertentu
- d. Jumlah nilai uang dari transaksi yang dilakukan

Keempat variabel tersebut lazim disebut sebagai model LRFM (*Length, Recency, Frequency, dan Monetary*). Dalam model LRFM, *Length* mewakili interval transaksi awal hingga akhir, *Recency* mewakili waktu pembelian terakhir, *Frequency* mewakili jumlah transaksi, dan *Monetary* mewakili jumlah pemasukan yang diterima [6] [7]. Keempat variabel ini didapatkan berdasarkan transaksi yang dilakukan dari waktu ke waktu. Pola perilaku bisnis pelanggan juga dapat berubah sehingga dengan mengelompokkan berdasarkan dengan karakteristiknya, hasil pengelompokan menjadi lebih dinamis dan lebih sesuai dengan keadaan bisnis sesungguhnya.

Keempat variabel tersebut kemudian diproses dengan menggunakan algoritma penggalian data *clustering*. *Clustering* adalah sebuah metode yang digunakan untuk mengidentifikasi *cluster* yang terdapat dalam data, dimana *cluster* adalah sekumpulan objek data yang sama antara satu sama lain [7]. Metode *clustering* yang digunakan untuk melakukan segmentasi retailer pada PT. XYZ adalah K-Means *clustering*. K-Means *clustering* adalah teknik *clustering* berdasarkan partisi N objek pengamatan ke dalam K kelompok (*cluster*) dimana setiap objek pengamatan dimiliki oleh sebuah kelompok dengan rata-rata (*mean*) terdekat [8]. Metode ini digunakan karena hasil pengelompokan yang diinginkan anggota kelompoknya hanya benar-benar berada hanya ada di

kelompok tersebut. Selain itu, metode ini mudah diinterpretasikan, mudah diimplementasikan dan dapat beradaptasi pada data yang tersebar [9]. Namun, metode ini memiliki kekurangan dimana nilai k harus ditentukan terlebih dahulu untuk menghasilkan hasil *clustering* yang terbaik. Metode elbow dapat membantu pencarian nilai k yang optimal dengan membandingkan nilai k dengan nilai SSE atau presentase biaya pada nilai *cluster* yang telah ditentukan. Hasil presentase akan ditampilkan melalui grafik yang mempermudah untuk dipahami. Apabila nilai SSE atau presentase biaya pada grafik tersebut turun drastis dan membentuk siku, maka nilai tersebut yang digunakan sebagai nilai k [10]

Tugas akhir ini akan menghasilkan segmentasi pelanggan yakni para retailer PT. XYZ untuk menampilkan perilaku retailer berdasarkan analisis LRFM. Hasil segmentasi akan ditampilkan melalui visualisasi berbasis web dengan metode *k-means clustering*. Visualisasi dilakukan untuk memudahkan perusahaan untuk memahami hasil segmentasi pelanggan.

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 *Clustering K-Means* yang menampilkan perilaku bisnis pelanggan berdasarkan model LRFM?
2. Bagaimana performa *Clustering K-Means* dalam segmentasi pelanggan?
3. Bagaimana analisa hasil segmentasi yang dihasilkan?
4. Bagaimana visualisasi dapat mempermudah analisis segmentasi pelanggan?

1.3 Batasan Masalah

Dalam pengerjaan tugas akhir ini, penulis memberikan batasan-batasan masalah dalam usulan tugas akhir ini agar pembahasan tugas akhir terorganisir dengan baik. Berikut adalah batasan tugas akhir :

1. Data yang diolah dalam tugas akhir ini adalah data transaksi penjualan produk saldo pulsa (*Sell In*) Kantor Cabang PT. XYZ pada tanggal 1 Januari 2017 – 13 Desember 2017
2. Penelitian ini terbatas pada segmentasi pelanggan/retailer area Gresik-Madura yang mengikuti program K3 dan Non-CW
3. Penentuan variabel *clustering* menggunakan metode LRFM
4. Penentuan jumlah *cluster* menggunakan metode elbow
5. Penelitian ini menggunakan software R sebagai *tools clustering*
6. Visualisasi *cluster* berbasis web

1.4 Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan yang ingin dicapai diantaranya sebagai berikut:

1. Menghasilkan segmentasi retailer PT. XYZ area Gresik-Madura berdasarkan metode *clustering k-means* dan model LRFM
2. Mendapatkan hasil performa klustering menggunakan metode k-means
3. Mendapatkan hasil analisa terhadap segmentasi retailer PT. XYZ area Gresik-Madura
4. Visualisasi hasil *clustering* berbasis web

1.5 Manfaat Penelitian

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

1.6 Relevansi

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

1.7 Sistematika Penulisan

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

BAB I PENDAHULUAN

Bab ini memuat hal-hal yang melatarbelakangi pengerjaan tugas akhir ini yang meliputi latar belakang masalah, perumusan masalah, batasan masalah, perumusan masalah, tujuan, manfaat penelitian, relevansi dan yang terakhir adalah sistematika penulisan tugas akhir.

BAB II TINJAUAN PUSTAKA

Bab ini menjelaskan tentang studi sebelumnya yang mendukung tugas akhir serta teori-teori yang berhubungan dengan segmentasi pelanggan, *clustering*, LRFM model dan k-means yang akan digunakan untuk menyelesaikan permasalahan pada tugas akhir.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan mengenai alur pengerjaan tugas akhir yang berawal dari identifikasi masalah, studi literatur, pengumpulan data, praproses data, penentuan variabel LRFM, pembobotan LRFM, menentukan nilai CLV, *clustering* data, visualisasi data, penarikan kesimpulan dan saran, dan diakhiri dengan pembuatan laporan tugas akhir.

BAB IV PERANCANGAN

Pada bab ini akan dibahas mengenai perancangan pengolahan data dengan melakukan praproses data, analisis LRFM dan juga penentuan jumlah *clustering* (nilai k) yang akan digunakan pada K-means *clustering*.

BAB V IMPLEMENTASI

Pada bab ini tentang pengolahan data yang didapat dari tahap sebelumnya dan implementasi *clustering* dengan menggunakan R.

BAB VI HASIL DAN PEMBAHASAN

Pada bab ini akan ditampilkan hasil *clustering* dan uji performa hasil *clustering* berdasarkan metode K-Means yang diharapkan dapat menjadi penyelesaian permasalahan.

BAB VII PENUTUP

Pada bab ini akan dijelaskan kesimpulan dan saran untuk penelitian sebelumnya dari pengerjaan tugas akhir yang menjawab rumusan masalah yang telah dijabarkan pada bab Pendahuluan.

Halaman ini sengaja dikosongkan

BAB II TINJAUAN PUSTAKA

Bab ini berisi mengenai studi sebelumnya yang berhubungan dengan tugas akhir dan teori - teori yang berkaitan dengan permasalahan tugas akhir.

2.1 Penelitian Sebelumnya

Bagian ini akan menjelaskan penelitian-penelitian sebelumnya yang digunakan sebagai dasar penelitian tugas akhir. Hal-hal yang didapat dari penelitian sebelumnya dijabarkan pada tabel 2.1 – 2.4.

Tabel 2.1 Penelitian Sebelumnya 1

| | |
|----------------------------------|---|
| Judul Penelitian | Analyzing Patients' Values by Applying Cluster Analysis and LRFM Model in a Pediatric Dental Clinic in Taiwan [11] |
| Penulis (Tahun) | Hsin-Hung Wu, Shih-Yen Lin, Chih-Wei Liu (2014) |
| Deskripsi Umum Penelitian | Penelitian ini mengkombinasikan analisa kluster dan pemodelan LRFM untuk mengelompokkan nilai pasien. Diperoleh 12 kluster dari 1462 pasien dengan menggunakan metode k-means dan model LRFM dalam menganalisa nilai pelanggan yang kemudian juga diperoleh 3 kluster untuk loyal pasien dan 3 kluster <i>lost</i> pasien dalam menganalisa hubungan para pelanggan terhadap klinik gigi terkait. |
| Keterkaitan Penelitian | Penggunaan Model <i>LRFM</i> pada penelitian ini dapat dijadikan acuan referensi pada penelitian tugas akhir bahwa model <i>LRFM</i> dapat dijadikan sebuah model dalam membangun segmentasi pelanggan |

Tabel 2.2 Penelitian Sebelumnya 2

| | |
|----------------------------------|---|
| Judul Penelitian | Generating Customer Profiles for Retail Stores Using Clustering Techniques [12] |
| Penulis (Tahun) | Pramod Prasad (2011) |
| Deskripsi Umum Penelitian | Penelitian ini memanfaatkan data yang berlimpah yang didapatkan dalam industri retail. Data transaksi penjualan di beberapa retail di india dijadikan sumber data untuk dilakukan penggalian data dalam proses <i>clustering</i> untuk mengidentifikasi segmentasi pelanggan. Penelitian ini memberikan kesimpulan bahwa proses <i>clustering</i> dengan menggunakan metode <i>k-means</i> didapatkan 4 segmen pelanggan yang diharapkan dapat meningkatkan pemahaman terhadap perilaku pelanggan dan menjadikan hasil segmentasi dalam pembuatan strategi pemasaran serta meningkatkan pelayanan pelanggan |
| Keterkaitan Penelitian | Penelitian ini memberikan acuan proses menentukan pengelompokan jenis retail menggunakan teknik <i>clustering</i> dengan metode <i>k-means</i> yang dapat dijadikan referensi dalam penelitian tugas akhir |

Tabel 2.3 Penelitian Sebelumnya 3

| | |
|----------------------------------|---|
| Judul Penelitian | Proposing a Model for Customer Segmentation using WLRFM Analysis (Case Study: an ISP Company) [13] |
| Penulis (Tahun) | Babak Hazaveh Hesar Maskan (2014) |
| Deskripsi Umum Penelitian | Penelitian ini meneliti mengenai <i>Customer Lifetime Value</i> dengan melakukan segmentasi pelanggan |

| | |
|-------------------------------|---|
| | <p>untuk menghasilkan pelayanan yang lebih baik serta menentukan strategi di setiap segmen pelanggan. Dalam menentukan <i>CLV</i>, peneliti menggunakan model <i>LRFM</i> dan <i>AHP</i> dalam mengalokasi bobot di tiap-tiap variabel <i>LRFM</i> serta menggunakan metode <i>k-means</i> dalam proses <i>clustering</i>. Kesimpulan yang didapatkan dalam penelitian ini diperoleh lima jenis segmentasi yang dapat digunakan dalam menghitung <i>CLV</i> tiap segmen pelanggan, <i>WLRFM</i> dapat membantu dalam mengidentifikasi nilai tiap pelanggan serta dapat membantu dalam membangun strategi sesuai perilaku tiap segmen pelanggan dan <i>CLV</i> pelanggan</p> |
| Keterkaitan Penelitian | <p>Penelitian ini dijadikan referensi dalam pembuatan pembobotan variabel <i>LRFM</i> dengan menggunakan metode <i>Analytic Hierarchy Process (AHP)</i> serta penggunaan metode <i>k-means</i> dalam proses <i>clustering</i> yang diangkat dalam topik penelitian tugas akhir ini</p> |

Tabel 2.4 Penelitian Sebelumnya 4

| | |
|-------------------------|---|
| Judul Penelitian | K-Mean Clustering Method For Analysis Customer Lifetime Value With LRFM Relationship Model In Banking Services [14] |
| Penulis (Tahun) | Mohsen Alvandi, Safar Fazli, Farzaneh Seifi Abdoli |

| | |
|----------------------------------|--|
| Deskripsi Umum Penelitian | Penelitian ini mengidentifikasi nilai pelanggan terhadap perusahaan, dalam melakukan identifikasi, peneliti menggunakan metode k-means dan model LRFM dan diperoleh 9 jenis kluster pelanggan yang selanjutnya dilanjutkan dengan menentukan strategi yang tepat terhadap identifikasi kluster pelanggan yang diperoleh. |
| Keterkaitan Penelitian | Penelitian ini sebagai referensi penggunaan analisis model <i>LRFM</i> dan penggunaan metode <i>k-means</i> dalam proses <i>clustering</i> untuk menentukan strategi dalam manajemen hubungan pelanggan yang sesuai dengan topik tugas akhir ini |

2.2 Dasar Teori

Pada bagian ini akan dijelaskan mengenai teori-teori yang digunakan oleh penulis dalam penelitian tugas akhir.

2.2.1 PT. XYZ Cabang Gresik

PT. XYZ adalah perusahaan penyedia jasa telekomunikasi yang berkembang pesat dan beroperasi dengan lisensi nasional 2G/GSM 1800 MHz dan 3G/WCDMA di Indonesia serta pertama kali beroperasi secara komersial di Indonesia pada tanggal 30 Maret 2007. PT. XYZ menyediakan layanan internet bergerak yang berkualitas dan inovatif, serta layanan komunikasi telepon dan SMS yang terjangkau dan terus melakukan ekspansi cakupan layanan hingga ke seluruh wilayah negeri untuk menghadirkan pengalaman berinternet kelas dunia bagi Indonesia.

PT. XYZ Cabang Gresik merupakan rekanan perusahaan resmi yang bergerak di sektor jasa telekomunikasi seluler PT.XYZ yang bertugas dalam memasarkan produk dan melayani pendistribusian produk PT.XYZ di beberapa wilayah Jawa

Timur. Wilayah pemasaran dari kantor cabang PT.XYZ ini meliputi beberapa kawasan wilayah area diantaranya:

- Wilayah 1 Gresik – Madura (Gresik, Lamongan, Madura)
- Wilayah 2 Mojokerto – Jombang
- Wilayah 3 Bojonegoro – Tuban

PT. XYZ Cabang Gresik memasarkan berbagai produk dan jasa diantaranya

- SELL IN (Pengisian Saldo Pulsa Retailer)
- Kartu Perdana (Internet dan Pulsa)
- Voucher Pulsa
- Voucher Internet

Pada saat ini PT.XYZ memiliki 3 program untuk para retailer diantaranya

- Program CW yakni program dimana para retailer dapat mendapatkan keuntungan berupa diskon harga saldo pulsa dengan beberapa persyaratan diantaranya jumlah besaran transaksi yang ditentukan.
- Program K3 yakni program undian berhadiah dengan mengumpulkan koin-koin dari penjualan yang diundi di waktu tertentu
- Program Non-CW atau D3 yakni program yang diberikan kepada pelanggan dengan menukarkan poin transaksi dengan hadiah langsung

2.2.2 Customer Relationship Management

Customer Relationship Management (CRM) adalah sebuah strategi perusahaan untuk memahami dan mempengaruhi perilaku pelanggan melalui komunikasi intensif dengan tujuan untuk meningkatkan pendapatan, penyimpanan, loyalitas, dan profitabilitas dari pelanggan [15]. Proses CRM menyediakan struktur untuk menciptakan hubungan yang sukses dengan pelanggan.

CRM merupakan komitmen dari perusahaan untuk menempatkan pengalaman pelanggan pada pusat dari prioritas perusahaan dan untuk memastikan sistem bahwa yang insentif, proses, dan sumber daya informasi dapat mempengaruhi hubungan dengan menambah pengalaman [16]. Dalam konteks teknologi, CRM merupakan desain komunikasi dan penggunaan informasi untuk memastikan bahwa pelanggan tumbuh lebih yakin, percaya, dan rasa nilai pribadi dalam hubungannya dengan perusahaan.

Fokus utama dalam strategi CRM adalah untuk membuat organisasi menciptakan dan mempertahankan keuntungan pelanggan [15]. Sebagian besar strategi dikembangkan dari tiga aspek, yaitu *customer profitability*, *customer acquisition*, dan *customer retention*, berdasarkan pengurangan biaya mempertahankan pelanggan daripada memperoleh pelanggan baru [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 [18]. Dalam penggunaan teknologi, sejumlah aplikasi teknologi dapat diidentifikasi yang digunakan pada pengembangan strategi CRM. Tiga komponen utama komponen sistem CRM adalah [19]:

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 Model LRFM

Model LRFM merupakan perkembangan dari model RFM yang dimana model RFM tidak dapat membedakan pelanggan yang memiliki hubungan yang panjang atau pendek dengan perusahaan. Model RFM pertama kali diperkenalkan oleh Hughes dan saat ini banyak digunakan oleh industry termasuk manufaktur, *retailer*, dan jasa [9], dan model LRFM diperkenalkan oleh Chang dan Tsay yang dapat membedakan panjang atau pendeknya kerjasama pelanggan dan perusahaan [20]. Model ini menganalisa empat variabel yakni *length*, *recency*, *frequency*, dan *monetary* [20]. Adapun penjelasan dari keempat variabel tersebut adalah:

1. *Length (L)*

Interval transaksi awal dan akhir pelanggan yang dilambangkan dengan L atau *Length*. Semakin panjang interval maka semakin besar nilai L

2. *Recency (R)*

Selisih jarak dari waktu transaksi terakhir kali dilakukan dengan waktu periode penelitian, dilambangkan dengan R atau yang disebut dengan *recency*. Semakin pendek intervalnya, maka semakin besar nilai R.

3. *Frequency (F)*

F merepresentasikan frekuensi, yaitu jumlah transaksi dalam periode tertentu, contohnya dua kali dalam satu tahun atau dua kali dalam satu bulan. Semakin tinggi frekuensinya semakin besar nilai F.

4. *Monetary Value (M)*

M merepresentasikan *monetary*, yaitu nilai produk dalam bentuk uang pada periode tertentu. Semakin banyak jumlah uang pada periode tersebut maka semakin tinggi nilai M.

2.2.4 Analytical Hierarchy Process

Analytical Hierarchy Process (AHP) merupakan salah satu alat pendukung keputusan yang pertama kali diperkenalkan oleh Saaty pada 1977. AHP digunakan untuk mengembangkan nilai numerik untuk peringkat setiap alternatif keputusan, berdasarkan seberapa baik setiap alternatif memenuhi kriteria pembuat keputusan. Metode peringkat alternatif keputusan ini akan memilih yang terbaik ketika pembuat keputusan memiliki beberapa tujuan, atau kriteria yang menjadi dasar pada keputusan [23]. Berikut ini adalah tahapan penggunaan AHP [13].

1. Melakukan *Pairwise Comparison*

Dengan melakukan survey terlebih dahulu melalui kuesioner untuk memberikan nilai setiap kriteria dimana kriteria yang akan dinilai adalah kriteria dari variabel LRFM dengan mengacu pada skala preferensi tabel 2.5:

Tabel 2.5 Skala Preferensi AHP

| 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 disbanding kriteria pasangannya |
| 5 | Lebih Penting | Penilaian dan pengalaman memihak pada salah satu kriteria tertentu disbanding 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. Normalisasi Matrix

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. Perhitungan Bobot Relative

Menghitung bobot setiap kriteria dengan cara menghitung rata-rata angka setiap baris dalam matriks *normal comparison*.

$$W = \frac{X_1 + \dots + X_n}{n}; \text{ keterangan n: jumlah variabel}$$

4. Uji Rasio Konsistensi

Uji rasio konsistensi merupakan instrument yang menentukan konsistensi dan menunjukkan tingkat kepercayaan prioritas yang diperoleh dari perbandingan. Terdapat dua tahapan dalam uji rasio konsistensi, yaitu :

- a. Setiap elemen pada tabel yang berisi matriks *pairwise comparison* dikalikan dengan bobot dari setiap kriteria yang didapatkan pada langkah ke 3, lalu menghitung total skor dengan menjumlahkan setiap barisnya.
- b. Mencari nilai x sebagai masukan untuk menghitung CI dengan cara mengalikan total skor dengan bobot kemudian hasilnya dirata-rata.
- c. Menghitung indeks konsistensi dengan persamaan berikut :

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

Keterangan:

CI = Indeks Konsistensi

X = hasil dari langkah 4b

n = banyaknya kriteria

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

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

Keterangan:

CI = Indeks konsistensi

RI = Random indeks

Tabel 2.6 Tabel Random Index

| 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.

2.2.5 Customer Lifetime Value (CLV)

Customer Lifetime Value (CLV) adalah sebuah metric pemasaran yang menggambarkan nilai dari seorang pelanggan sepanjang hubungan pelanggan dengan perusahaan berlangsung [13]. CLV menentukan nilai pelanggan ke perusahaan selama siklus hidup pelanggan. CLV berusaha memaksimalkan keuntungan dengan menganalisis perilaku pelanggan dan siklus bisnis untuk mengidentifikasi dan menargetkan pelanggan dengan potensi nilai terbesar dari waktu ke waktu.

Dalam penelitian tugas akhir ini, penggunaan CLV mengacu pada indeks nilai CLV. Indeks CLV ditentukan melalui bobot variabel LRFM yang sebelumnya telah diperoleh dan nilai normalisasi masing-masing variabel

LRFM. Adapun formula untuk menghitung nilai CLV adalah sebagai berikut [15]:

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

Keterangan:

- CLV = nilai CLV yang dicari
- NL = nilai normalisasi variabel *length*
- WL = nilai bobot variabel *length*
- NR = nilai normalisasi variabel *recency*
- WR = nilai bobot variabel *recency*
- NF = nilai normalisasi variabel *frequency*
- WF = nilai bobot variabel *frequency*
- NM = nilai normalisasi variabel *monetary*
- WM = nilai bobot variabel *monetary*

2.2.6 Normalisasi Min-Max

Normalisasi merupakan proses yang digunakan untuk standarisasi semua variabel dataset dan memberikan bobot yang sama [26]. Normalisasi dapat mengeliminasi data yang redundan dan pencilan, sehingga data yang dihasilkan valid, dapat diandalkan dan meningkatkan akurasi hasil. Metode normalisasi yang digunakan tergantung dengan data yang akan dinormalisasi, pada penelitian ini metode yang digunakan adalah min-max.

Normalisasi min-max adalah proses perubahan data yang diukur menjadi nilai antara 0.0 dan 1.0. Dimana nilai terendah adalah 0.0 dan tertinggi adalah 1.0. Normalisasi min-max akan mempermudah perbandingan nilai yang memiliki satuan ukuran yang berbeda. Berikut adalah rumus normalisasi min-max :

$$d' = \frac{d - d_{min}}{d_{max} - d_{min}}$$

Keterangan:

d' = nilai yang telah dinormalisasi

d = nilai asli

d_{min} = nilai terkecil

d_{max} = nilai terbesar

2.2.7 Clustering

Clustering dalam konteks penggalian data adalah mengelompokkan kumpulan objek tertentu berdasarkan karakteristiknya, mengumpulkan berdasarkan kesamaan tertentu [20]. *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 [27]:

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

2.2.8 K-Means

K-Means *clustering* adalah teknik *clustering* berdasarkan pemartisian N objek pengamatan ke dalam K kelompok (*cluster*) dimana setiap objek pengamatan dimiliki oleh sebuah kelompok dengan rata-rata (*mean*) terdekat [8].

Teknik K-Means sangatlah simpel, pertama-tama tentukan K centroid awal, dimana K adalah parameter yang ditentukan dalam penelitian ini. Setiap poin ke ditetapkan pada centroid terdekat dan setiap kumpulan poin ditentukan pada centroid adalah *cluster*. Kemudian hal ini dilakukan kembali dan memperbarui langkah hingga *cluster* tidak berubah atau hingga centroidnya tetap. Tabel 2.7 adalah tahapan dari algoritma K-Means [27].

Tabel 2.7 Algoritma K-Means

| Algoritma : Algoritma K-Means Dasar | |
|--|---|
| 1 | Tentukan poin K sebagai centroid awal |
| 2 | Ulangi |
| 3 | Buatlah <i>cluster</i> K dengan menempatkan setiap point pada centroid terdekat |
| 4 | Hitung ulang centroid setiap <i>cluster</i> |
| 5 | Hingga centroid tidak berganti |

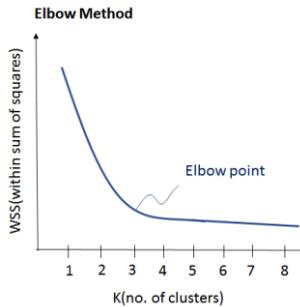
K-Means dapat digunakan diberbagai macam tipe data. K-Means juga sangat efisien walaupun terkadang dijalankan beberapa kali. Selain itu, metode ini mudah diinterpretasikan, mudah diimplementasikan dan dapat beradaptasi pada data yang tersebar. Namun, K-Means tidak dapat mengatasi cluster yang memiliki ukuran dan kepadatan yang berbeda. K-Means juga memiliki kesulitan untuk mengelompokan data yang mengandung pencilan. Pendeteksian dan penghapusan pencilan akan membantu proses K-Means [27].

Metode *clustering* K-Means digunakan dalam penelitian ini karena hasil yang diharapkan dari *clustering* adalah kelompok-

kelompok retailer yang objek didalamnya benar-benar hanya milik kelompok tersebut.

2.2.9 Metode Elbow

Metode *Elbow* adalah suatu metode untuk menghasilkan informasi dalam menentukan jumlah *cluster* terbaik dengan cara melihat prosentase hasil perbandingan antara jumlah *cluster* yang akan membentuk sudut siku pada suatu titik.



Gambar 2.1 Metode Elbow

Hasil persentase 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 hal ini juga dapat dilihat dari perbandingan nilai rasio SSE terbesar. Langkah-langkah dalam perhitungan metode *Elbow* adalah sebagai berikut [28]:

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

2.2.10 Performa Cluster

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 [26]:

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

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

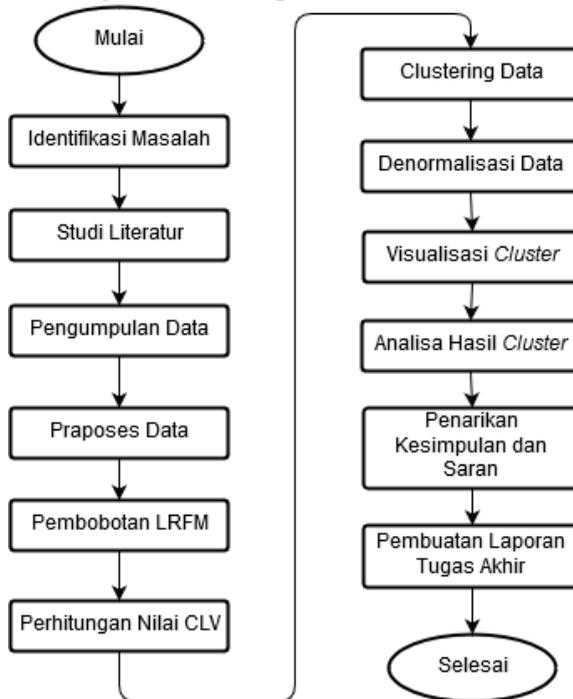
C_i = *cluster* ke- i

BAB III METODOLOGI PENELITIAN

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

3.1 Tahapan Pelaksanaan Tugas Akhir

Pada bagian ini akan dijelaskan mengenai proses pengerjaan maupun metode yang digunakan dalam pengerjaan dari tugas akhir ini. Tahapan pelaksanaan tugas akhir digambarkan seperti Gambar 3.1.



Gambar 3.1 Metodologi Penelitian

3.2 Uraian Metodologi

Berdasarkan diagram alur metodologi penelitian yang digambarkan pada sub bab sebelumnya, berikut ini merupakan penjelasan dari setiap prosesnya.

3.2.1 Identifikasi Masalah

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

3.2.2 Studi Literatur

Tahap ini membahas mengenai studi-studi terdahulu yang membahas segmentasi pelanggan, *clustering*, k-means, normalisasi, analisa model LRFM, metode AHP dan penggalan data yang dibahas di tugas akhir. Studi literatur dilakukan berdasarkan narasumber dengan melakukan wawancara, buku, penelitian sebelumnya dan dokumen terkait lainnya.

3.2.3 Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data-data yang berhubungan dengan segmentasi retailer PT. XYZ yang terbagi menjadi empat bagian yaitu :

- 1) Data retailer atau pelanggan dari PT. XYZ
- 2) Data transaksi pembelian produk ke retailer atau pelanggan dari 1 Januari 2017 hingga 13 Desember 2017
- 3) Data preferensi LRFM
- 4) Data pendukung lainnya yang berkaitan dengan tugas akhir

Data dikumpulkan menggunakan metode primer, meminta data mentah pada perusahaan dan juga metode sekunder melalui wawancara dan kuisisioner. Keluaran pada tahap ini adalah data yang dibutuhkan untuk proses *clustering*.

3.2.4 Praproses Data

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

3.2.4.1 Seleksi Data

Seleksi data dilakukan untuk mempermudah proses *clustering* sehingga data yang diolah tidak terlalu banyak. Pada tahapan ini data yang telah dikumpulkan akan dipilih variabel yang memiliki keterkaitan dengan proses clustering dan analisis LRFM. Dari variabel yang ada pada data penjualan produk PT.XYZ, variabel yang dipilih adalah kode retailer, nama retailer, tanggal transaksi dan jumlah transaksi.

3.2.4.2 Data Cleaning

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

3.2.4.3 Pemodelan LRFM

Pada tahapan ini data yang dipilih akan dianalisis menggunakan analisis LRFM. Masukan dari proses ini adalah data yang telah dilakukan praproses data dan keluarannya berupa variabel LRFM. Adapun pemodelan variabel-variabel tersebut diantaranya:

- a) *Length* : variabel *length* diperoleh dari interval transaksi awal hingga akhir yang dilakukan oleh pelanggan atau retailer. Atribut yang berkaitan dengan variabel ini diperoleh dari tanggal transaksi
- b) *Recency* : variabel *recency* diperoleh dari selisih waktu transaksi yang terakhir kali dilakukan oleh masing-masing retailer terhadap periode penelitian. Atribut yang berkaitan dengan variabel ini adalah tanggal transaksi
- c) *Frequency* : variabel *frequency* diperoleh dari jumlah transaksi yang dilakukan oleh masing-masing retailer dengan mengagregasikan semua data berdasarkan retailer.
- d) *Monetary* : variabel *monetary* diperoleh dari jumlah total pendapatan yang diperoleh perusahaan berdasarkan masing-masing pelanggan. Atribut yang berkaitan dengan variabel ini adalah *quantity* yang dijumlah berdasarkan transaksi tiap retailer.

3.2.4.4 Normalisasi Data

Pada tahapan ini dilakukan proses normalisasi keempat variabel LRFM yang memiliki range nilai berbeda antara satu sama lainnya, dimana *monetary* dapat memiliki nilai hingga jutaan sementara *length*, *recency* dan *frequency* memiliki nilai hanya berkisar pada puluhan hingga ratusan. Pada tahap ini akan dilakukan normalisasi sehingga keempat

variabel LRFM memiliki nilai yang sama. Metode normalisasi yang digunakan adalah Metode Min-Max, dimana semua variabel akan diubah menjadi nilai nol hingga satu.

3.2.5 Pembobotan LRFM

Pada tahap ini dilakukan pembobotan variabel LRFM yang dibobotkan sesuai dengan kepentingan yang sama. Hasil pembobotan akan digunakan untuk mencari nilai *customer lifetime value* yang digunakan untuk memberi ranking pada retailer. Metode yang digunakan dalam pembobotan ini adalah *Analytic Hierarchical Process* (AHP). Masukan pada tahapan ini adalah hasil kuisioner atau wawancara pada pihak yang terkait mengenai ketiga variabel LRFM.

3.2.6 Perhitungan Nilai CLV

Tahapan selanjutnya yaitu menghitung indeks CLV terhadap retailer atau 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 retailer dapat diketahui dari besar kecilnya indeks CLV. Luaran dari tahap ini adalah nilai atau ranking CLV dari masing-masing pelanggan atau retailer.

3.2.7 Clustering Data

Setelah melakukan pengumpulan data dan praproses data, serta perhitungan indeks CLV kemudian dilakukan pengelompokan menggunakan metode clustering. Metode *clustering* yang digunakan metode K-Means. *Clustering* terdiri dari 3 tahapan yaitu menentukan jumlah *cluster*, proses *cluster* dan yang terakhir adalah uji performa *cluster*. Masukan dari proses ini adalah data retailer yang telah dilakukan perhitungan CLV. Keluarannya berupa *cluster* yang telah optimum.

3.2.7.1 Menentukan Nilai K Menggunakan Metode Elbow

Metode *clustering* yang akan digunakan adalah non hiraki *cluster* atau *partition clustering* karena tujuan utamanya untuk mendapatkan hasil *cluster* yang benar-benar hanya jadi anggota di suatu kelompok. Pada tahap ini akan dilakukan penentuan nilai k atau jumlah kelompok yang digunakan menggunakan *elbow method*. Setelah itu dilakukan proses *clustering* menggunakan metode k-means.

3.2.7.2 Proses Clustering Menggunakan Metode K-Means

Pada tahap ini akan dilakukan proses *clustering* menggunakan metode K-Means. Hasil *cluster* yang didapatkan akan diberikan label nama sesuai dengan perilaku retailer atau pelanggan.

3.2.7.3 Uji Performa Cluster

Setelah proses *clustering* selesai, selanjutnya adalah tahap pengujian kualitas *cluster*. 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.

3.2.8 Denormalisasi Data

Pada tahapan ini data yang telah dinormalisasi dan di transformasi dikembalikan ke nilainya yang semula. Tahapan ini dilakukan untuk mempermudah analisis *cluster*.

3.2.9 Visualisasi Cluster

Pada tahap visualisasi *cluster*, hasil *cluster* akan divisualisasikan dengan media dashboard berbasis web dengan alat bantu RShiny. Visualisasi ini meliputi grafik *pie chart*, *scatter 3D*

dan *Box and Whisker*. Visualisasi ini bertujuan untuk memudahkan dalam pembacaan hasil *clustering* untuk menganalisis segmentasi pelanggan pada setiap *cluster*-nya.

3.2.10 Analisa Hasil Cluster

Pada tahapan ini, hasil *cluster* dianalisis dengan membandingkan nilai atribut LRFM dengan *cluster* lainnya. Selain itu ranking dari CLV dan jumlah anggota retailer yang termasuk dalam *cluster* tersebut serta perbandingan hasil *cluster* terhadap suatu area.

3.2.11 Penarikan Kesimpulan dan Saran

Pada tahapan ini ditarik kesimpulan berdasarkan *clustering* atau segmentasi yang dilakukan pada data retailer. Hasil *clustering* akan menghasilkan kelompok-kelompok retailer yang memiliki ciri yang sejenis. Kesimpulan yang dihasilkan akan menjawab tujuan penelitian. Serta memberikan saran untuk penelitian selanjutnya.

3.2.12 Pembuatan Laporan Tugas Akhir

Pembuatan laporan tugas akhir adalah tahapan akhir pada pengerjaan tugas akhir ini. Tahapan ini dilakukan sebagai dokumentasi seluruh proses pengerjaan tugas akhir. Laporan yang dibuat telah mengikuti format dokumen yang telah ditentukan.

Halaman ini sengaja dikosongkan

BAB IV PERANCANGAN

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

4.1 Pengumpulan Data

Pengumpulan data dilakukan secara langsung pada divisi Management Information System, PT. XYZ Cabang Gresik. Data diperoleh dari salah satu karyawan yang bertanggung jawab dalam pelaporan penjualan bulanan. Data yang dibutuhkan adalah riwayat transaksi penjualan dari tanggal 1 Januari 2017 hingga 13 Desember 2017 dan daftar pelanggan yang pernah melakukan transaksi. Data transaksi penjualan dan pelanggan didapatkan dalam bentuk dokumen *excel* dengan atribut dijelaskan pada tabel 4.1 dan 4.2.

Tabel 4.1 Data Mentah Tabel Transaksi

| Atribut | Tipe Data | Keterangan |
|----------------|------------------|----------------------------------|
| Tanggal | Date | Tanggal transaksi dilakukan |
| ID | Number | ID retailer |
| Produk | Text | Jenis produk |
| Qty | Number | Jumlah produk/uang yang diterima |
| Nama Ret | Text | Nama retail |
| SE Area | Text | Lokasi retailer |
| TM Area | Text | Kawasan retailer |
| SB01 | Text | Nama Sales Cabang Area |

Data transaksi yang diperoleh berjumlah 30.044 baris data atau transaksi, terhitung mulai Januari hingga awal Desember 2017.

Tabel 4.2 Data Mentah Tabel Retailer

| Atribut | Tipe Data | Keterangan |
|----------|-----------|------------------|
| ID | Number | ID retailer |
| Nama Ret | Text | Nama retail |
| SE Area | Text | Lokasi retailer |
| TM Area | Text | Kawasan retailer |

Data retailer yang diperoleh berjumlah kurang lebih 624 baris data, terhitung melakukan transaksi dari Januari hingga awal Desember 2017.

4.2 Praproses Data

Sebelum melakukan pengolahan data, sebelumnya data mentah memasuki tahapan praproses data, sehingga data mentah yang diperoleh dapat diolah dan menghasilkan hasil yang diharapkan. Tahapan praproses data yang dilakukan antara lain:

1. Pemilihan data untuk menentukan atribut yang akan digunakan sebagai variabel untuk dilakukan proses *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 minus ataupun tidak valid. Pembersihan data dilakukan agar hasil *clustering* lebih baik dengan tingkat kesalahan yang lebih kecil.

4.3 Pemodelan LRFM

Pemodelan variabel LRFM untuk mendapatkan nilai variabel LRFM yang dibutuhkan berdasarkan atribut yang tersedia.

- a. Variabel *length* diperoleh dari lamanya transaksi yang dilakukan oleh masing-masing retailer dari awal hingga akhir transaksi yang dilakukan. Atribut yang berkaitan dengan variabel ini adalah tanggal transaksi.
- b. Variabel *recency* diperoleh dari selisih waktu transaksi yang terakhir kali dilakukan oleh masing-masing retailer per tanggal 18 Maret 2018. Atribut yang berkaitan dengan variabel ini adalah tanggal transaksi.

- c. Variabel *frequency* diperoleh dari jumlah transaksi yang dilakukan oleh masing-masing retailer dengan mengagregasikan semua data berdasarkan setiap retailer.
- d. Variabel *monetary* diperoleh dari jumlah total pendapatan yang diperoleh oleh perusahaan berdasarkan masing-masing retailer. Atribut yang berkaitan dengan variabel ini adalah qty yang diakumulasi berdasarkan tiap retailer.

4.4 Normalisasi Data

Normalisasi data dilakukan untuk menyetaraan data dengan menyetarakan skala nilai variabel *monetary* yang cukup jauh dengan nilai variabel *length*, *recency*, dan *frequency*. Metode normalisasi yang digunakan adalah metode *min-max*.

4.5 Pembobotan LRFM

Tahapan selanjutnya adalah dalam analisis LRFM adalah menentukan bobot untuk setiap variabel LRFM. Bobot yang diberikan pada setiap variabel menentukan tingkat prioritas atau kepentingan setiap variabel yang ada. Pembobotan variabel ini diperoleh dengan cara proses pengisian kuesioner oleh pihak terkait di fungsi *Management Information System* maupun di pihak terkait lainnya di PT. XYZ Cabang Gresik. Responden yang mengisi kuesioner ini terdapat 3 orang responden, yaitu *Head of Territory Madura-Gresik*, *Deputy Head of Branch Madura-Gresik*, *Deputy Head of Branch Jombang*. Kuesioner yang diberikan berisikan mengenai data responden, petunjuk pengisian kuesioner, penjelasan mengenai variabel LRFM, skala preferensi, contoh pengisian, dan daftar pertanyaan. Rancangan kuesioner dapat dilihat pada Lampiran A.

4.6 Penentuan Jumlah Cluster

Sebelum proses *clustering* dilakukan, tahap sebelumnya adalah menentukan nilai k atau jumlah *cluster* yang akan dibentuk. Dalam penentuan nilai k digunakan metode *Elbow*. Hasil dari metode *Elbow* adalah grafik yang menunjukkan nilai SSE pada

setiap nilai k . Nilai k dipilih berdasarkan nilai terbesar penurunan SSE. Aplikasi yang digunakan untuk menghitung nilai k pada penelitian ini adalah RStudio.

4.7 Rancangan Visualisasi

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

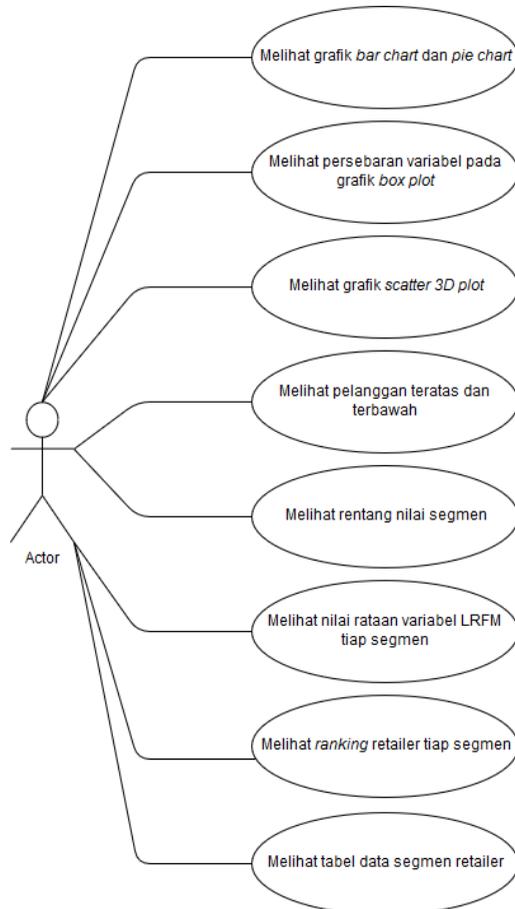
4.7.1 Use Case Diagram

Visualisasi dibuat untuk menampilkan setiap grafik dan data yang menjelaskan informasi mengenai karakteristik tiap segmen. Gambar 4.1 merupakan gambaran dari *use case diagram* yang terbagi dari 8 *use case* utama, yaitu:

1. Melihat grafik *bar chart* dan *pie chart*
2. Melihat persebaran variabel pada grafik *box-plot*
3. Melihat grafik *scatter 3D plot*
4. Melihat retailer teratas dan terbawah
5. Melihat rentang nilai segmen
6. Melihat nilai rata-rata variabel LRFM tiap segmen
7. Melihat *ranking* retailer tiap segmen
8. Melihat tabel data segmen retailer

4.7.2 Deskripsi Use Case

Use case pertama yang akan dijelaskan adalah *use case* untuk melihat grafik *bar chart* dan *pie chart* yang dapat dilihat pada Tabel 4.3



Gambar 4.1 Use Case Diagram

Untuk tujuh deskripsi *use case* lainnya dapat dilihat pada Lampiran B.

Tabel 4.3 Use case melihat grafik bar chart dan pie chart

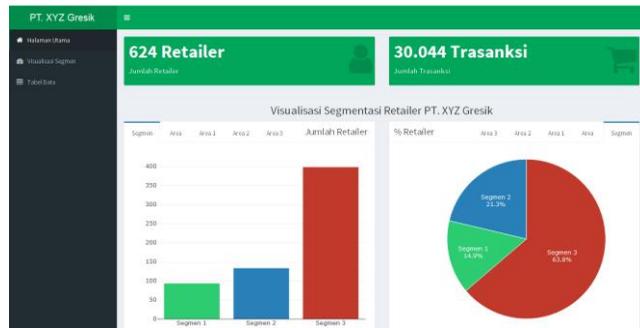
| UC-1 Melihat grafik bar chart dan pie chart | | |
|--|---|--|
| Purpose | Mengetahui jumlah dan presentase retailer pada masing-masing segmen serta jumlah dan peresentase retailer pada setiap area tiap segmen dalam grafik <i>bar chart</i> dan <i>pie chart</i> | |
| Overview | Dimulai ketika pengguna mengakses aplikasi atau memasuki laman Halaman Utama Aplikasi | |
| Actors | Pengguna Aplikasi | |
| Pre Condition | Data hasil <i>clustering</i> sudah dimasukkan ke dalam aplikasi | |
| Post Condition | Grafik <i>bar chart</i> dan <i>pie chart</i> ditampilkan | |
| Typical Cause Event | Actor | System |
| | 1. Mengklik ikon  <i>slide bar</i> 3. Mengklik menu “Halaman Utama” | 2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan grafik jumlah retailer per segmen dan per area dan presentase retailer per segmen dan area |
| Alternate Flow Event | - | |
| Exceptional Flow Event | Aplikasi tertutup akibat <i>error</i> | |

4.7.3 Desain Antarmuka Aplikasi

Desain antarmuka aplikasi menjelaskan mengenai tampilan dari aplikasi yang diakses oleh pengguna. Aplikasi untuk menampilkan visualisasi hasil *clustering* berbasis web dengan desain 3 halaman yang dapat diakses oleh pengguna, diantara lain:

a. *Halaman Utama*

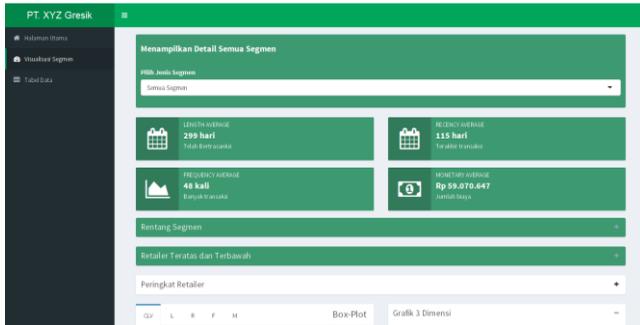
Halaman utama aplikasi menampilkan informasi jumlah retailer, jumlah transaksi, grafik jumlah retailer tiap segmen dan area dalam bentuk *bar chart* serta menampilkan presentase retailer tiap segmen dan areanya dalam bentuk *pie chart*. Rancangan halaman utama aplikasi disajikan pada Gambar 4.2.



Gambar 4.2 Halaman Utama Aplikasi

b. *Halaman Visualisasi Segmen*

Halaman visualisasi segmen akan menampilkan informasi mengenai kriteria retailer tiap segmen diantaranya informasi rata-rata nilai variabel LRFM, rentang nilai variabel LRFM, informasi retailer teratas dan terbawah, informasi mengenai peringkat retailer dan yang terakhir informasi mengenai grafik *box plot* variabel LRFM setiap segmen yang dipilih. Rancangan visualisasi halaman visualisasi segmen dapat dilihat pada Gambar 4.3.



Gambar 4.3 Halaman Visualisasi Segmen

c. *Halaman Tabel Data*

Halaman tabel data akan menampilkan informasi keseluruhan tabel secara detail mengenai peringkat retailer. Rancangan halaman tabel data dapat dilihat pada Gambar 4.4.

| ID | Retailer | Length | Rececy | Frequency | Monetary | CLV | Segmen | Area |
|--------|---------------|--------|--------|-----------|----------|-----------|--------|--------------|
| 10372 | msi call | 323 | 111 | 105 | 1073000 | 0.202942 | 2 | LAR-KINGGIRI |
| 10373 | msi call | 281 | 179 | 26 | 1879300 | 0.181013 | 1 | LAR-KINGGIRI |
| 10377 | eko call | 352 | 90 | 41 | 1718200 | 0.2319073 | 3 | LAR-KINGGIRI |
| 10378 | top call | 520 | 110 | 110 | 5007000 | 0.3313091 | 2 | LAR-KINGGIRI |
| 10374 | amy call | 210 | 110 | 26 | 490000 | 0.1320990 | 3 | LAR-KINGGIRI |
| 103742 | ronggawe call | 326 | 90 | 18 | 320000 | 0.1710376 | 1 | LAR-KINGGIRI |
| 103743 | mira call | 394 | 132 | 79 | 1049000 | 0.240104 | 2 | LAR-KINGGIRI |
| 103745 | pagul call | 208 | 144 | 10 | 200000 | 0.1481605 | 2 | LAR-KINGGIRI |
| 103747 | rkg call | 333 | 113 | 40 | 1190000 | 0.2444007 | 3 | LAR-KINGGIRI |
| 103941 | ip call | 204 | 120 | 20 | 750000 | 0.1877087 | 3 | LAR-KINGGIRI |

Gambar 4.4 Halaman Tabel Data

BAB V IMPLEMENTASI

Pada bab ini menjelaskan proses implementasi berdasarkan dari rancangan yang dibuat yaitu praproses data, transformasi dan normalisasi data, pembobotan LRFM, perhitungan CLV, proses *clustering*, uji performa *clustering*, proses denormalisasi data dan proses pembuatan visualisasi.

5.1 Praproses Data

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

5.1.1 Seleksi Data

Pada tahapan ini akan dilakukan pemilihan atribut yang digunakan dalam proses *clustering*. Dari atribut data yang ada pada tabel 4.1, atribut yang digunakan adalah tanggal, ID, Qty, Nama Ret, SE Area. Sementara atribut yang digunakan pada tabel retailer adalah ID, Nama Ret. Pada tabel 5.1 merupakan kumpulan variabel yang digunakan pada penelitian ini.

Tabel 5.1 Atribut Terpilih

| Atribut | Keterangan |
|----------------|---|
| Tanggal | Atribut tanggal digunakan untuk memperoleh variabel <i>length</i> , <i>recency</i> dan <i>frequency</i> |
| ID | Atribut ID menunjukkan kode retailer |
| Qty | Atribut Qty digunakan untuk memperoleh variabel monetary tiap retailer |
| Nama Ret | Atribut Nama Ret menjelaskan nama retail |
| SE Area | Atribut SE Area menunjukkan lokasi retailer |

5.1.2 Data Cleaning

Tahapan kedua dari praproses data adalah melakukan pembersihan data atau penghapusan data yang tidak valid pada atribut-atribut tersebut. Saat dilakukan validasi data, terdapat beberapa baris yang bernilai 0 dan negatif pada atribut Qty. Pencarian baris yang bernilai 0 dan negative dilakukan dengan fitur *sort* pada *excel*. Setelah diurutkan, didapatkan 134 baris data yang tidak valid yang termasuk dalam data bernilai 0 maupun negatif. Adapun potongan data yang dihapus ditampilkan pada Tabel 5.2.

Tabel 5.2 Baris Data Tidak Valid

| ID | Qty |
|-----------|--------------|
| 219205 | 0 |
| 301751 | 0 |
| 301753 | 0 |
| 302049 | 0 |
| 302056 | 0 |
| 152914 | 0 |
| 126446 | -50.000.000 |
| 20886 | -300.000.000 |
| 362751 | -200.000.000 |
| 20886 | -145.948.320 |
| 126449 | -5.000.000 |
| 73168 | -5.000.000 |
| 20886 | -43.263.550 |
| 73780 | -1.500.000 |
| 74765 | -15.000.000 |

5.2 Pemodelan LRFM

Pemodelan LRFM dilakukan untuk membentuk nilai variabel *length*, *recency*, *frequency*, dan *monetary* sebagai masukan proses *clustering*. Nilai LRFM didapatkan dengan mengekstrak nilai atribut yang sebelumnya dipilih. Pemodelan LRFM ini

dilakukan dengan melakukan proses *query* pada basis data *MySQL*.

Nilai variabel *length* atau L adalah lamanya waktu yang telah ditempuh oleh retailer dalam bertransaksi dengan perusahaan, dimana variabel ini dapat dihitung dengan mencari selisih waktu terakhir retailer melakukan transaksi dan awal retailer melakukan transaksi. Atribut yang dibutuhkan dalam memperoleh nilai L adalah ID dan Tanggal. Nilai L didapatkan dari tanggal transaksi yang paling terakhir menggunakan fungsi MAX dan tanggal awal melakukan transaksi dengan fungsi MIN berdasarkan tiap retailer 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 retailer dengan hari pengerjaan, yaitu 18 Maret 2018. Atribut yang dibutuhkan untuk memperoleh nilai R adalah ID dan Tanggal. Nilai R diperoleh dari tanggal transaksi terakhir menggunakan fungsi MAX berdasarkan tiap retailer dengan fungsi GROUP BY, sedangkan untuk mencari selisihnya menggunakan fungsi DATEDIFF.

Nilai variabel *frequency* atau F adalah jumlah transaksi yang dilakukan oleh setiap retailer. Atribut yang diperlukan untuk mendapatkan nilai F adalah ID dan Tanggal. Nilai F diperoleh dengan menghitung banyaknya Tanggal dengan fungsi COUNT dan GROUP BY berdasarkan ID retailer.

Nilai variabel *monetary* atau M adalah jumlah pemasukan yang diterima oleh pihak perusahaan dari tiap retailer. Atribut yang dibutuhkan untuk memperoleh nilai M adalah ID dan Qty. Nilai M didapatkan dengan menjumlah atribut Qty dengan fungsi SUM dan fungsi GROUP BY berdasarkan ID retailer.

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

```
SELECT `ID`, `NAME RET`,
DATEDIFF(MAX(TANGGAL),MIN(TANGGAL)) AS L,
DATEDIFF('2018-03-18',MAX(TANGGAL)) AS R,
COUNT(DISTINCT TANGGAL) AS F,
SUM(QTY) AS M
FROM `trasanksi madura-gresik`
GROUP BY `ID`
ORDER BY `ID`;
```

Script 5.1 Query Ekstraksi Data

Setelah *query* dijalankan, diperoleh data sejumlah 624 retailer yang melakukan transaksi sejak 1 Januari 2017 hingga 13 Desember 2017. 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

| ID | NAME RET | L | R | F | M |
|--------|-----------------|-----|-----|-----|------------|
| 103072 | new cell | 321 | 111 | 105 | 16.710.000 |
| 103073 | mila cell | 216 | 170 | 26 | 13.575.000 |
| 103473 | eka cell | 332 | 96 | 41 | 17.012.500 |
| 103478 | top cell | 329 | 110 | 130 | 50.070.000 |
| 103741 | army cell | 239 | 103 | 28 | 4.960.000 |
| 103742 | ronggolawe cell | 316 | 99 | 18 | 3.250.000 |
| 103743 | mitra cell | 304 | 132 | 79 | 19.863.000 |
| 103745 | jagad cell | 286 | 144 | 10 | 2.600.000 |

5.3 Normalisasi Data

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 *Excel Worksheet* (.xlsx) diubah ke

dalam format *Comma Separated Value* (.csv) agar dapat diproses.

Kemudian, normalisasi dilakukan dengan menggunakan metode *Min-Max* dari skala 0 hingga 1. Normalisasi dilakukan pada semua nilai variabel L,R,F, dan M dengan memperhatikan nilai maksimum dan minimum setiap variabel yang dapat dilihat pada tabel 5.4.

Tabel 5.4 Rentang Nilai Atribut LRFM

| Variabel | Minimum | Maksimum |
|------------------|---------|---------------|
| <i>Length</i> | 49 | 345 |
| <i>Recency</i> | 95 | 373 |
| <i>Frequency</i> | 2 | 241 |
| <i>Monetary</i> | 448.000 | 5.832.824.162 |

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

```
#normalisasi MIN-MAX
L_Norm <- (Model_LRFM$L-min(Model_LRFM$L))/(max(Model_LRFM$L)-min(Model_LRFM$L))
R_Norm <- (Model_LRFM$R-min(Model_LRFM$R))/(max(Model_LRFM$R)-min(Model_LRFM$R))
F_Norm <- (Model_LRFM$F-min(Model_LRFM$F))/(max(Model_LRFM$F)-min(Model_LRFM$F))
M_Norm <- (Model_LRFM$M-min(Model_LRFM$M))/(max(Model_LRFM$M)-min(Model_LRFM$M))
HasilNorm <- data.frame(Model_LRFM,L_Norm,R_Norm,F_Norm,M_Norm)
View(HasilNorm)
write.csv(HasilNorm,"D:/KULIAH/SKRIPSI/Model LRFM Normalisasi.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 L, F dan M yang semakin kecil menunjukkan nilai yang buruk. Untuk

itu, variabel R diubah dengan cara mengurangi nilai R terhadap 1 pada kolom baru bernama R Norm Rev. Hasil lengkap nilai normalisasi LRFM dapat dilihat pada Lampiran D

5.4 Pembobotan LRFM

Dari proses pengisian kuesioner, didapatkan nilai preferensi terhadap empat 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 keempat 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 sebaliknya. Uji konsistensi ini dilakukan dengan cara mengalikan setiap nilai perbandingan pada *normal comparisons matrix* dengan bobot masing-masing kriteria. Kemudian setiap baris kriteria dijumlahkan menjadi *overall score*. Selanjutnya, dicari nilai lambda dengan merata-rata hasil pembagian *overall score* dengan nilai bobot masing-masing kriteria.

5.5 Perhitungan CLV

Perhitungan CLV dilakukan untuk mengetahui nilai retailer terhadap perusahaan di masa yang akan datang. Selain itu, nilai CLV juga digunakan sebagai masukan untuk proses *clustering*, akan tetapi hanya nilai variabel LRFM saja yang digunakan. Data yang dibutuhkan adalah data LRFM setiap retailer. Nilai variabel LRFM tiap Retailer dikalikan dengan bobot tiap

variabel. Selanjutnya, hasil perkalian tersebut dijumlahkan untuk memperoleh nilai CLV setiap retailer, dan digunakan untuk bahan analisa segmentasi.

```
#Menghitung Nilai CLV
LBobot <- Model_LRFM_Normalisasi$L_Norm*0.11
RBobot <- Model_LRFM_Normalisasi$R_Norm_Rev*0.06
FBobot <- Model_LRFM_Normalisasi$F_Norm*0.31
MBobot <- Model_LRFM_Normalisasi$M_Norm*0.51
CLV <- LBobot+RBobot+FBobot+MBobot

HasilCLV <- data.frame(Model_LRFM_Normalisasi,LBobot,RBobot,FBobot,MBobot,CLV)
View(HasilCLV)

write.csv(HasilCLV,"D:/KULIAH/SKRIPSI/Nilai CLV.csv")
```

Script 5.3 Menghitung Nilai CLV

5.6 Clustering Data

Clustering data dilakukan untuk menghasilkan *cluster* yang beranggotakan retailer dengan karakteristik yang sama. Masukan dari proses ini adalah data nilai variabel LRFM semua retailer. Dalam proses ini terdapat 3 subproses, diantaranya penentuan nilai atau jumlah K dan *clustering* dengan menggunakan metode *K-Means* serta uji performa *cluster*. Keseluruhan proses *clustering* dilakukan dengan menggunakan aplikasi RStudio.

5.6.1 Menentukan Jumlah K dengan Menggunakan Metode Elbow

Data dari proses sebelumnya, yaitu nilai variabel LRFM semua retailer, digunakan sebagai dataset untuk menentukan jumlah K atau jumlah *cluster* yang dihasilkan. Metode yang digunakan adalah metode Elbow, dimana pada metode ini akan dihasilkan sebuah grafik dua dimensi. Pada sumbu X grafik menggambarkan jumlah K dan sumbu Y mempresentasikan nilai SSE yang diperoleh pada setiap nilai K.

Dalam menentukan nilai K, rentang K yang digunakan adalah 1 hingga 15. Kemudian akan 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 5.4.

```
#input data LRFM
library(readxl)
LRFMNormal <- read_excel("D:/KULIAH/SKRIPSI/Model LRFM
Normalisasi.xlsx")
View(LRFMNormal)

#variabel yang digunakan
HitungElbow2 <-data.frame
(LRFMNormal$L_Norm, LRFMNormal$R_Norm_Rev,
LRFMNormal$F_Norm, LRFMNormal$M_Norm)
View(HitungElbow2)

#Metode Elbow
k.max <- 15
elbow2 <- sapply
(1:k.max, function(k) {kmeans(HitungElbow2, k, nstart =
10)$tot.withinss})

plot (1:k.max, elbow2,type="b", pch = 19, xlab="Jumlah K",
ylab="Sum of Squares")
View(elbow2)
```

Script 5.4 Metode Elbow

5.6.2 Proses Clustering Menggunakan Metode K-Means

Setelah memperoleh nilai K optimal atau terbaik, tahap selanjutnya adalah proses *clustering*. Dataset yang digunakan masih sama seperti tahapan sebelumnya, yakni dataset nilai variabel LRFM semua retailer. Proses *clustering* menggunakan metode *K-Means* untuk menghasilkan derajat keanggotaan setiap retailer terhadap setiap *cluster*. Hasil *clustering* kemudian disimpan dalam format file *.csv*. Berikut ini adalah

potongan kode yang digunakan dalam proses *clustering* yang ditampilkan pada Script 5.5.

```
#masukan data yang telah dinormalisasi
library(readxl)
LRFMNormal <- read_excel
("D:/KULIAH/SKRIPSI/Model LRFM Normalisasi.xlsx")
View(LRFMNormal)

#Variabel yang digunakan
datanorm <- data.frame
(LRFMNormal$L_Norm,          LRFMNormal$R_Norm_Rev,
LRFMNormal$F_Norm, LRFMNormal$M_Norm)
View(datanorm)

#Clustering K-Means 3 Segmen
NCluster3 <- kmeans (datanorm, 3)
Aggregate(datanorm,by=list(NCluster3$cluster),
FUN=mean)

HasilNCluster3 <- data.frame
(LRFMNormal,NCluster3$cluster)
View(HasilNCluster3)

write.csv (HasilNCluster3,"D:/KULIAH/SKRIPSI/Cluster 3
Normal.csv")
```

Script 5.5 Proses Clustering

5.6.3 Uji Performa Cluster

Setelah proses *clustering* dilakukan, maka dihasilkan beberapa *cluster* atau segmen yang membagi semua retailer ke setiap *cluster*. Untuk mengetahui apakah jumlah segmen yang ditentukan telah optimal, diperlukan uji performa *cluster*. Pengujian ini dibantu dengan menggunakan metode SSE. Data yang diperlukan adalah data nilai variabel LRFM semua retailer 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 dapat dilihat pada Script 5.6

```
#masukan data yang telah dinormalisasi
library(readxl)
LRFMNormal <- read_excel
("D:/KULIAH/SKRIPSI/Model LRFM Normalisasi.xlsx")
View(LRFMNormal)

#Variabel yang digunakan
NSSE <- data.frame
(LRFMNormal$L_Norm, LRFMNormal$R_Norm_Rev,
LRFMNormal$F_Norm, LRFMNormal$M_Norm)
View(NSSE)

#Hitung SSE Pada data normalisasi
nwss = kmeans(NSSE, centers = 1)$tot.withinss
for (i in 2:15)
nwss[i] = kmeans(NSSE, centers=i)$tot.withinss

plot(nwss,xlab="Jumlah K",ylab="Sum of Squares")
View(nwss)
```

Script 5.6 Uji Performa Cluster

5.7 Denormalisasi Data

Pada tahapan ini dilakukan pembalikan nilai dari normalisasi data untuk mempermudah proses analisa hasil *clustering*. Script 5.7 menunjukkan proses denormalisasi data yang dilakukan dengan menggunakan *tools* RStudio

```

#denormalisasi
R_Rev <- (1-LRFMdeNorm$R_Norm_Rev)

L <- (LRFMdeNorm$L_Norm*(max(LRFMdeNorm$L)-
min(LRFMdeNorm$L)))+min(LRFMdeNorm$L)

R <- (R_Rev*(max(LRFMdeNorm$R)
-min(LRFMdeNorm$R)))+min(LRFMdeNorm$R)

F <- (LRFMdeNorm$F_Norm*(max(LRFMdeNorm$F)-
min(LRFMdeNorm$F)))+min(LRFMdeNorm$F)

M <- (LRFMdeNorm$M_Norm*(max(LRFMdeNorm$M)
-min(LRFMdeNorm$M)))+min(LRFMdeNorm$M)

```

Script 5.7 Denormalisasi Data

5.8 Pembuatan Halaman Visualisasi

Proses rancang bangun halaman visualisasi menggunakan aplikasi Rshiny. *Package* yang dijalankan dalam aplikasi ini yaitu *package* shiny dan *package* shinydashboard untuk merancang desain tampilan visualisasi. Penjelasan lebih lengkap mengenai rancang bangun halaman visualisasi dijelaskan pada bagian-bagian berikut.

5.8.1 Halaman Utama Aplikasi

Halaman utama menampilkan beberapa informasi diantaranya adalah 2 jenis grafik yang menjelaskan jumlah dan presentase retailer tiap segmen maupun tiap area.

a. Tampilan Bar Chart

Dalam pembuatan *bar chart* dibutuhkan masukan data berupa hasil *clustering*. Data hasil *clustering* tersebut kemudian diagregasikan berdasarkan kelompok *cluster* atau area. Kemudian dalam proses rancang bangun grafik dibutuhkan *package* *plotly*, dan dalam proses visualisasi *bar chart* digambarkan dengan sumbu X berupa jenis *cluster* atau area, sedangkan sumbu Y dipresentasikan sebagai jumlah retailer dalam *cluster*

atau area. Proses rancang bangun visualisasi *bar chart* dapat dilihat pada Script 5.8.

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

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

#Kode Warna
warna <- c("#2ecc71", "#2980b9", "#c0392b")
```

Script 5.8 Rancang Bangun Visual Bar Chart

b. *Tampilan Pie Chart*

Seperti halnya rancang bangun *bar chart*, rancang bangun visualisasi *pie chart* membutuhkan data masukan dari hasil proses *clustering* dan membutuhkan *package plotly* dalam proses rancang bangun visualisasi *pie chart*. Script 5.9 merupakan proses rancang bangun visualisasi *pie chart*.

```

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

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

#Kode Warna
warna <- c("#2ecc71", "#2980b9", "#c0392b")

```

Script 5.9 Rancang Bangun Visual Pie Chart

5.8.2 Halaman Visualisi Segmen

Halaman visualisasi segmen akan menampilkan informasi mengenai kriteria retailer tiap segmen diantaranya informasi rata-rata nilai variabel LRFM, rentang nilai variabel LRFM, informasi retailer teratas dan terbawah, informasi mengenai peringkat retailer dan yang terakhir informasi mengenai grafik *box plot* variabel LRFM setiap segmen yang dipilih. Rancang bangun visualisasi halaman visualisasi segmen dibagi atas beberapa bagian visualisasi. Berikut ini adalah bagian-bagian visualisasi pada halaman visualisasi segmen.

a. *Visualisasi Nilai Rataan Variabel LRFM*

Rancang bangun visual nilai rata-rata variabel LRFM ditampilkan dalam bentuk kotak-kotak informasi yang terpisah. Data masukan yang diperlukan untuk menampilkan informasi rata-rata LRFM adalah rata-rata dari variabel LRFM tiap segmennya. Proses rancang bangun visualisasi nilai rata-rata variabel LRFM dapat dilihat pada Script 5.10 dan hasil dari proses rancang bangun visualisasi nilai rata-rata variabel LRFM dapat dilihat pada Gambar 5.1

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

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

```

Script 5.10 Rancangan Bangun Visual Nilai Rataan LRFM



Gambar 5.1 Hasil Visual Nilai Rataan LRFM

b. Visualisasi Rentang Nilai Variabel LRFM

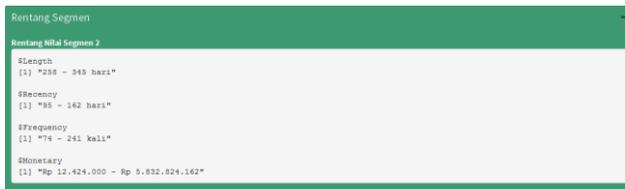
Rancang bangun visual rentang nilai variabel LRFM ditampilkan dalam bentuk kotak-kotak informasi. Data masukan yang diperlukan untuk menampilkan informasi rentang nilai LRFM adalah nilai variabel LRFM tiap segmennya. Proses rancang bangun

visualisasi rentang nilai variabel LRFM dapat dilihat pada Script 5.11 dan hasil dari proses rancang bangun visualisasi rentang nilai variabel LRFM dapat dilihat pada Gambar 5.2

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

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

Script 5.11 Rancang Bangun Visual Rentang Nilai LRFM



Gambar 5.2 Hasil Visual Rentang Nilai LRFM

c. *Visualisasi Informasi Retailer Teratas & Terbawah*

Informasi retailer teratas dan terbawah divisualkan dalam bentuk kotak informasi. Data masukan yang diproses adalah nilai CLV retailer setiap segmennya, dimana retailer teratas diperoleh dari nilai CLV tertinggi tiap segmennya dan begitupun sebaliknya retailer terbawah diperoleh dari nilai CLV terendah tiap segmennya. Proses rancang bangun visual informasi retailer teratas dan terbawah dapat dilihat pada Script 5.12 dan hasil visualisasi dari retailer teratas dan terbawah dapat dilihat pada Gambar 5.3.

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

#Retailer Terbawah
output$sterbawah<- renderValueBox({
  inputan <- pilihdata()
  cari <- sqldf("Select Retailer From(select
Retailer,min(CLV) From inputan)" )
  valueBox( value = tags$h5( "RETAILER
TERBAWAH"),
    subtitle = tags$h4(tags$strong (cari)),
    icon = icon("thumbs-o-down"),
    color = "olive" )
  })
```

Script 5.12 Rancang Bangun Visual Retailer Teratas & Terbawah



Gambar 5.3 Hasil Visual Retailer Teratas & Terbawah

d. *Visualisasi Peringkat Retailer*

Rancang bangun visualisasi peringkat retailer ditampilkan dalam bentuk tabel yang berisi beberapa data pelanggan dan nilai CLVnya yang kemudian diurutkan berdasarkan nilai CLV dari terbesar hingga terkecil. Data masukan yang dibutuhkan adalah data retailer dan nilai CLVnya. Proses rancang bangun visualisasi peringkat retailer dapat dilihat pada Script 5.13 serta hasil visualisasinya dapat dilihat pada Gambar 5.4.

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

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

Script 5.13 Rancang Bangun Tabel Peringkat Retailer

Peringkat Retailer

Show 5 entries Search:

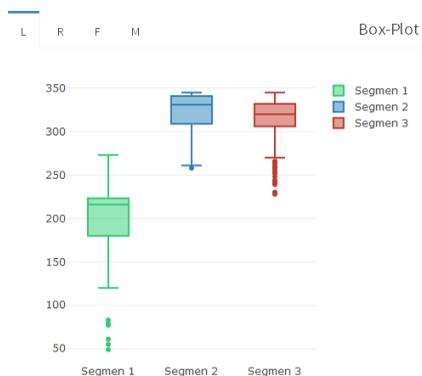
| Retailer | Length | Recency | Frequency | Monetary | CLV | Area |
|----------------------------|--------|---------|-----------|------------|-----------|-----------|
| lancar bola batu sejattira | 340 | 95 | 76 | 5832024162 | 0.7742252 | LAMONGAN |
| AFRA 2 cell | 338 | 100 | 127 | 2937544213 | 0.5852817 | LAMONGAN |
| Naja cell 2 | 340 | 97 | 75 | 3306050500 | 0.5515179 | PAMEKASAN |
| zahera cell | 343 | 96 | 241 | 69075000 | 0.4051118 | GRESIK |
| caapr cell | 344 | 95 | 232 | 130600000 | 0.4793356 | GRESIK |
| merah putih | 343 | 95 | 180 | 635500000 | 0.4557056 | LAMONGAN |
| perdana cell | 317 | 95 | 197 | 293205000 | 0.4301300 | GRESIK |
| ardan cell | 344 | 95 | 199 | 640605000 | 0.4307619 | GRESIK |

Showing 1 to 8 of 133 entries Previous 1 2 3 4 5 ... 17 Next

Gambar 5.4 Hasil Visual Tabel Peringkat Retailer

e. Visualisasi Grafik Box-Plot Variabel LRFM

Rancang bangun grafik *box plot* variabel LRFM divisualisasikan dalam bentuk grafik *box plot*, data masukan dalam rancang bangun ini adalah nilai LRFM serta nilai CLV tiap segmennya. Grafik *box plot* akan menampilkan nilai rata-rata, minimal, maksimal, kuartil I dan kuartil III tiap variabel setiap segmennya. Proses rancang bangun grafik *box plot* variabel LRFM dapat dilihat pada Script 5.14 dan hasil visualisasinya dapat dilihat pada Gambar 5.5.



Gambar 5.5 Hasil Visual Box-Plot

```

#Data box plot
bp1 <- subset(data,data$Cluster==1)
bp2 <- subset(data,data$Cluster==2)
bp3 <- subset(data,data$Cluster==3)
#Box Plot
output$boxl <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = bp1$L, name = "Segmen 1",
               marker = list(color = "#2ecc71"),
               line = list(color = "#2ecc71")) %>%
    add_boxplot(y = bp2$L, name = "Segmen 2",
               marker = list(color = "#2980b9"),
               line = list(color = "#2980b9")) %>%
    add_boxplot(y = bp3$L, name = "Segmen 3",
               marker = list(color = "#c0392b"),
               line = list(color = "#c0392b")) })
output$boxr <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = bp1$R, name = "Segmen 1",
               marker = list(color = "#2ecc71"),
               line = list(color = "#2ecc71")) %>%
    add_boxplot(y = bp2$R, name = "Segmen 2",
               marker = list(color = "#2980b9"),
               line = list(color = "#2980b9")) %>%
    add_boxplot(y = bp3$R, name = "Segmen 3",
               marker = list(color = "#c0392b"),
               line = list(color = "#c0392b"))
})

output$boxf <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = bp1$F, name = "Segmen 1",
               marker = list(color = "#2ecc71"),
               line = list(color = "#2ecc71")) %>%
    add_boxplot(y = bp2$F, name = "Segmen 2",
               marker = list(color = "#2980b9"),
               line = list(color = "#2980b9")) %>%
    add_boxplot(y = bp3$F, name = "Segmen 3",
               marker = list(color = "#c0392b"),
               line = list(color = "#c0392b"))
})

```

```

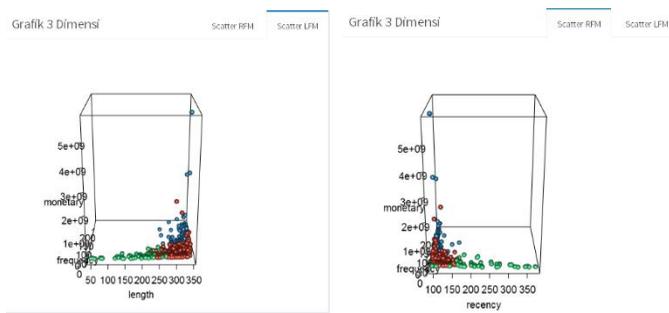
output$boxm <- renderPlotly ({
  plot_ly(type = 'box') %>%
    add_boxplot(y = bp1$M, name = "Segmen 1",
      marker = list(color = "#2ecc71"),
      line = list(color = "#2ecc71")) %>%
    add_boxplot(y = bp2$M, name = "Segmen 2",
      marker = list(color = "#2980b9"),
      line = list(color = "#2980b9")) %>%
    add_boxplot(y = bp3$M, name = "Segmen 3",
      marker = list(color = "#c0392b"),
      line = list(color = "#c0392b"))
})

```

Script 5.14 Rancang Bangun Visual Blox-Plot

f. *Visualisasi Scatter Plot 3D Variabel LRFM*

Proses visualisasi *scatter plot 3D* dibutuhkan masukan data berupa variabel LRFM. Bentuk visualisasi ini ditampilkan dalam bentuk kubus 3 dimensi dimana terdiri dari sumbu x,y,z yang mewakili variabel LRFM. Namun disini dikarenakan terdapat 4 variabel, maka *scatter plot 3D* akan dibuat 2 bentuk yakni mewakili variabel LFM dan mewakili variabel RFM. Berikut ini adalah proses visualisasi grafik *scatter plot 3D* yang dapat dilihat pada Script 5.15 dan hasil visualisasi dapat dilihat pada Gambar 5.6.



Gambar 5.6 Hasil Visual Scatter Plot 3D

```

#Scatter Plot 3 Dimensi LFM
output$scatter <- renderRglwidget ({
  dplot <- pilihdata()
  open3d()
  plot3d(x=dplot$L,
        y=dplot$F,
        z=dplot$M,
col      =      c("#2ecc71",      "#2980b9",
"#c0392b")[as.numeric(dplot$Cluster)],
        type = 's', size = 1,
        xlab = "length",
        ylab = "frequency",
        zlab = "monetary",
        colkey = list(length = 0.5, width = 0.5,
        cex.clab = 0.75)
  )
  scene1 <- scene3d()
  rgl.close()
  save <- options(rgl.inShiny = TRUE)
  on.exit(options(save))
  )
  rglwidget(scene1) })
output$scatterr <- renderRglwidget ({
  dplot <- pilihdata()
  open3d()
  plot3d(x=dplot$R,
        y=dplot$F,
        z=dplot$M,
col      =      c("#2ecc71",      "#2980b9",
"#c0392b")[as.numeric(dplot$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 5.15 Rancang Bangun Visual Scatter Plot 3D

5.8.3 Halaman Tabel Data

Rancang bangun halaman tabel data dengan memvisualisasikan tabel retailer dengan keseluruhan detail seperti, peringkat seluruh retailer, nilai variabel LRFM, id dan nama retailer, nilai CLV serta area retailer. Proses rancang bangun visualisasi halaman tabel data dapat dilihat pada Script 5.15 dan hasil dari visualisasinya dapat dilihat pada Gambar 5.6

```

#Data Tabel Keseluruhan
output$Table = renderDataTable({
  tabel <- data
  tampil <- data.frame(tabel$ID, tabel$Retailer, tabel$L,
    tabel$R, tabel$F, tabel$M, tabel$CLV, tabel$Cluster,
    tabel$Area )
  colnames(tampil) <- c("ID", "Retailer", "Length", "Recency",
    "Frequency", "Monetary", "CLV", "Segmen", "Area" )
  print(tampil)
},
  list(lengthMenu = c(5, 10, 15), pageLength = 10)) }

```

Script 5.16 Rancang Bangun Halaman Tabel Data

| ID | Retailer | Length | Recency | Frequency | Monetary | CLV | Segmen | Area |
|-------|---------------|--------|---------|-----------|----------|----------|--------|-----------|
| 10372 | new cell | 111 | 111 | 105 | 1073000 | 0.202412 | 2 | LAROGSIAN |
| 10373 | mlu cell | 215 | 175 | 25 | 1557300 | 0.319113 | 1 | LAROGSIAN |
| 10374 | mlu cell | 102 | 95 | 41 | 17162200 | 0.220873 | 3 | LAROGSIAN |
| 10375 | top cell | 129 | 110 | 130 | 5007000 | 0.181188 | 2 | LAROGSIAN |
| 10376 | amny cell | 239 | 103 | 20 | 4560000 | 0.162089 | 3 | LAROGSIAN |
| 10377 | ranggohe cell | 115 | 99 | 11 | 3258000 | 0.175074 | 3 | LAROGSIAN |
| 10378 | mlu cell | 304 | 132 | 79 | 10683000 | 0.348701 | 2 | LAROGSIAN |
| 10379 | jugal cell | 208 | 148 | 10 | 2600000 | 0.148835 | 3 | LAROGSIAN |
| 10380 | mlu cell | 113 | 113 | 40 | 11788000 | 0.294487 | 3 | LAROGSIAN |
| 10381 | ml cell | 204 | 120 | 20 | 7000000 | 0.107707 | 3 | LAROGSIAN |

Gambar 5.7 Hasil Visual Halaman Tabel Data

Halaman ini sengaja dikosongkan

BAB VI HASIL DAN PEMBAHASAN

Pada bab ini akan menjelaskan hasil dan pembahasan dari proses pembobotan variabel LRFM, perhitungan CLV, proses *clustering* menggunakan metode *K-Means*, dan uji performa *cluster*. Bab ini juga akan membahas analisa hasil dari 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>Length</i> | <i>Recency</i> | <i>Frequency</i> | <i>Monetary</i> |
|-------------------------|----------------------|-----------------------|-------------------------|------------------------|
| <i>Length</i> | 1 | 3 | 0.2 | 0.14 |
| <i>Recency</i> | 0.33 | 1 | 0.25 | 0.17 |
| <i>Frequency</i> | 5 | 4 | 1 | 0.50 |
| <i>Monetary</i> | 7 | 6 | 2 | 1 |

Tabel 6.2 Hasil Kuesioner Responden 2

| <i>Kriteria</i> | <i>Length</i> | <i>Recency</i> | <i>Frequency</i> | <i>Monetary</i> |
|-------------------------|----------------------|-----------------------|-------------------------|------------------------|
| <i>Length</i> | 1 | 4 | 0.2 | 0.2 |
| <i>Recency</i> | 0.25 | 1 | 0.2 | 0.17 |
| <i>Frequency</i> | 5 | 5 | 1 | 0.5 |
| <i>Monetary</i> | 5 | 6 | 2 | 1 |

Tabel 6.3 Hasil Kuesioner Responden 3

| <i>Kriteria</i> | <i>Length</i> | <i>Recency</i> | <i>Frequency</i> | <i>Monetary</i> |
|-------------------------|----------------------|-----------------------|-------------------------|------------------------|
| <i>Length</i> | 1 | 3 | 0.2 | 0.17 |
| <i>Recency</i> | 0.33 | 1 | 0.33 | 0.14 |
| <i>Frequency</i> | 5 | 3 | 1 | 0.5 |
| <i>Monetary</i> | 6 | 7 | 2 | 1 |

Data kuesioner selanjutnya diolah dengan cara merata-rata semua pengisian yang dilakukan oleh responden terhadap perbandingan keempat kriteria variabel LRFM. Hasil rata-rata pengisian kuesioner oleh 3 responden dapat dilihat pada Tabel 6.4.

Tabel 6.4 Rata-rata Pengisian Kuesioner

| <i>Kriteria</i> | <i>Length</i> | <i>Recency</i> | <i>Frequency</i> | <i>Monetary</i> |
|-------------------------|----------------------|-----------------------|-------------------------|------------------------|
| <i>Length</i> | 1 | 3.33 | 0.2 | 0.17 |
| <i>Recency</i> | 0.31 | 1 | 0.26 | 0.16 |
| <i>Frequency</i> | 5 | 4.00 | 1 | 0.50 |
| <i>Monetary</i> | 6.00 | 6.33 | 2.00 | 1 |
| <i>Jumlah</i> | 12.31 | 14.67 | 3.46 | 1.83 |

Kemudian langkah selanjutnya dilakukan normalisasi matriks pada hasil rata-rata nilai setiap kriteria yang akan menghasilkan tabel *normalitation comparisons matrix*. Hasil dari normalisasi matriks ini dapat dilihat pada Tabel 6.5 yang juga terdapat hasil bobot yang didapatkan.

Tabel 6.5 Normalitation Comparisons Matrix

| <i>Kriteria</i> | <i>Length</i> | <i>Recency</i> | <i>Frequency</i> | <i>Monetary</i> | <i>Bobot</i> |
|-------------------------|----------------------|-----------------------|-------------------------|------------------------|---------------------|
| <i>Length</i> | 0.08 | 0.23 | 0.06 | 0.09 | 0.11 |
| <i>Recency</i> | 0.02 | 0.07 | 0.08 | 0.09 | 0.06 |
| <i>Frequency</i> | 0.41 | 0.27 | 0.29 | 0.27 | 0.31 |
| <i>Monetary</i> | 0.49 | 0.43 | 0.58 | 0.55 | 0.51 |

Bobot yang dihasilkan perlu dilakukan uji konsistensi untuk mengetahui apakah kuesioner yang diisi secara konsisten atau tidak. Uji konsistensi ini dilakukan dengan cara perkalian matrix pada *pairwase comparisons* atau hasil dari rata-rata pengisian kuesioner sebelumnya dengan bobot masing-masing kriteria. Berikut ini adalah hasil dari perkalian matrix nilai perbandingan kriteria terhadap masing-masing bobot kriteria pada Tabel 6.6.

Tabel 6.6 Perkalian Matrix Uji Konsistensi AHP

| Kriteria | L | R | F | M | Bobot | Hasil Perkalian Matrix |
|----------|------|------|------|------|-------|------------------------|
| L | 1 | 3.33 | 0.2 | 0.17 | 0.11 | 0.46669658 |
| R | 0.31 | 1 | 0.25 | 0.15 | 0.06 | 0.251373014 |
| F | 5 | 4.33 | 1 | 0.5 | 0.31 | 1.397943403 |
| M | 6 | 6.67 | 2 | 1 | 0.51 | 2.223305233 |

Langkah selanjutnya setelah mendapatkan nilai hasil perkalian matrix antara nilai *pairwise comparisons* kriteria terhadap bobot, maka dilanjutkan dengan mencari nilai **Lambda** (Λ). Berikut ini adalah perhitungan untuk mencari nilai lambda.

$$\Lambda = \frac{0.47 + 0.25 + 1.39 + 2.22}{4} = 4.26$$

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

$$CI = \frac{4.26 - 4}{4 - 1} = 0.087$$

Diperoleh nilai CI sebesar 0.87, berdasarkan hukum uji konsistensi, nilai CI diharuskan untuk dilakukan uji ketidakkosistenan karena CI tidak bernilai sama dengan nol. Mengacu pada Bab II, nilai ketidakkonsistenan 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

$$\frac{0.87}{0.9} = 0.096$$

Diperoleh nilai CR adalah 0.096 atau kurang dari 0.1, maka oleh karena itu hasil kuesioner telah bersifat konsisten. Oleh karena itu, bobot kriteria telah valid dan dapat digunakan pada perhitungan CLV. Berikut ini adalah bobot masing-masing kriteria 0.11, 0.06, 0.31 dan 0.51 berturut-turut untuk *length*, *rency*, *frequency*, dan *monetary*. Sehingga diperoleh temuan bahwa bobot untuk *monetary* memiliki bobot terbesar dibandingkan dengan kriteria lainnya.

6.2 Perhitungan Nilai CLV

Dari nilai bobot tiap kriteria yang diperoleh, langkah selanjutnya adalah melakukan perhitungan nilai CLV. Perhitungan nilai CLV dilakukan dengan mengalikan nilai masing-masing variabel LRFM tiap retailer terhadap bobotnya kemudian dijumlah hasil perkalian tersebut untuk mendapatkan nilai CLV. Berikut ini beberapa hasil perhitungan nilai CLV tiap retailer yang dapat dilihat pada Tabel 6.7 dan hasil lengkapnya pada Lampiran E.

Tabel 6.7 Sebagian Hasil Perhitungan CLV Retailer

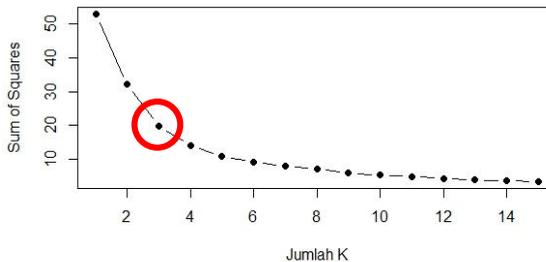
| ID | LBobot | RBobot | FBobot | MBobot | CLV |
|------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 1030 72 | 0.101081 081 | 0.056546 763 | 0.133598 326 | 0.001421 997 | 0.292648 167 |
| 1030 73 | 0.062060 811 | 0.043812 95 | 0.031129 707 | 0.001147 863 | 0.138151 331 |
| 1034 73 | 0.105168 919 | 0.059784 173 | 0.050585 774 | 0.001448 448 | 0.216987 314 |
| 1034 78 | 0.104054 054 | 0.056762 59 | 0.166025 105 | 0.004339 093 | 0.331180 841 |
| 1037 41 | 0.070608 108 | 0.058273 381 | 0.033723 849 | 0.000394 543 | 0.162999 881 |

6.3 Clustering Data

Dalam proses *clustering* data dijelaskan mengenai hasil pentuan nilai K dengan metode Elbow dan proses *clustering* menggunakan metode *K-Means* serta uji performa *cluster*.

6.3.1 Menentukan Nilai K Menggunakan Metode Elbow

Menentukan nilai K dilakukan dengan menggunakan metode Elbow pada *tools* RStudio. Hasil perhitungan metode Elbow ditampilkan dalam plot dua dimensi pada gambar 6.1. Nilai pada sumbu X menggambarkan nilai K yaitu 1 hingga 15. Sedangkan sumbu Y merepresentasikan nilai SSE yang dihasilkan oleh setiap nilai K. Metode Elbow menjelaskan bahwa nilai K yang dipilih adalah nilai rasio terbesar dari penurunan nilai SSE. Pada plot tersebut terlihat bahwa seiring naiknya nilai K, maka nilai SSE akan semakin menurun. Dan berdasarkan Gambar 6.1 penurunan terjadi secara drastis pada nilai K = 3.



Gambar 6.1 Metode Elbow

Pada Tabel 6.8 terlihat bahwa K=3 memiliki nilai terbesar rasio penurunan nilai SSE.

Tabel 6.8 Rasio Penurunan SSE

| Klaster | SSE | Rasio | Klaster | SSE | Rasio |
|---------|-----------|----------|---------|----------|----------|
| 1 | 52.99641 | | 9 | 5.922095 | 2.064059 |
| 2 | 32.120365 | 1.649181 | 10 | 5.310771 | 1.393617 |
| 3 | 19.461936 | 2.305425 | 11 | 4.872111 | 0.832181 |
| 4 | 13.971221 | 1.736546 | 12 | 4.34499 | 1.115546 |
| 5 | 10.809361 | 2.118961 | 13 | 3.872467 | 1.947793 |
| 6 | 9.317186 | 1.090636 | 14 | 3.629873 | 1.725101 |
| 7 | 7.949017 | 1.788192 | 15 | 3.489247 | 0.040303 |
| 8 | 7.183904 | 0.606362 | | | |

6.3.2 Proses Clustering Menggunakan Metode K-Means

Proses *clustering* dengan menggunakan metode K-Means dan *tools* RStudio menghasilkan sebuah tabel pengelompokan retailer sesuai *cluster* yang disimpan dalam bentuk CSV agar mudah dalam proses pengelolaan analisa segmentasi maupun visualisasi. Hasil *cluster* dalam bentuk tabel akan terlihat pada Tabel 6.9.

Tabel 6.9 Sebagian Hasil Clustering

| ID | NAME.RET | L | R | F | M | Cluster |
|--------|-----------------|-------|-------|-------|-------|---------|
| 103072 | new cell | 0.919 | 0.942 | 0.431 | 0.003 | 2 |
| 103073 | mila cell | 0.564 | 0.730 | 0.100 | 0.002 | 1 |
| 103473 | eka cell | 0.956 | 0.996 | 0.163 | 0.003 | 3 |
| 103478 | top cell | 0.946 | 0.946 | 0.536 | 0.009 | 2 |
| 103741 | army cell | 0.642 | 0.971 | 0.109 | 0.001 | 3 |
| 103742 | ronggolawe cell | 0.902 | 0.986 | 0.067 | 0.000 | 3 |
| 103743 | mitra cell | 0.861 | 0.867 | 0.322 | 0.003 | 2 |
| 103745 | jagad cell | 0.801 | 0.824 | 0.033 | 0.000 | 3 |
| 103747 | rifqi cell | 0.892 | 0.935 | 0.159 | 0.002 | 3 |

Kemudian data dipisahkan pada masing-masing *clusternya*. Diketahui *cluster* pertama terdiri dari 93 retailer, *cluster* kedua terdiri dari 133 retailer dan *cluster* ketiga memiliki anggota 398 retailer.

6.3.3 Uji Performa Cluster

Pengujian performa cluster menggunakan metode SSE yang dibantu dengan *tools* RStudio. Berikut ini adalah hasil dari nilai SSE pada setiap titik yang terlihat pada Tabel 6.10.

Tabel 6.10 menunjukkan adanya nilai rasio penurunan SSE, dimana nilai rasio penurunan terbesar berada pada nilai $K = 3$ sehingga nilai SSE yang didapatkan adalah 19.461936.

Tabel 6.10 Hasil Uji SSE

| Klaster | SSE | Selisih | Rasio |
|---------|-----------|----------|----------|
| 1 | 52.99641 | | |
| 2 | 32.120365 | 20.87605 | 1.649181 |
| 3 | 19.461936 | 12.65843 | 2.305425 |
| 4 | 13.971221 | 5.990715 | 1.736546 |
| 5 | 10.809361 | 3.16186 | 2.118961 |
| 6 | 9.317186 | 1.492175 | 1.090636 |
| 7 | 7.949017 | 1.368169 | 1.788192 |
| 8 | 7.183904 | 0.765113 | 0.606362 |
| 9 | 5.922095 | 1.261809 | 2.064059 |
| 10 | 5.310771 | 0.611324 | 1.393617 |
| 11 | 4.872111 | 0.43866 | 0.832181 |
| 12 | 4.34499 | 0.527121 | 1.115546 |
| 13 | 3.872467 | 0.472523 | 1.947793 |
| 14 | 3.629873 | 0.242594 | 1.725101 |
| 15 | 3.489247 | 0.140626 | 0.040303 |

6.4 Analisa Hasil Clustering

Hasil dari proses *clustering* menghasilkan tiga segmen dengan jumlah total retailer sebanyak 624 anggota. Masing-masing segmen memiliki karakteristik yang berbeda-beda dimana karakteristik tersebut dipengaruhi oleh variabel *length*, *recency*, *frequency*, dan *monetary*. Keempat variabel tersebut juga mempengaruhi nilai CLV tiap retailer. Nilai CLV inilah yang dijadikan sebagai ukuran tingkat loyalitas setiap retailer tiap segmen.

6.4.1 Pemingkatan Segmen

Pemingkatan *cluster* atau segmen dilakukan untuk menentukan segmen terbaik hingga segmen terburuk. Tingkatan segmen ditentukan berdasarkan rata-rata nilai CLV

yang diperoleh oleh tiap retailer dalam satu segmen. Segmen yang memiliki nilai CLV terbesar menunjukkan bahwa segmen tersebut beranggotakan retailer yang loyal dan layak disebut sebagai segment terbaik, begitupun sebaliknya.

Tabel 6.11 Peringkat Segmen Berdasarkan Nilai CLV

| Segmen | CLV | Peringkat Segmen |
|---------------|------------|-------------------------|
| 1 | 0.116 | 3 |
| 2 | 0.323 | 1 |
| 3 | 0.199 | 2 |

Tabel 6.10 merupakan hasil perhitungan rata-rata nilai CLV tiap segmen. Dari nilai tersebut diperoleh bahwa segmen 2 merupakan segmen yang terbaik dan segmen 1 merupakan segmen yang terburuk.

6.4.2 Analisa Karakteristik Segmen 1

Segmen 1 terdiri dari 93 anggota retailer. Rata-rata nilai CLV segmen 1 adalah 0.12 atau nilai terendah dari kedua segmen lainnya, sehingga segmen 1 dapat dikategorikan sebagai segmen terburuk. Rentang jumlah uang yang dikeluarkan oleh retailer pada segmen 1 antara Rp. 448.000 hingga Rp. 337.407.000 dengan rata-rata Rp 11.218.739. Hal ini menunjukkan bahwa retailer mengeluarkan uang yang relatif rendah pada setiap transaksi, meskipun terdapat retailer yang mengeluarkan uang dengan jumlah besar pada sekali transaksi.

Rentang jumlah transaksi atau frekuensi transaksi yang dilakukan para retailer pada segmen 1 antara 2 hingga 47 kali, dengan rata-rata 16 transaksi tiap retailer pada saat periode penelitian. Hal ini menunjukkan bahwa para retailer pada segmen ini jarang melakukan transaksi pembelian. Kemudian rentang waktu transaksi terakhir kali dilakukan adalah sekitar 95 hingga 373 hari atau sekitar 3 hingga 6 bulan yang lalu dihitung dari periode penelitian pada bulan Maret 2018. Dilihat dari berapa lama retailer telah melakukan transaksi terdapat rentang waktu selama 49 hingga 273 hari dengan kata lain terdapat retailer yang hanya melakukan transaksi hanya selama

49 hari dan terdapat retailer yang telah melakukan transaksi dengan perusahaan selama 273 hari dan rata-rata para retailer pada segmen 1 telah melakukan transaksi dengan perusahaan selama 197 hari. Karakteristik segmen 1 lebih jelasnya dapat dilihat pada Tabel 6.12.

Tabel 6.12 Karakteristik Retailer Segmen 1

| | | |
|-------------------------|------------------|--|
| Segmen | 1 | |
| Peringkat Segmen | 3 | |
| Jumlah Anggota | 93 | |
| Karakteristik | Length | 49 – 273 hari (Rata-rata: 197 hari) |
| | Recency | 95 – 373 hari (Rata-rata: 179 hari) |
| | Frequency | 2 – 47 kali (Rata-rata: 16 kali) |
| | Monetary | Rp. 448.000 – Rp. 337.407.000 (Rata-rata: Rp. 11.218.739) |

6.4.3 Analisa Karakteristik Segmen 2

Segmen 2 terdiri dari 133 anggota retailer. Rata-rata nilai CLV pada segmen ini adalah 0.32 atau tertinggi dari kedua segmen lainnya, sehingga segmen 2 dapat dikategorikan sebagai segmen terbaik. CLV yang tinggi merepresentasikan bahwa pelanggan yakni para retailer memiliki tingkat loyal yang tinggi terhadap perusahaan. Rentang jumlah uang yang dilakukan selama transaksi oleh para retailer antara Rp. 12.424.000 hingga Rp. 5.832.824.162, dengan rata-rata Rp. 155.718.628. Hal ini menunjukkan bahwa para retailer mengeluarkan uang yang besar pada saat transaksi.

Rentang jumlah transaksi yang dilakukan pada segmen ini antara 74 hingga 241 kali, dengan rata-rata transaksi sebanyak 116 kali tiap retailer. Hal ini menunjukkan bahwa retailer sangat sering melakukan transaksi pada perusahaan. Rentang waktu terakhir kali dilakukan oleh retailer adalah 95 hingga 162 hari. Hal ini menunjukkan para retailer aktif melakukan transaksi sekitar 3 bulan yang lalu dimana apabila dihitung mundur dari

periode penelitian pada bulan maret maka retailer terakhir yang melakukan transaksi sekitar bulan Desember 2017. Kemudian dapat diketahui rentang waktu para retailer telah melakukan transaksi pada perusahaan adalah sekitar 258 hingga 345 hari dengan rata-rata 323 hari, hal ini menunjukkan bahwa para retailer sangat loyal melakukan transaksi pada perusahaan. Karakteristik segmen 2 lebih lengkap dapat dilihat pada Tabel 6.13.

Tabel 6.13 Karakteristik Retailer Segmen 2

| | | |
|-------------------------|------------------|--|
| Segmen | | 2 |
| Peringkat Segmen | | 1 |
| Jumlah Anggota | | 133 |
| Karakteristik | Length | 258 – 345 hari (Rata-rata: 323 hari) |
| | Recency | 95 – 162 hari (Rata-rata: 98 hari) |
| | Frequency | 74 – 241 kali (Rata-rata: 116 kali) |
| | Monetary | Rp. 12.424.000 – Rp. 5.832.824.162 (Rata-rata: Rp.155.718.628) |

6.4.4 Analisa Karakteristik Segmen 3

Segmen 3 memiliki anggota atau retailer sebanyak 398 retailer, dimana rata-rata nilai CLV pada segmen ini adalah 0.19 atau menempati peringkat kedua dari 3 segmen, sehingga segmen 2 dikategorikan dalam segmen menengah. Nilai CLV yang sedang menunjukkan bahwa retailer pada segmen ini cukup loyal terhadap perusahaan.

Untuk rentang uang yang dikeluarkan pada saat transaksi, para retailer dalam segmen 3 ini mengeluarkan uang dalam rentang Rp. 1.250.000 hingga Rp. 2.317.734.080 dengan rata-rata Rp. 37.955.185. Hal ini menunjukkan bahwa retailer pada segmen 3 ini mengeluarkan uang yang relatif tinggi pada setiap transaksi.

Rentang jumlah transaksi yang dilakukan oleh para retailer segmen 3 ini sekitar 8 hingga 74 kali transaksi dengan rata-rata transaksi 33 kali tiap retailernya, dapat dikatakan bahwa para retailer pada segmen ini cukup sering melakukan transaksi pada perusahaan. Kemudian diketahui bahwa rentang transaksi terakhir yang dilakukan oleh para retailer segmen ini adalah 95 hingga 160 hari yang lalu terhitung dari tanggal 16 Maret 2018 dan diketahui juga para retailer telah melakukan transaksi terhadap perusahaan dengan rentang waktu 228 hingga 345 hari dengan rata-rata lama bertransaksi dengan perusahaan sebanyak 315 hari. Hal ini menunjukkan bahwa para retailer pada segmen ini cukup loyal bertransaksi pada perusahaan. Karakteristik segmen 3 lebih lengkap dapat dilihat pada Tabel 6.14.

Tabel 6.14 Karakteristik Retailer Segmen 3

| | | |
|-------------------------|------------------|--|
| Segmen | 3 | |
| Peringkat Segmen | 2 | |
| Jumlah Anggota | 398 | |
| Karakteristik | Length | 228 – 345 hari (Rata-rata: 315 hari) |
| | Recency | 95 – 160 hari (Rata-rata: 106 hari) |
| | Frequency | 8 – 74 kali (Rata-rata: 33 kali) |
| | Monetary | Rp. 1.250.000 – Rp. 2.317.734.080 (Rata-rata: Rp. 37.955.185) |

6.4.5 Perbandingan Karakteristik Segmen

Perbandingan karakteristik ketiga segmen dapat dilihat pada Tabel 6.15. Dari tabel tersebut terlihat bahwa variabel *length* terjadi *overlap* data yang terbukti dengan adanya rentang data yang sama pada ketiga segmen. Hal ini dikarenakan pemberian bobot pada variabel ini relatif kecil dan dari perbandingan ini diketahui bahwa retailer pada segmen 2 dan 3 merupakan retailer yang loyal dibuktikan dengan lamanya waktu yang telah dilalui oleh retailer dalam bertransaksi dengan perusahaan.

Kemudian pada variabel *recency* diketahui dari perbandingan ketiga segmen bahwa terdapat *overlap* data hal ini terbukti dengan adanya rentang nilai yang sama. Namun dari perbandingan ini diketahui bahwa tiap segmen memiliki catatan transaksi terkakhir 95 hari yang lalu dari periode penelitian dan disimpulkan bahwa dari segi variabel *recency*, segmen 3 merupakan segmen yang terbaik.

Pada variabel *frequency* dimana variabel ini memiliki bobot terbesar kedua diketahui bahwa segmen kedua merupakan segmen yang terbaik dimana segmen ini memiliki jumlah frekuensi transaksi paling besar dibandingkan dengan segmen 1 maupun segmen 3.

Terakhir, variabel *monetary* yang merupakan variabel dengan bobot terbesar atau paling penting dalam menentukan peringkat retailer, ditemukan bahwa segmen 2 merupakan segmen terbaik dari segi variabel *monetary* dengan catatan *monetary* paling tinggi dan diikuti dengan segmen 3 dan yang terakhir segmen 1 yang paling rendah.

Tabel 6.15 Perbandingan Karakteristik 3 Segmen

| | Segmen 1 | Segmen 2 | Segmen 3 |
|-----------------------|------------------------------|-----------------------------------|----------------------------------|
| Peringkat | 3 | 1 | 2 |
| Jumlah Anggota | 93 | 133 | 398 |
| Length | 49 – 273 hari | 258 – 345 hari | 228 – 345 hari |
| Recency | 95 – 373 hari | 95 – 162 hari | 95 – 160 hari |
| Frequency | 2 – 47 kali | 74 – 241 kali | 8 – 74 kali |
| Monetary | Rp. 448.000 – Rp.337.407.000 | Rp. 12.424.000 – Rp.5.832.824.162 | Rp. 1.250.000 – Rp.2.317.734.080 |

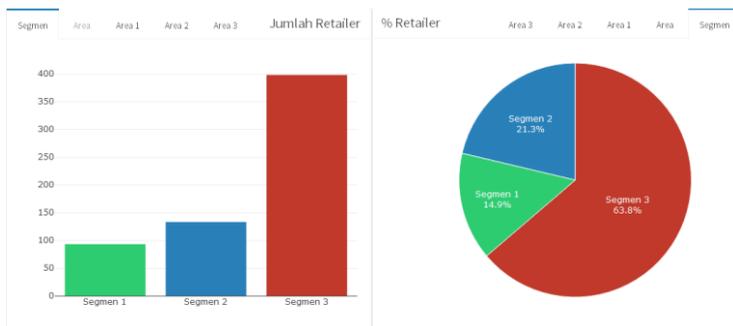
6.5 Analisa Hasil Visualisasi

Hasil *clustering* atau segmentasi kemudian ditampilkan berbasis web dengan bentuk beberapa grafik untuk mempermudah analisa perilaku retailer tiap segmennya. Berikut ini adalah analisa data dari perilaku retailer dalam bentuk beberapa visualiasi grafik.

6.5.1 Bar Chart dan Pie Chart

Grafik ini digunakan untuk menjelaskan persebaran dan persentase persebaran retailer pada tiap segmen dan area. Grafik ini dibuat dengan menggunakan bantuan *package ggplot()*.

Persebaran dan persentase persebaran retailer pada masing-masing segmen ditampilkan dengan grafik *bar chart* dan *pie chart* seperti yang nampak pada Gambar 6.2.

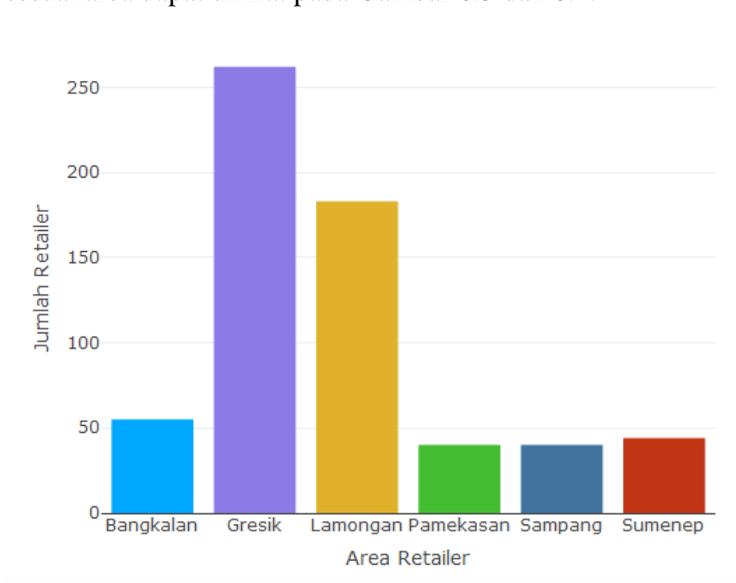


Gambar 6.2 Grafik Bar Chart dan Pie Chart

Segmen 1 yang digambarkan dengan warna hijau memiliki persentase retailer sebesar 14.9% dari keseluruhan retailer atau sebanyak 93 retailer. Segmen 2 yang digambarkan dengan warna biru memiliki anggota retailer sebanyak 133 retailer dengan kata lain memiliki persentase persebaran retailer sebesar 21.3% dari total keseluruhan retailer. Selanjutnya adalah segmen 3 yang dipresentasikan dengan warna grafik warna merah dengan memiliki persentase anggota retailer terbesar yakni sebesar 63.8% atau dengan jumlah retailer sebanyak 398 retailer.

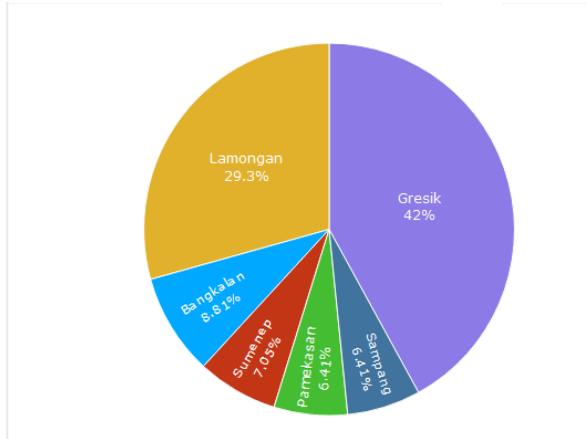
Selain menunjukkan jumlah dan presentase persebaran retailer tiap segmen, grafik *bar chart* dan *pie chart* menunjukkan jumlah dan persentase persebaran retailer tiap areanya. Diketahui bahwa untuk keseluruhan area, wilayah Gresik menempati posisi teratas dalam hal jumlah dan persentase

persebaran retailer dengan jumlah retailer 262 atau 42% dari keseluruhan retailer yang tersebar di 6 area, kemudian diikuti oleh area Lamongan dengan jumlah retailer sebanyak 183 atau 29.3% dari total keseluruhan retailer. Untuk posisi ketiga ditempati oleh area Bangkalan dengan total retailer 55 atau 8.81% dari total keseluruhan retailer, posisi keempat terdapat area Sumenep dengan total retailer sebanyak 44 atau 7.05% dari keseluruhan retailer. Kemudian untuk posisi kelima dan keenam ditempati oleh wilayah Pamekasan dan Sampang dengan sama-sama memiliki total retailer sebanyak 40 retailer atau sekitar 6.41%. Jumlah dan persentase persebaran retailer sesuai area dapat dilihat pada Gambar 6.3 dan 6.4.



Gambar 6.3 Persebaran Jumlah Retailer

Grafik *bar chart* dan *pie chart* juga menampilkan visualisasi dari jumlah dan persentase persebaran area tiap segmennya. Untuk lebih jelasnya berikut ini uraian jumlah dan persentase persebaran retailer dalam bentuk beberapa poin penjelasan.



Gambar 6.4 Persentase Persebaran Retailer

a. *Segmen 1*

Diketahui bahwa segmen 1 merupakan segmen dengan jumlah dan persebaran retailer terkecil dibandingkan dengan kedua segmen lainnya. Dengan total persebaran retailer sebanyak 93 retailer, diketahui terdapat beberapa komposisi persebaran retailer dari tiap area yang ada dengan area Gresik menempati posisi pertama persebaran retailer terbanyak dan area Sampang menempati posisi terakhir sebagai area persebaran paling sedikit pada segmen 1 ini. Untuk lebih jelasnya, berikut ini keterangan persebaran retailer pada segmen 1 ini dalam bentuk beberapa poin:

1. Area Gresik : 41 retailer atau 44.1% dari total persebaran retailer segmen 1
2. Area Lamongan : 26 retailer atau 28% dari total persebaran retailer segmen 1
3. Area Bangkalan : 15 retailer atau 16.1% dari total persebaran retailer segmen 1
4. Area Pamekasan : 4 retailer atau 4.3% dari total persebaran retailer segmen 1

5. Area Sumenep : 4 retailer atau 4.3% dari total persebaran retailer segmen 1
6. Area Sampang : 3 retailer atau 3.23% dari total persebaran retailer segmen 1

b. Segmen 2

Segmen 2 yang merupakan segmen terbaik menempati posisi kedua dalam jumlah dan persentase persebaran retailer. Pada segmen ini area Lamongan menempati posisi teratas dalam jumlah dan persentase persebaran retailer pada segmen ini dan area Pamekasan menempati posisi paling sedikit dalam hal jumlah anggota retailer. Berikut ini adalah jumlah dan persentase persebaran retailer tiap area pada segmen 2:

1. Area Lamongan : 63 retailer atau 47.4% dari total persebaran retailer segmen 2
2. Area Gresik : 43 retailer atau 32.3% dari total persebaran retailer segmen 2
3. Area Bangkalan : 12 retailer atau 9.02% dari total persebaran retailer segmen 2
4. Area Sumenep : 7 retailer atau 5.26% dari total persebaran retailer segmen 2
5. Area Sampang : 5 retailer atau 3.76% dari total persebaran retailer segmen 2
6. Area Pamekasan : 3 retailer atau 2.26% dari total persebaran retailer segmen 2

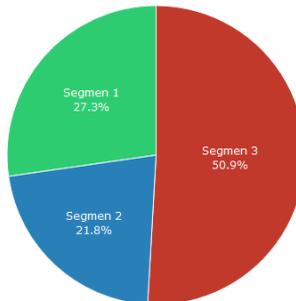
c. Segmen 3

Selanjutnya adalah segmen yang menduduki peringkat pertama dalam hal jumlah dan persentase persebaran retailer yakni segmen 3, dengan jumlah total persebaran retailer pada segmen ini adalah 398 retailer dengan area Gresik yang menduduki posisi pertama dan area Bangkalan posisi terakhir dalam jumlah persebaran retailer. Berikut ini adalah jumlah dan persentase persebaran retailer tiap area pada segmen ini.

1. Area Gresik : 178 retailer atau 44.7% dari total persebaran retailer segmen 3
2. Area Lamongan : 94 retailer atau 23.6% dari total persebaran retailer segmen 3
3. Area Pamekasan : 33 retailer atau 8.29% dari total persebaran retailer segmen 1
4. Area Sumenep : 33 retailer atau 8.29% dari total persebaran retailer segmen 3
5. Area Sampang : 32 retailer atau 8.04% dari total persebaran retailer segmen 3
6. Area Bangkalan : 28 retailer atau 7.04 % dari total persebaran retailer segmen 3

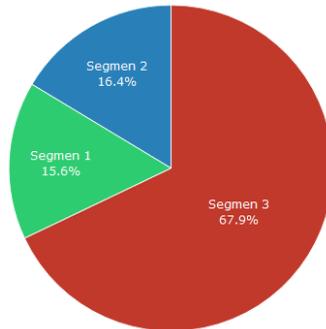
Berikut ini adalah detail persebaran retailer tiap area berdasarkan persebaran terhadap keanggotaan segmen.

a. Bangkalan



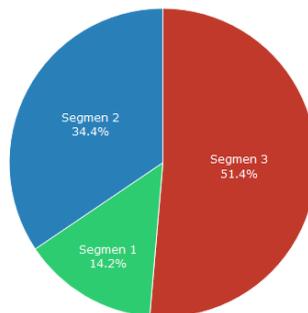
Gambar 6.5 Presentase Persebaran Retailer Area Bangkalan

Dilihat dari Gambar 6.5 diketahui bahwa persebaran retailer area bangkalan paling banyak menempati segmen 3 dengan persentase 50.9% kemudian terdapat pada segmen 2 dengan 27.3% dan terakhir menempati segmen 2 dengan 21.8%.

b. Gresik

Gambar 6.6 Persentase Persebaran Retailer Area Gresik

Diketahui dari grafik pada Gambar 6.6, persebaran retailer terbanyak pada area ini terdapat pada segmen 3 dengan persentase 67.9% diikuti dengan segmen 2 dengan 16.4% dan segmen 1 dengan persentase persebaran retailer 15.6% dari keseluruhan retailer yang berasal dari area Gresik

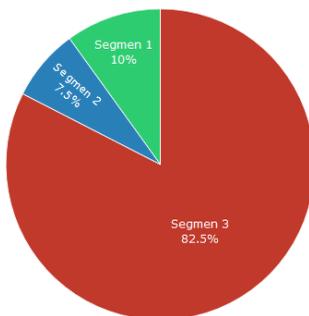
c. Lamongan

Gambar 6.7 Persentase Persebaran Retailer Area Lamongan

Pada Gambar 6.7 terlihat bahwa, segmen 3 menjadi persebaran retailer terbesar pada area Lamongan dengan persentase 51.4%, diikuti dengan segmen 2

dengan 34.4% dan terakhir segmen 1 dengan persentase 14.2%.

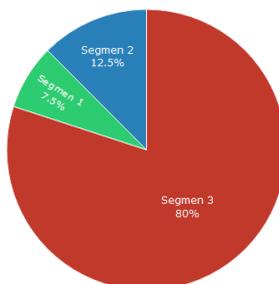
d. *Pamekasan*



Gambar 6.8 Persentase Persebaran Retailer Area Pamekasan

Melihat grafik pada Gambar 6.8, segmen 3 menjadi segmen dengan persebaran retailer terbesar pada area Pamekasan dengan persentase 82.5% dan diikuti dengan segmen 1 dengan persentase persebaran 10% dan segmen 2 dengan persentase 7.5% dari keseluruhan retailer pada area Pamekasan.

e. *Sampang*

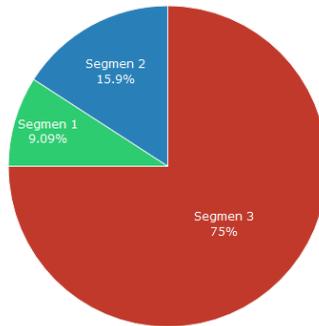


Gambar 6.9 Persentase Persebaran Retailer Area Sampang

Pada Gambar 6.9, terlihat bahwa segmen 3 dengan persentase 80% menjadikan segmen dengan persebaran terbanyak pada area Sampang. Diikuti oleh Segmen 2

dengan persentase 12.5% menempati posisi kedua dan segmen 1 dengan persentase 7.5% yang menjadikan segmen 1 sebagai segmen dengan persebaran retailer terkecil dari area Sampang.

f. *Sumenep*



Gambar 6.10 Persentase Persebaran Retailer Area Sumenep

Diketahui dari Gambar 6.10, segmen 3 menempati posisi pertama dalam presentase persebaran retailer terbanyak pada area Sumenep dengan 75%, diikuti oleh segmen 2 dengan persentase 15.9% dan segmen 1 dengan persentase persebaran retailer 9.09%

Dari visualisasi *bar chart* dan *pie chart* terhadap persebaran retailer, dapat diketahui persebaran retailer tiap area dan segmennya, untuk lebih jelasnya lihat Tabel 6.16.

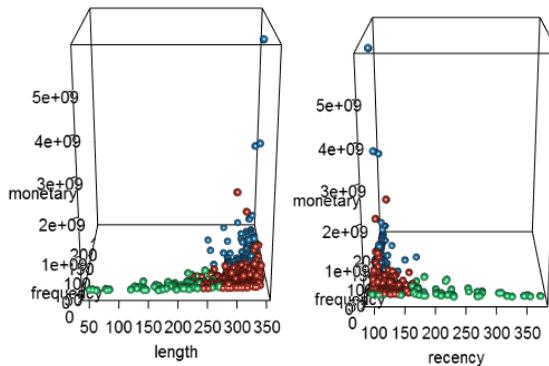
Tabel 6.16 Persebaran Retailer

| Area | Segmen 1 | Segmen 2 | Segmen 3 |
|-----------|----------|----------|----------|
| Bangkalan | 15 | 12 | 28 |
| Gresik | 41 | 43 | 178 |
| Lamongan | 26 | 63 | 94 |
| Pamekasan | 4 | 3 | 33 |
| Sampang | 3 | 5 | 32 |
| Sumenep | 4 | 7 | 33 |

Dapat diketahui dari Tabel 6.15, persebaran retailer dari keseluruhan area banyak menyebar pada segmen 3, kemudian diikuti penyebaran pada segmen 2, kecuali area Bangkalan dan Pamekasan yang persebaran retailer terbanyak kedua pada segmen 1.

6.5.2 Scatter 3D Plot

Grafik 3D *scatter plot* digunakan untuk melihat persebaran retailer dan bagaimana retailer tersebut berkelompok membentuk beberapa segmen berdasarkan variabel LRFM. Sumbu X mewakili variabel *recency/length*, sumbu Y mewakili variabel *Frequency* dan variabel Z mewakili variabel *monetary*. Kemudian diketahui segmen 1 digambarkan dengan warna hijau, segmen 2 dengan warna biru dan segmen 3 dipresentasikan dengan warna merah. Berikut ini hasil visualisasi grafik 3D *scatter plot* ditampilkan pada Gambar 6.11

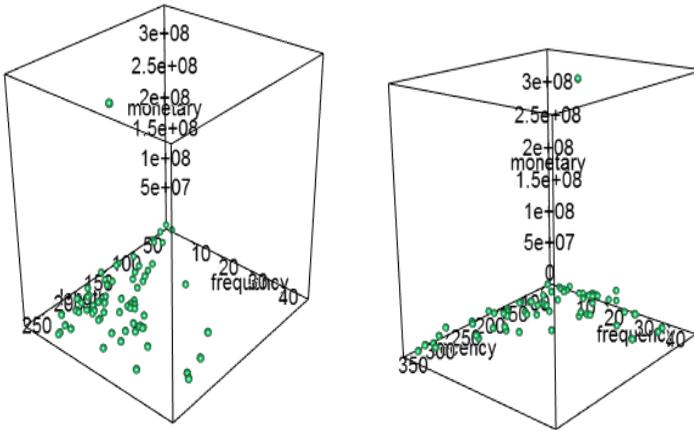


Gambar 6.11 Grafik Scatter 3D Plot Keseluruhan Segmen

Terlihat dari gambar tersebut dapat diketahui bahwa retailer pada segmen 1 cenderung tersebar menjauhi titik 0 terhadap *recency* dan mendekati titik 0 terhadap *length*. Kemudian retailer pada segmen 2 terlihat cenderung mengarah pada variabel *monetary* yang semakin tinggi dan mengarah pada variabel *frequency* yang semakin menyebar. Retailer pada

segmen 3 cenderung berkumpul mendekati titik 0 pada variabel *frequency* dan *recency*.

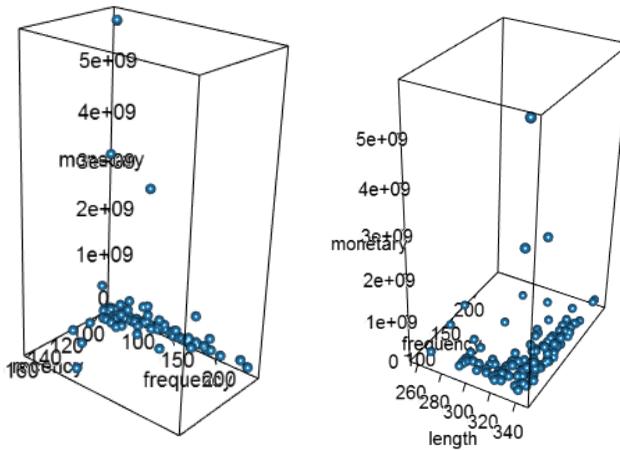
Gambar 6.12 menjelaskan lebih rinci mengenai persebaran retailer pada segmen 1. Dari gambar tersebut dapat dilihat bahwa persebaran nilai *recency* dan *length* cenderung tinggi dan persebaran nilai *frequency* dan *monetary* yang rendah.



Gambar 6.12 Grafik Scatter 3D Plot Segmen 1

Hal ini menunjukkan bahwa segmen 1 memiliki karakteristik retailer yang memiliki nilai *trasanksi* yang rendah.

Selanjutnya pada Gambar 6.13 menjelaskan lebih rinci mengenai persebaran nilai LRFM terhadap perilaku retailer pada segmen 2. Dapat dilihat dari gambar tersebut bahwa persebaran nilai *recency* yang rendah, diikuti dengan persebaran nilai *length*, *frequency* dan *monetary* yang cenderung tinggi.

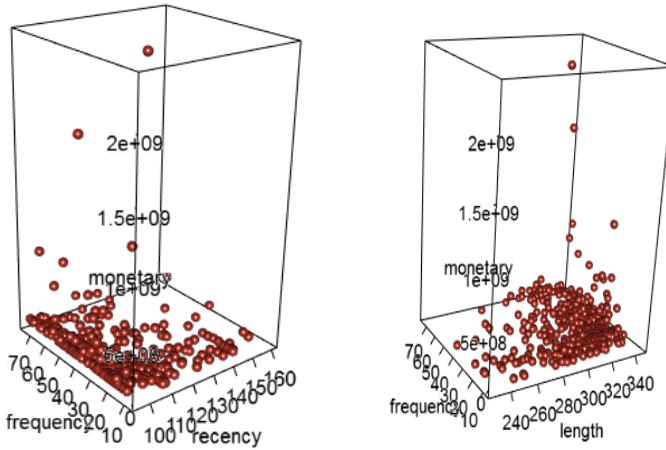


Gambar 6.13 Grafik Scatter 3D Plot Segmen 2

Oleh karena itu, dapat disimpulkan bahwa perilaku retailer pada segmen 2 ini memiliki karakteristik sangat sering melakukan transaksi dengan nilai yang cukup besar.

Pada Gambar 6.14 menunjukkan karakteristik perilaku retailer pada segmen 3. Dapat dilihat bahwa perilaku retailer pada segmen ini memiliki nilai *recency* yang cukup rendah meskipun masih terdapat retailer memiliki nilai *recency* yang tinggi, kemudian diikuti dengan nilai *length* yang cukup tinggi yang menandakan retailer dapat dikatakan cukup loyal. Selanjutnya terlihat nilai *frequency* yang cukup tinggi dan diikuti dengan nilai *monetary* yang tidak terlalu besar walaupun terdapat nilai *monetary* yang besar.

Diliat dari perilaku retailer pada segmen 3 ini, dapat disimpulkan bahwa retailer pada segmen ini memiliki karakteristik yang cukup baik dalam melakukan transaksi meskipun dengan nilai transaksi yang tidak cukup besar.



Gambar 6.14 Grafik Scatter 3D Plot Segmen 3

6.5.3 Box and Whisker

Grafik ini digunakan untuk menunjukkan rentang persebaran nilai tiap segmennya. Pada grafik *box and whisker* 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*.

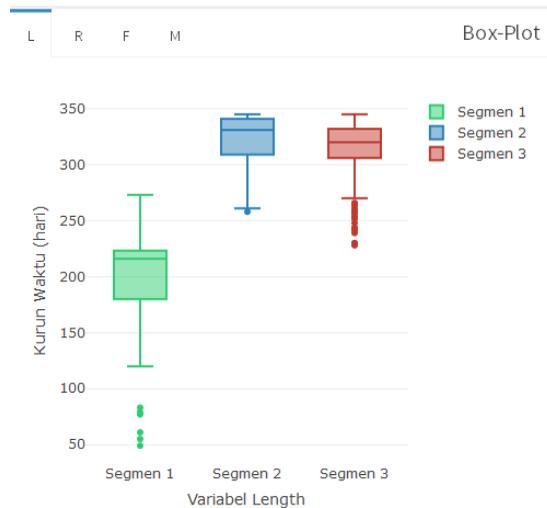
Tabel 6.17 menampilkan nilai minimal, kuartil 1, median, kuartil 3 dan nilai maksimum dari variabel *length* setiap segmennya.

Tabel 6.17 Range Variabel Length

| Var: L | Segmen 1 | Segmen 2 | Segmen 3 |
|--------|----------|----------|----------|
| Max | 273 | 345 | 345 |
| Q3 | 223,25 | 341 | 332 |
| Median | 216 | 331 | 320 |
| Q1 | 180 | 309 | 306 |
| Min | 120 | 261 | 270 |

Berdasarkan Gambar 6.15, berikut ini adalah analisa *box and whisker* untuk variabel *length*:

- *Box and whisker* segmen 1 berada di bawah *box and whisker* segmen lainnya yang berarti bahwa nilai *length* segmen 1 memiliki nilai *length* atau waktu telah bertransaksi dengan perusahaan dengan waktu yang pendek.
- *Box and whisker* segmen 1 lebih panjang dibandingkan dengan segmen lainnya yang menandakan bahwa nilai *length* pada segmen 1 lebih beragam.
- Segmen 2 memiliki *box and whisker* dengan posisi yang sama dengan segmen 3 yang menandakan nilai *length* terjadi persamaan, namun segmen 2 memiliki nilai *length* yang lebih beragam dibandingkan dengan segmen 3.
- Terdapat *outlier* pada segmen 2 dan segmen 3 yang terletak dibawah nilai minimum *length* pada segmen 2 dan segmen 3.



Gambar 6.15 Grafik Box Plot Variabel Length

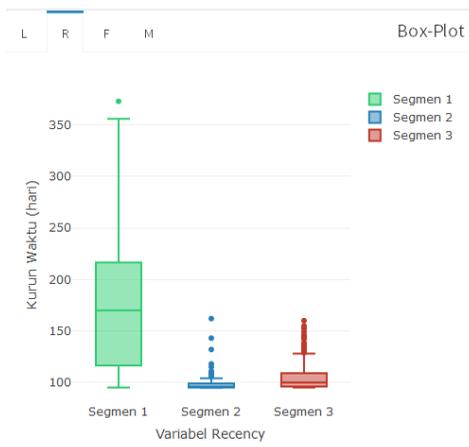
Tabel 6.18 menunjukkan nilai minimal, kuartil 1, median, kuartil 3 dan nilai maksimum dari variabel *recency* pada setiap segmennya. Berikut ini adalah analisa untuk variabel *recency*:

- *Box and whisker* pada segmen 1 lebih panjang dan berada diatas *box and whisker* segmen lainnya yang mengartikan bahwa segmen ini memiliki selisih waktu trasanksi yang beragam dan trasanksi terakhir terjadi pada waktu yang sangat lampau.

Tabel 6.18 Range Variabel Recency

| Var: R | Segmen 1 | Segmen 2 | Segmen 3 |
|--------|----------|----------|----------|
| Max | 356 | 104 | 128 |
| Q3 | 216.5 | 99 | 109 |
| Median | 170 | 95 | 96 |
| Q1 | 116.5 | 95 | 95 |
| Min | 95 | 95 | 95 |

- Terdapat titik-titik pada seluruh segmen yang memiliki arti bahwa terdapat *outlier* yang memiliki nilai lebih dari nilai maksimum variabel *recency*.
- Diketahui bahwa pada segmen 2 dan segmen 3 nilai kuartil 1 sama dengan nilai minimal dari nilai *recency*.
- Segmen 2 memiliki *box and whisker* terpendek yang mengartikan bahwa nilai *recency* pada segmen 2 paling seragam.



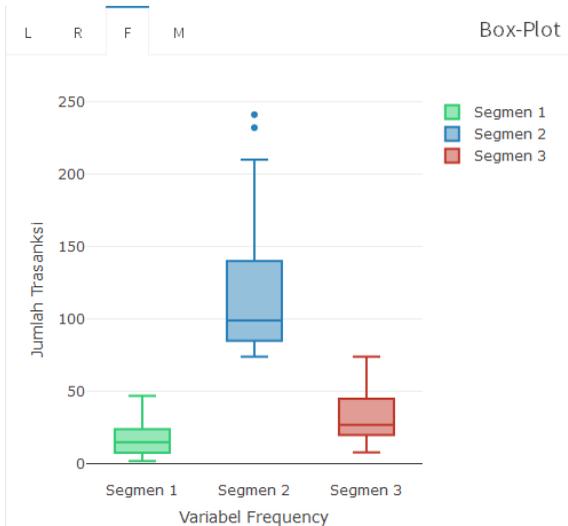
Gambar 6.16 Grafik Box Plot Variabel Recency

Tabel 6.19 menunjukkan nilai minimal, kuartil 1, median, kuartil 3 dan nilai maksimum dari variabel *frequency* pada setiap segmennya. Berikut ini adalah analisa untuk variabel *frequency* yang didapat dari Gambar 6.17.

Tabel 6.19 Range Variabel Frequency

| Var: <i>F</i> | Segmen 1 | Segmen 2 | Segmen 3 |
|---------------|----------|----------|----------|
| Max | 47 | 210 | 74 |
| Q3 | 24 | 140 | 45 |
| Median | 15 | 99 | 27 |
| Q1 | 7.75 | 85 | 20 |
| Min | 2 | 74 | 8 |

- Dapat dilihat bahwa terdapat titik pada segmen 2, hal ini menunjukkan bahwa terdapat *outlier* dimana nilai dari *outlier* tersebut lebih dari nilai maksimum.
- Segmen 1 berada dibawah segmen lainnya yang mengartikan bahwa nilai *frequency*nya lebih rendah dari segmen lainnya. Berbeda dengan segmen 2, yang memiliki *box and whisker* yang berada diatas kedua segmen lainnya yang menandakan bahwa nilai *frequency* pada segmen 2 lebih besar dari yang lain.
- Segmen 2 memiliki panjang *box plot* yang lebih panjang dari yang lain, yang berarti nilai *frequency* pada segmen 2 lebih beragam dari segmen lainnya.
- *Box* bagian atas segmen 3 lebih panjang dibandingkan dengan bagian bawah mengartikan bahwa data diatas nilai *mean*/median memiliki nilai yang beragam.
- Segmen 1 memiliki nilai *frequency* yang seragam dibandingkan dengan segmen lainnya, terlihat dari *box plot* segmen 1 yang paling terkecil dibandingkan dengan segmen 2 maupun segmen 3.



Gambar 6.17 Grafik Box Plot Variabel Frequency

Selanjutnya Tabel 6.20 akan menunjukkan nilai minimum, kuartil 1, median, kuartil 3 dan maksimum dari variabel monetary setiap segmennya.

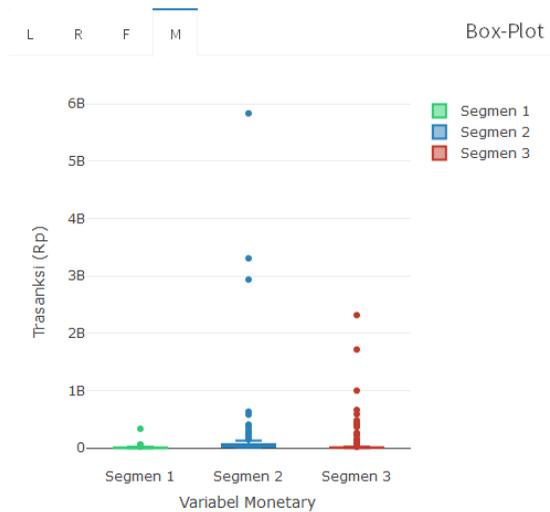
Tabel 6.20 Range Variabel Monetary

| Var: M | Segmen 1 | Segmen 2 | Segmen 3 |
|--------|------------|------------|-------------|
| Max | 23.7 juta | 130.6 juta | 26 .15 juta |
| Q3 | 12.14 juta | 68 juta | 12.69 juta |
| Median | 2.84 juta | 29.7 juta | 6.58 juta |
| Q1 | 1.45 juta | 22.8 juta | 3.54 juta |
| Min | 448 ribu | 12.42 juta | 1.25 juta |

Berikut ini analisa variabel *monetary* yang didapatkan dari Gambar 6.18:

- Semua segmen terdapat titik pada bagian atas yang berarti bahwa semua segmen memiliki nilai *outlier* yang lebih dari nilai maksimal atau lebih dari $3/2$ IQ.
- Segmen 1 memiliki *box plot* yang terpendek yang mengartikan bahwa nilai yang ada pada segmen tersebut yang paling seragam

- Segmen 2 memiliki *box plot* paling tinggi yang mengartikan bahwa segmen ini memiliki nilai *monetary* yang tertinggi dari kedua segmen lainnya, sedangkan segmen 1 memiliki *box plot* terendah yang berarti bahwa nilai *monetary* segmen 1 paling rendah dari kedua segmen lainnya.



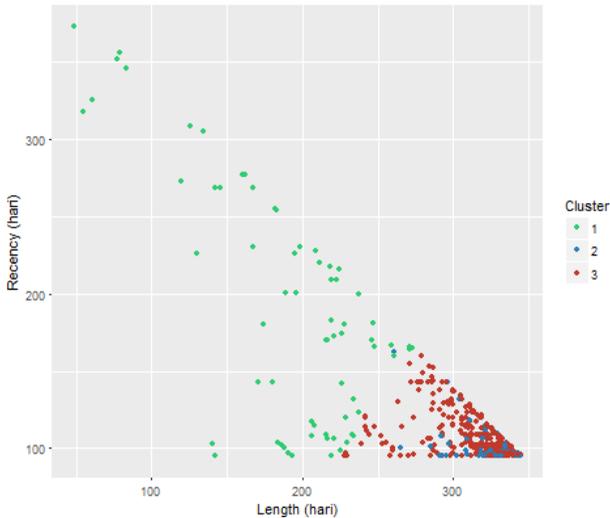
Gambar 6.18 Grafik Box Plot Variabel Monetary

6.5.4 Scatter 2D Plot

Grafik *scatter 2D plot* digunakan untuk melihat perbandingan persebaran antar variabel LRFM tiap segmennya. Dari perbandingan tersebut akan didapatkan perbedaan karakteristik tiap segmen.

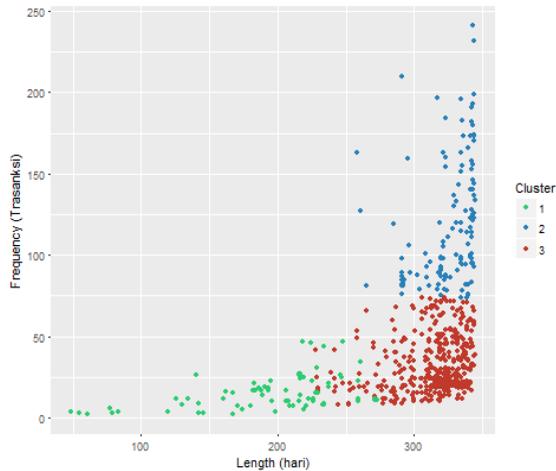
Pada Gambar 6.19 terlihat bahwa persebaran variabel *length* dan *recency* tiap segmen. Segmen 1 memiliki persebaran variabel *recency* yang tinggi dan persebaran variabel *length* yang tidak terlalu tinggi yang mengartikan bahwa retailer pada segmen 1 tergolong retailer baru namun dalam melakukan transaksi terakhir kali terhitung telah lama dilakukan

berdasarkan tanggal penelitian. Sedangkan segmen 2 dan segmen 3 memiliki persebaran variabel *length* yang tinggi namun pada segmen 3 diikuti dengan persebaran variabel *recency* yang lebih tinggi dibandingkan dengan variabel *recency* pada segmen 2 yang mengartikan bahwa rata-rata transaksi terakhir kali yang dilakukan oleh retailer pada segmen 2 lebih baik atau lebih sering melakukan transaksi dibandingkan dengan segmen 3.



Gambar 6.19 Scatter 2D Plot Variabel Length, Recency

Selanjutnya pada Gambar 6.20 menjelaskan tentang perbandingan variabel *length* dan variabel *frequency*.

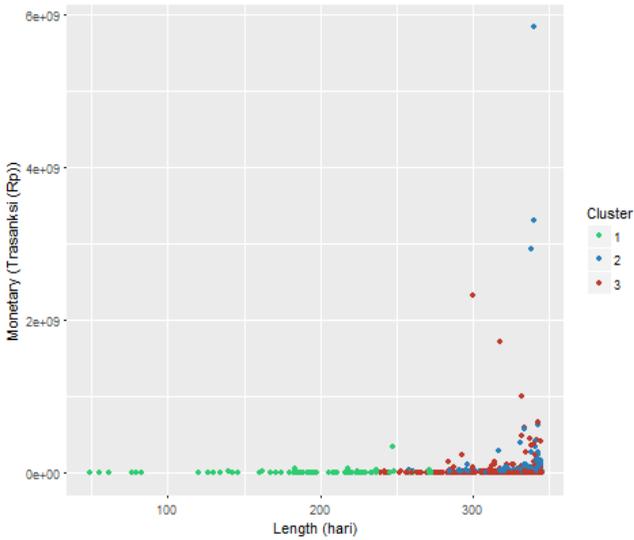


Gambar 6.20 Scatter 2D Plot Variabel Length, Frequency

Dari gambar diatas terlihat bahwa segmen 2 memiliki nilai variabel *length* dan *frequency* diatas rata-rata yang mengartikan bahwa retailer pada segmen ini sangat *intens* dan loyal melakukan transaksi sedangkan pada segmen 1 memiliki nilai variabel *length* dan *frequency* dibawah rata-rata yang berarti bahwa segmen ini jarang melakukan transaksi dan segmen 3 memiliki nilai variabel *length* diatas rata-rata namun tidak dengan nilai *frequency*nya yang ada dibawah nilai *frequency* segmen 2 yang berarti retailer pada segmen 3 tidak terlalu banyak melakukan transaksi.

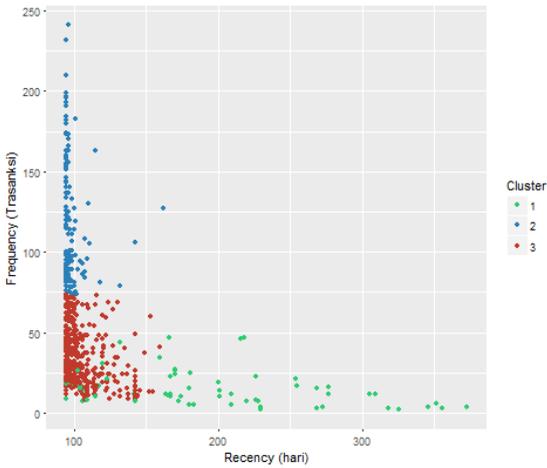
Pada Gambar 6.21 menggambarkan perbandingan nilai variabel *length* terhadap nilai variabel *monetary*. Terlihat pada segmen 2 dan 3 memiliki nilai variabel *monetary* yang tinggi namun nilai variabel *length* pada segmen 2 lebih tinggi daripada nilai variabel *length* pada segmen 3 yang mengartikan bahwa pengeluaran atau nilai transaksi yang dikeluarkan oleh segmen 2 sangat tinggi dan cukup loyal dalam melakukan transaksi. Sedangkan pada segmen 1 memiliki nilai variabel *monetary* dan *length* yang rendah yang menjadikan segmen ini menjadu segmen terburuk dikarenakan nilai transaksi yang rendah dan

lama kurun waktu bergabung dengan perusahaan tidak lama dibandingkan dengan kedua segmen lainnya.



Gambar 6.21 Scatter 2D Plot Variabel Length, Monetary

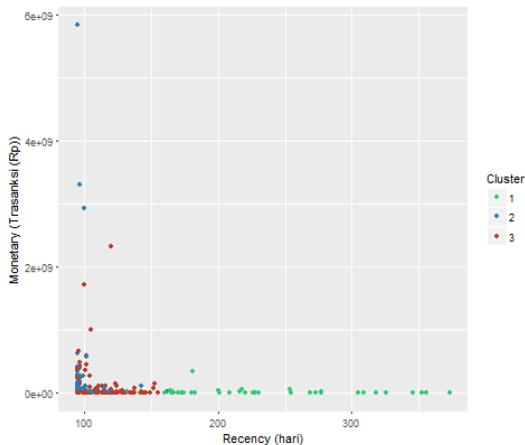
Perbandingan nilai variabel selanjutnya terlihat pada Gambar 6.22 yakni antara nilai *Recency* terhadap nilai *Frequency*.



Gambar 6.22 Scatter 2D Plot Variabel Recency, Frequency

Diketahui segmen 2 dan segmen 3 memiliki nilai *recency* yang kecil, namun nilai *frequency* yang tinggi dimiliki oleh segmen 2 yang ada diatas rata-rata yang berarti segmen 2 dan segmen 3 sering melakukan transaksi hanya saja banyaknya transaksi yang dilakukan membedakan segmen 2 dan 3. Segmen 1 memiliki nilai *recency* yang tinggi namun diikuti dengan nilai *frequency* yang rendah yang menjadikan segmen 1 ini segmen terburuk dikarenakan banyaknya transaksi yang dilakukan sangat rendah dan transaksi terakhir yang dilakukan telah lama dilakukan.

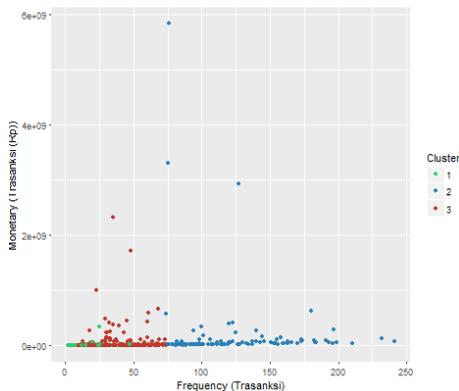
Kemudian pada Gambar 6.23 terlihat perbandingan nilai variabel *recency* dan *monetary*. Terlihat bahwa segmen 1 memiliki nilai *recency* dan *monetary* yang rendah yang menjadikan segmen ini menjadi segmen terburuk dimana nilai transaksi yang rendah diikuti dengan transaksi terakhir kali yang dilakukan sudah cukup lama dihitung dari tanggal penelitian.



Gambar 6.23 Scatter 2D Plot Variabel Recency, Monetary

Pada perbandingan variabel *recency* dan *monetary*, segmen 2 memiliki nilai *recency* yang kecil atau waktu terakhir kali melakukan transaksi tidak terpaut jauh dari waktu penelitian dan nilai rata-rata *monetary* atau nilai transaksi yang tinggi.

Pada Gambar 6.24 menggambarkan perbandingan nilai variabel *frequency* dan *monetary* dimana segmen 2 memiliki nilai tertinggi pada kedua variabel tersebut yang. Kemudian diikuti dengan segmen 3 dengan nilai *frequency* yang berada pada rata-rata dan segmen 1 dengan nilai terkecil dari kedua variabel tersebut yang menjadikan retailer pada segmen ini menjadi retailer yang paling buruk dalam hal transaksi.



Gambar 6.24 Scatter 2D Plot Variabel Frequency, Monetary

6.6 Kesimpulan Hasil Analisa

Setelah melakukan proses *clustering* dan analisa visualisasi yang telah dilakukan, diperoleh kesimpulan sebagai berikut:

1. Hasil dari proses AHP menunjukkan bahwa variabel *monetary* memiliki bobot tertinggi yaitu sebesar 0.51. Hal ini menunjukkan bahwa variabel *monetary* merupakan pertimbangan yang paling penting dalam pemeringkatan retailer. Sedangkan variabel *length*, *recency*, *frequency* memiliki bobot 0.11, 0.06, 0.31.
2. Berdasarkan metode Elbow, nilai K yang paling optimal untuk digunakan adalah 3, sehingga jumlah segmen yang dihasilkan pada proses *clustering* berjumlah 3 segmen.

3. Dari total retailer sejumlah 624 retailer, setelah dilakukan proses *clustering* terdapat persebaran retailer tiap segmennya yakni 93 retailer pada segmen 1, 133 retailer pada segmen 2, dan segmen 3 dengan jumlah reatailer terbesar berjumlah 398 retailer
4. Dari keseluruhan area, retailer terbanyak berasal dari area Gresik dengan jumlah 262 retailer, diikuti oleh area Lamongan dengan jumlah retailer sebanyak 183 retailer, berikutnya area Bangkalan dengan 55 retailer, Sumenep 44 retailer dan Pamekasan serta Sampang yang masing-masing berjumlah 40 retailer.
5. Segmen 1 mempunyai rataan nilai CLV terkecil yakni 0.116. Hal ini didukung oleh dengan nilai *length*, *frequency*, *monetary* yang kecil dan nilai *recency* yang besar yang mengartikan bahwa retailer yang tergabung pada segmen ini rata-rata sudah lama tidak melakukan transaksi dan atau merupakan retailer yang baru bergabung dan apabila melakukan transaksi tidak diikuti dengan jumlah transaksi yang besar maupun aktivitas transaksi yang cukup sering dilakukan. Sehingga dapat disimpulkan bahwa segmen 1 merupakan segmen terburuk. Dengan karakteristik tersebut, perusahaan dapat melakukan strategi *pick but be choosy*, yakni mengidentifikasi retailer yang lebih memberikan kontribusi, kemudian meningkatkan komunikasi dan perhatian kepada retailer tersebut, *do nothing* untuk retailer sisanya.
6. Segmen 2 merupakan segmen terbaik, terbukti dengan rata-rata nilai CLV 0.323, nilai *length*, nilai *frequency* dan nilai *monetary* yang tinggi serta nilai *recency* yang rendah mengidentifikasi rata-rata retailer pada segmen 2 ini sangat sering melakukan transaksi diikuti dengan jumlah transaksi yang cukup besar. Perusahaan dapat mempertahankan loyalitas para retailer dengan menerapkan strategi *up selling* dengan cara memberikan *reward* seperti potongan harga setelah pelanggan bertransaksi x rupiah.

7. Segmen 3 merupakan segmen yang memiliki rata-rata CLV yang sedang yakni 0.199 dan diikuti dengan nilai *length* yang cukup tinggi, nilai *recency* yang rendah serta nilai *frequency* dan *monetary* yang sedang. Dengan karakteristik tersebut, perusahaan dapat menerapkan strategi dengan cara mengembangkan aktivitas promosi yang bertujuan untuk meningkatkan frekuensi transaksi retailer, salah satu caranya adalah dengan memberikan reward seperti potongan harga setelah para retailer telah bertransaksi x kali.

6.7 Rekomendasi Strategi Marketing

Berdasarkan hasil analisa sebelumnya, terdapat beberapa temuan yang didapatkan terhadap perilaku-perilaku para retailer tiap segmennya, diantaranya:

- a. Segmen 1: memiliki karakteristik perilaku retailer yang baru bergabung dengan perusahaan dilihat dari nilai variabel *length* yang rendah, namun retailer pada segmen ini diketahui memiliki nilai variabel *recency* yang tinggi yang berarti waktu terakhir kali melakukan transaksi cenderung telah lama dilakukan terhitung dari waktu penelitian atau cenderung jarang melakukan transaksi. Kemudian nilai transaksi dan banyaknya melakukan transaksi tergolong rendah, sehingga retailer pada segmen ini dapat digolongkan pada *uncertain new customer*.
- b. Segmen 2: karakteristik perilaku retailer ini memiliki nilai transaksi dan jumlah transaksi diatas rata-rata. Selanjutnya para retailer pada segmen ini cukup loyal terhadap perusahaan dengan terbukti telah melakukan transaksi pertama yang telah cukup lama serta terakhir kali melakukan transaksi tidak terlalu jauh dari periode penelitian, sehingga retailer ini dapat digolongkan sebagai retailer yang loyal dan menjadi retailer utama oleh perusahaan.

- c. Segmen 3: retailer pada segmen ini memiliki karakteristik perilaku retailer yang loyal dengan telah beragabung dan melakukan awal transaksi yang telah cukup lama serta transaksi akhir yang dilakukan tidak terlalu jauh dilakukan dari tanggal penelitian, namun jumlah dan nilai transaksi yang dilakukan tidak terlalu tinggi, sehingga segmen ini tergolong retailer yang potensial untuk menjadi retailer utama perusahaan.

Dari karakteristik retailer tersebut maka terdapat usulan strategi yang dapat digunakan untuk mengelola ketiga segmen retailer yang dihasilkan. Usulan strategi berdasarkan beberapa referensi yang terkait dengan penelitian ini [27] dapat dilihat pada Tabel 6.21

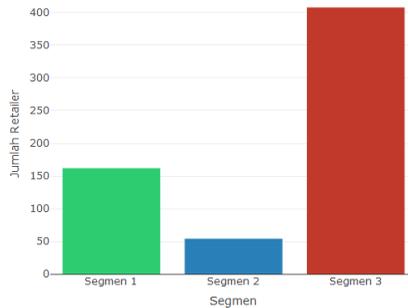
Tabel 6.21 Usulan Strategi CRM

| Mark Strategi | Segmen | Strategi | Detail Strategi |
|--------------------------------------|----------|---|--|
| <i>“Pick But Be Choosy Strategy”</i> | Segmen 1 | Fokus pada <i>“Pick But Be Choosy</i> | Aktivitas CRM yang disarankan: <ol style="list-style-type: none"> 1. Mengidentifikasi retailer yang lebih memberikan kontribusi 2. Tingkatkan komunikasi dan <i>follow up</i> 3. <i>Do nothing</i> pada retailer sisanya |
| <i>Offensive Strategy</i> | Segmen 2 | <ol style="list-style-type: none"> 1. Mempertahankan loyalitas dengan <i>up selling</i> 2. Mengembangkan promosi untuk meningkatkan nilai transaksi | Aktivitas CRM yang disarankan: <ol style="list-style-type: none"> 1. Memberikan reward potongan harga setelah bertransaksi x rupiah |
| <i>Defensive Strategy</i> | Segmen 3 | Mengembangkan promosi untuk meningkatkan frekuensi transaksi | Aktivitas CRM yang disarankan: Memberikan reward potongan harga setelah bertransaksi x rupiah |

6.8 Perbandingan Nilai K

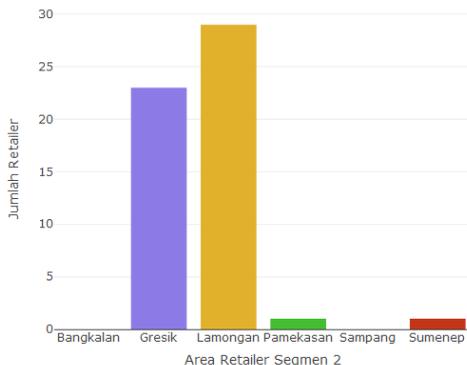
Dari nilai K yang didapatkan pada proses *clustering* sebelumnya, terdapat beberapa nilai K yang dapat dijadikan perbandingan penentuan jumlah segmen, berikut ini adalah perbandingan nilai K lainnya:

6.8.1 Pembobotan dengan Nilai K=3



Gambar 6.25 Hasil Segmentasi K = 3 Bobot

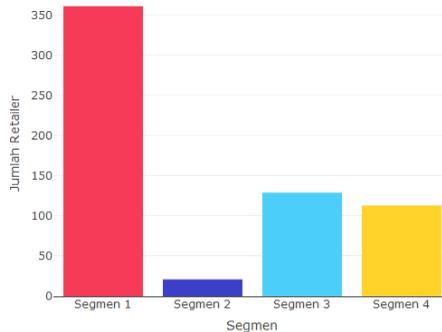
Pada Gambar 6.25 terlihat hasil segmen memiliki penyebaran retailer yang tidak merata, hal ini diperjelas pada Gambar 6.26 bahwa pada segmen nilai K=3 bobot ini terdapat area yang tidak tercangkup pada segmen 2 yakni area Bangkalan dan Sampang yang keduanya tidak memiliki anggota retailer yang tergabung pada segmen 2.



Gambar 6.26 Segmen 2, K=3 Bobot

6.8.2 Tanpa Pembobotan dengan Nilai $K=4$

Selanjutnya perbandingan nilai K pada penentuan jumlah segmen adalah nilai $K=4$ tanpa pembobotan, dimana terlihat pada Gambar 6.27, penyebaran retailer sangat tidak merata dimana pada segmen 2 hanya terdiri dari 21 retailer.

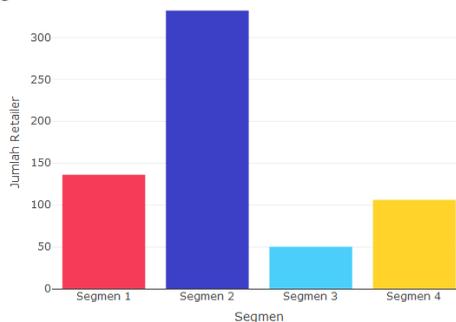


Gambar 6.27 Hasil Segmentasi $K=4$ Tanpa Bobot

Sedangkan jumlah retailer pada segmen 1 berjumlah 361, segmen 3 berjumlah 129 dan segmen 4 berjumlah 133 retailer.

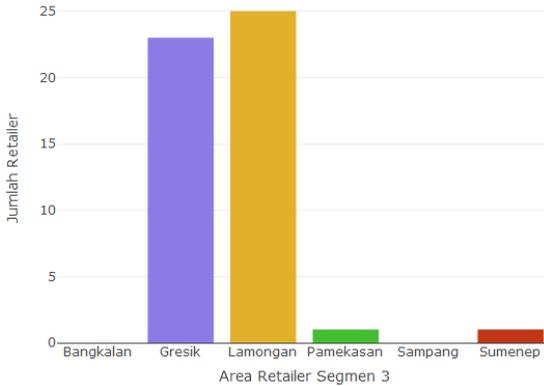
6.8.3 Pembobotan dengan Nilai $K=4$

Selanjutnya nilai $k = 4$ dengan bobot menjadi perbandingan hasil segmentasi. Pada Gambar 6.28 terlihat penyebaran retailer terlihat merata, namun pada segmen 3 terdapat persebaran retailer yang tidak merata.



Gambar 6.28 Hasil Segmentasi $K=4$ Bobot

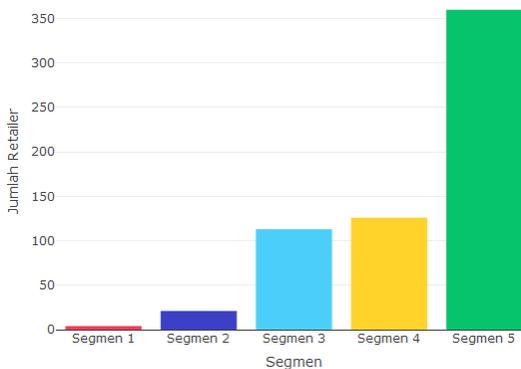
Diketahui pada segmen 3 dengan area retailer Bangkalan dan Sampang tidak terdapat retailer yang bergabung dan persebaran retailer pada area lainnya pada segmen ini cukup tidak merata, untuk lebih jelasnya dapat dilihat pada Gambar 6.29.



Gambar 6.29 Segmen 3, $K=4$ Bobot

6.8.4 Tanpa Pembobotan dengan Nilai $K=5$

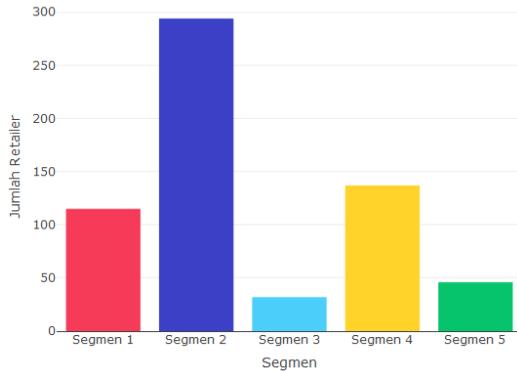
Pada Gambar 6.30 terlihat persebaran retailer sangat tidak merata, dimana retailer dominan berkumpul pada segmen 5 dan segmen 1 merupakan segmen dengan keanggotaan retailer yang sangat sedikit yang hanya beranggotakan retailer yang berasal dari area Lamongan dan Pamekasan.



Gambar 6.30 Hasil Segmentasi $K=5$ Tanpa Bobot

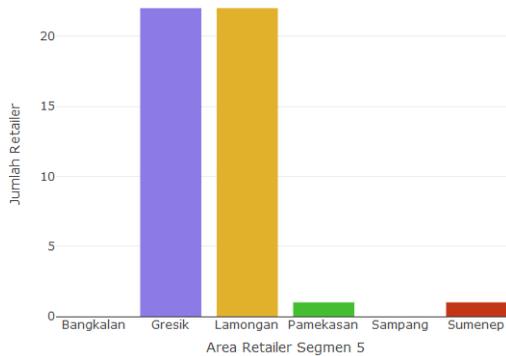
6.8.5 Pembobotan dengan Nilai $K=5$

Pada Gambar 6.31 terlihat penyebaran retailer pada tiap segmen dengan menggunakan nilai $K=5$ bobot cukup merata.



Gambar 6.31 Hasil Segmentasi $K=5$ Bobot

Namun pada segmen 5, terdapat persebaran retailer tiap area yang tidak merata. Terdapat area yang tidak memiliki retailer pada segmen ini yakni area Bangkalan dan Sampang.



Gambar 6.32 Segmen 5, $K=5$ Bobot

Halaman ini sengaja dikosongkan

BAB VII

KESIMPULAN DAN SARAN

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

7.1 Kesimpulan

Dari hasil dan proses penelitian tugas akhir ini didapatkan beberapa kesimpulan, antara lain:

1. Proses *clustering* yang digunakan dalam penggalian data dapat membantu perusahaan dalam menampilkan perilaku bisnis retailer yang sama yang dikelompokkan dalam segmen-segmen tertentu. Metode *K-Means* yang diimplementasikan menghasilkan 3 segmen retailer dengan karakteristik yang berbeda berdasarkan pertimbangan beberapa variabel diantaranya, lama waktu telah bertransaksi dengan perusahaan, rentang waktu terakhir kali bertransaksi, jumlah transaksi dan total rupiah yang diterima oleh perusahaan dalam transaksi. Dengan pembagian segmen ini, perusahaan dapat menentukan strategi pemasaran yang tepat untuk setiap segmen dilihat dari karakteristiknya.
2. Dari hasil *clustering* yang dilakukan diketahui bahwa retailer yang dominan merupakan retailer yang termasuk kategori medium dan retailer yang termasuk kategori loyal dan *profitable* hanya sekitar 21.3%, sehingga perusahaan dapat melakukan program-program untuk dapat meningkatkan jumlah retailer yang dapat menguntungkan perusahaan.
3. Visualisasi hasil *clustering* berbasis web dapat mempermudah perusahaan dalam membaca grafik

hasil *clustering* yang ditampilkan. Perusahaan dapat membaca jumlah dan persentase persebaran retailer tiap segmen maupun area dalam grafik *bar chart* dan *pie chart*. Perusahaan juga dapat melihat persebaran retailer tiap segmennya dalam grafik *scatter 3D plot* dan menganalisis rentang variabel LRFM pada grafik *box plot*. Selain itu, perusahaan juga dapat melihat urutan peringkat maupun retailer terbaik dan terburuk berdasarkan tabel peringkat yang diurutkan berdasarkan nilai *customer life value*.

7.2 Saran

Dalam pengerjaan tugas akhir ini, tentunya terdapat beberapa kekurangan yang dapat diperbaiki dan ditingkatkan. Oleh karena itu, berikut ini terdapat beberapa saran yang dapat dipertimbangkan dalam pengerjaan tugas akhir dengan topik yang sama kedepannya.

1. Penelitian ini dapat dilanjutkan dengan menggunakan metode *clustering* yang lain selain metode *K-Means*, misalkan menggunakan metode *Fuzzy C-Means*.
2. Penentuan nilai optimal terhadap K dapat dikembangkan dengan metode lain, misalkan SOM, agar dapat memberikan nilai K yang lebih optimal.
3. Rentang data transaksi dalam penelitian ini lebih baik diperbanyak lagi lebih dari satu tahun transaksi agar hasil dari proses *clustering* lebih baik lagi, misal 2 tahun transaksi
4. Penelitian ini dapat dilanjutkan dengan melakukan proses *clustering* untuk area yang lebih luas lagi, misalkan keseluruhan area distribusi PT.XYZ Gresik
5. Penentuan variabel *clustering* dapat menggunakan metode selain metode LRFM, misal metode RFMD.
6. Penelitian ini dapat dikembangkan lagi dengan menggunakan software selain R sebagai *tools clustering*, misal menggunakan Phyton sebagai pembandingan hasil *clustering*

DAFTAR PUSTAKA

- [1] A. KataData, "Katadata News and Research," 8 Agustus 2016. [Online]. Available: <https://databoks.katadata.co.id/datapublish/2016/08/08/pengguna-smartphone-di-indonesia-2016-2019>. [Accessed 6 Januari 2018].
- [2] P. R. Indonesia, "PERATURAN PEMERINTAH TENTANG PENYELENGGARAAN TELEKOMUNIKASI," 2000.
- [3] A. Katadata, "Katadata News and Research," 3 Mei 2017. [Online]. Available: <https://databoks.katadata.co.id/datapublish/2017/05/03/siapa-operator-seluler-yang-mempunyai-pelanggan-terbanyak>. [Accessed 7 Januari 2018].
- [4] Y. Nugraheni, *Data Mining Dengan Metode Fuzzy Untuk Customer Relationship Management (CRM) pada Perusahaan Retail*, 2011.
- [5] Procurement Leaders/ Sigaria Ltd, *STRATEGY GUIDE: SUPPLIER RELATIONSHIP MANAGEMENT*, London: Procurement Leaders/ Sigaria Ltd, 2013.
- [6] P. W. Schroder, "A better way to engage with supplier," in *CSCMP's Supply Chain*, Agile Business Media, 2012.
- [7] H. Chang and S. Tsay, "Integrating of SOM and K-mean in Data Mining Clustering: An Empirical Study of CRM and Provitability Evaluation," *Journal of Information Management*, vol. 11, no. 4, pp. 161-203, 2004.
- [8] J. C. Black, "Effective Supplier Relationship Management," in *Northeast Supply Chain Conference*, Southborough.
- [9] E. Prasetyo, *Data Mining Konsep dan Aplikasi Menggunakan Matlab*, Yogyakarta: Andi, 2013.

- [10] A. M. Hughes, H. A. M, Strategic database marketing, Chicago: Probus Publishing, 1994.
- [11] B. Stone and R. Jacobs, "Data Mining," in *Successful direct marketing methods*, Lincolnwood, NTC Business Book, 1995, pp. 35-57.
- [12] H.-H. Wu, S.-Y. Lin and C.-W. Liu, "Analyzing Patients' Values by Applying Cluster Analysis and," *The Scientific World Journal*, 2014.
- [13] P. Prasad, "Generating Customer Profiles for Retail Stores Using Clustering Techniques," *International Journal on Computer Science and Engineering*, vol. 3, pp. 2506-2510, 2011.
- [14] B. H. H. Maskan, "Proposing a Model for Customer Segmentation using WRFM Analysis (Case Study: an ISP Company)," *International Journal of Economy, Management and Social Sciences*, vol. 3, no. 12, pp. 77-80, 2014.
- [15] M. Alvandi, S. Fazli and F. S. Abdoli, "K-Mean Clustering Method For Analysis Customer Lifetime Value With LRFM Relationship Model in Banking Service," *International Research Journal of Applied and Basic Sciences*, vol. 3, no. 11, pp. 2294-2302, 2012.
- [16] E. Ngai, "Customer relationship management research (1992-2002)," *Marketing Intelligence & Planning*, vol. 23, no. 6, pp. 582-605, 2005.
- [17] P. Greenberg, CRM at the Speed of Light, Fourth Edition, New York: MCGRAW-HILL Education, 2010.
- [18] H. Peck, A. Payne, M. Christopher and M. Clark, Relationship Marketing – Strategy and Implementation, Burlington: Butterworth Heinemann, 2004.
- [19] A. Julie, "Data data everywhere – and not a byte of use?," *Market Research: An International Journal*, vol. 4, no. 3, pp. 182-192, 2007.
- [20] M. Viljoen, J. Bennet, A. Berndt and C. Van Zyl, "The Use Of Technology In Customer Relationship

- Management (CRM)," *Acta Commercii*, vol. 5, no. 1, pp. 106-116, 2005.
- [21] B. W. Taylor, *Introduction to Management Science*, Virginia: Pearson Education, Inc, 2013.
- [22] D. Virmani, S. Taneja and G. Malhotra, "Normalization based K means Clustering Algorithm," *International Journal of Advanced Engineering Research and Science*, 2015.
- [23] What is Data Mining, 1999.
- [24] P. Tan, M. Steinbach and V. Kumar, "Cluster Analysis: Basic Concepts and Algorithms," *Introduction to Data Mining*, pp. 487-568, 2005.
- [25] P. Bholowalia and A. Kumar, "EBK-Means: A Clustering Technique based on Elbow Method and K-Means in WSN," *International Journal of Computer Applications* , vol. 9, no. 105, pp. 17-24, 2014.
- [26] P. Tan, M. Steinbach and V. Kumar, *Introduction to Data Mining*, Boston: Addison-Wesley Longman Publishing Co, 2005.
- [27] S. Roberts, "Transform Your Data," *Nutrition*, vol. 21, no. 5, pp. 492-494, 2008.
- [28] D. C. Howell, *Statistical Methods for Psychology* (6th edition), Belmont: Thomson Wadsworth, 2007.

Halaman ini sengaja dikosongkan

BIODATA PENULIS



Penulis memiliki nama lengkap Syahrul Septa Perdana, lahir di Sampang, 29 September 1995. Penulis merupakan anak pertama dari dua bersaudara. Penulis mengenyam pendidikan formal Sekolah Dasar di SD Negeri Gunung Sekar 1 dan melanjutkan ke jenjang Sekolah Menengah Pertama di SMP Negeri 1 Sampang. Kemudian penulis melanjutkan

jenjang pendidikan di SMA 1 Sampang dan mengenyam bangku perkuliahan di Departemen Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi, Institut Teknologi Sepuluh Nopember, Surabaya. Selama menjadi mahasiswa, penulis aktif di berbagai kegiatan dan organisasi diluar kampus, diantaranya penulis terdaftar sebagai wakil ketua forum mahasiswa daerah Sampang, Institut Teknologi Sepuluh Nopember, staf ahli departemen sosial dan teknologi paguyuban penerima beasiswa Karya Salemba Empat, ITS 2017. Dalam pengerjaan Tugas Akhir, penulis terdaftar sebagai anggota laboratorium Rekayasa Data dan Intelegensi Bisnis, dengan bimbingan Bapak Edwin Riksakomara, S.Kom., M.T dengan topik mengenai penggalian data. Penulis dapat dihubungi melalui email syahrulsepta@gmail.com untuk keperluan penelitian. Semoga penulisan Tugas Akhir ini dapat memberikan kontribusi positif bagi semua pihak yang terkait

Halaman ini sengaja dikosongkan

LAMPIRAN A

KUESIONER PENELITIAN SURVEY PEMERINGKATAN KEPENTINGAN KRITERIA PERILAKU RETAILER OPERATOR TELEKOMUNIKASI SELULER

Kuesioner ini bertujuan untuk mendapatkan masukan pendapat/pemikiran dari Narasumber terkait dengan pemeringkatan kriteria yang akan digunakan dalam pemberian bobot pada variabel segmentasi retailer. Kami menjamin kerahasiaan data/informasi yang diberikan serta tidak akan membawa implikasi apapun bagi responden.

IDENTITAS RESPONDEN

Nama Responden : _____
Jabatan : _____
Tanda Tangan : _____

PETUNJUK PENGISIAN KUESIONER

1. Berilah tanda centang (√) pada kriteria yang menurut penilaian bapak/ibu lebih penting dibandingkan dengan kriteria lainnya, dengan cara membandingkan kriteria pada sebelah kiri dengan kriteria pada sebelah kanan. Kriteria yang dimaksud adalah perilaku *retailer* 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 disbanding kriteria pasangannya |
| 5 | Lebih Penting | Penilaian dan pengalaman memihak pada salah satu kriteria tertentu disbanding 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 *retailer* yang dibandingkan yaitu:
 - a. *Length* : Interval atau jarak waktu trasanksi awal hingga transaksi terakhir yang dilakukan oleh *retailer*

- b. *Recency* : Jarak dari waktu transaksi terakhir kali yang dilakukan oleh *retailer* dengan waktu saat ini
- c. *Frequency* : Total jumlah transaksi yang dilakukan oleh *retailer* selama periode tertentu
- d. *Monetary* : Total nilai transaksi suatu produk dalam bentuk uang yang diberikan *retailer* pada perusahaan

CONTOH PENGISIAN KUESIONER

| Kriteria | Skala Prioritas | | | | | | | | | | | | | | | | | | Kriteria |
|------------------|-----------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----------------|----------|
| <i>Frequency</i> | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | <i>Monetary</i> | |
| | | | | | √ | | | | | | | | | | | | | | |

Artinya:

Pada pernyataan contoh pengisian kuesioner diatas, tanda centang (√) di kolom jawaban sebelah kiri menunjukkan bahwa: ***frequency* 5 kali lebih penting dibandingkan dengan *monetary*** dalam hal pemeringkatan kriteria prioritas untuk *retailer*. Dengan kata lain bahwa *monetary* 1/5 kali lebih penting dibandingkan dengan *frequency*.

DAFTAR PERTANYAAN

Berikan tanda centang (√) untuk menunjukkan penilaian bapak/ibu terhadap perbandingan kriteria *retailer* 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 | Skala Prioritas | | | | | | | | | | | | | | | | | | Kriteria |
|---------------|-----------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|------------------|----------|
| <i>Length</i> | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | <i>Recency</i> | |
| <i>Length</i> | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | <i>Frequency</i> | |
| | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | | |

A-4

| | | | | | | | | | | | | | | | | | | | |
|---------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|------------------|----------------|
| <i>Length</i> | | | | | | | | | | | | | | | | | | | <i>Monthly</i> |
| <i>Rece</i> | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | <i>Frequency</i> | |
| <i>ncy</i> | | | | | | | | | | | | | | | | | | | |
| <i>Rece</i> | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | <i>Monthly</i> | |
| <i>ncy</i> | | | | | | | | | | | | | | | | | | | |
| <i>Freq</i> | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | <i>Monthly</i> | |
| <i>uency</i> | | | | | | | | | | | | | | | | | | | |

Terima kasih atas kesediaan bapak/ibu untuk mengisi kuesioner ini.

LAMPIRAN B

Tabel B.1 Use Case Melihat Grafik Box Plot

| UC-2 Melihat persebaran variabel pada grafik <i>box plot</i> | | | | | | | | | |
|---|---|--------------|---------------|---|---|---------------------------------------|---|--------------------------------------|---|
| Purpose | Mengetahui persebaran nilai variabel <i>length</i> , <i>recency</i> , <i>frequency</i> , dan <i>monetary</i> dalam grafik <i>box plot</i> | | | | | | | | |
| Overview | Dimulai ketika pengguna mengakses aplikasi atau memasuki laman Halaman Visualisasi segmen | | | | | | | | |
| Actors | Pengguna Aplikasi | | | | | | | | |
| Pre Condition | Pengguna sudah ada halaman utama | | | | | | | | |
| Post Condition | Grafik <i>box plot</i> ditampilkan | | | | | | | | |
| Typical Couse Event | <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 50%; text-align: center;"><i>Actor</i></th> <th style="width: 50%; text-align: center;"><i>System</i></th> </tr> </thead> <tbody> <tr> <td>1. Mengklik ikon  <i>slide bar</i></td> <td>2. Menampilkan pilihan menu pada <i>slide bar</i></td> </tr> <tr> <td>3. Mengklik menu “Visualisasi segmen”</td> <td>4. Menampilkan halaman visualisasi segmen</td> </tr> <tr> <td>5. Mengklik pilihan variabel L,R,F,M</td> <td>6. Menampilkan grafik <i>box plot</i> sesuai pilihan variabel L,R,F,M</td> </tr> </tbody> </table> | <i>Actor</i> | <i>System</i> | 1. Mengklik ikon  <i>slide bar</i> | 2. Menampilkan pilihan menu pada <i>slide bar</i> | 3. Mengklik menu “Visualisasi segmen” | 4. Menampilkan halaman visualisasi segmen | 5. Mengklik pilihan variabel L,R,F,M | 6. Menampilkan grafik <i>box plot</i> sesuai pilihan variabel L,R,F,M |
| <i>Actor</i> | <i>System</i> | | | | | | | | |
| 1. Mengklik ikon  <i>slide bar</i> | 2. Menampilkan pilihan menu pada <i>slide bar</i> | | | | | | | | |
| 3. Mengklik menu “Visualisasi segmen” | 4. Menampilkan halaman visualisasi segmen | | | | | | | | |
| 5. Mengklik pilihan variabel L,R,F,M | 6. Menampilkan grafik <i>box plot</i> sesuai pilihan variabel L,R,F,M | | | | | | | | |
| Alternate Flow Event | - | | | | | | | | |
| Exceptional Flow Event | Aplikasi tertutup akibat <i>error</i> | | | | | | | | |

Tabel B.02 Use Case Melihat Grafik Scatter 3D Plot

| UC-3 Melihat Grafik Scatter 3D Plot | | |
|--|---|--|
| Purpose | Mengetahui persebaran nilai variabel <i>length</i> , <i>recency</i> , <i>frequency</i> , dan <i>monetary</i> dalam grafik <i>scatter 3D plot</i> | |
| Overview | Dimulai ketika pengguna mengakses aplikasi atau memasuki laman Halaman Visualisasi segmen | |
| Actors | Pengguna Aplikasi | |
| Pre Condition | Pengguna sudah ada halaman utama | |
| Post Condition | Grafik <i>scatter 3D plot</i> ditampilkan | |
| Typical Cause Event | Actor | System |
| | 1. Mengklik ikon  <i>slide bar</i> 3. Mengklik menu “Visualisasi segmen” 5. Mengklik salah satu pilihan segmen 7. Mengklik pilihan <i>scatter lfm</i> atau <i>scatter rfm</i> pada kotak dengan judul “Grafik 3 Dimensi” | 2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman visualisasi segmen 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan grafik <i>scatter 3D Plot sesuai pilihan</i> |
| Alternate Flow Event | Pada langkah 5, jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan grafik semua segmen secara default | |
| Exceptional Flow Event | Aplikasi tertutup akibat <i>error</i> | |

Tabel B.3 Use Case Melihat Retailer Teratas & Terbawah

| UC-4 Melihat Retailer Teratas dan Terbawah | | |
|---|---|---|
| Purpose | Mengetahui retailer dengan nilai CLV teratas dan terbawah setiap segmen | |
| Overview | Dimulai ketika pengguna mengakses aplikasi atau memasuki laman Halaman Visualisasi segmen | |
| Actors | Pengguna Aplikasi | |
| Pre Condition | Pengguna sudah ada halaman utama | |
| Post Condition | Informasi retailer teratas dan terbawah ditampilkan | |
| Typical Cause Event | Actor | System |
| | 1. Mengklik ikon  <i>slide bar</i> 3. Mengklik menu “Visualisasi segmen” 5. Mengklik salah satu pilihan segmen 7. Mengklik ikon  pada kotak dengan judul “Retailer Teratas dan Terbawah” | 2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman visualisasi segmen 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan informasi retailer teratas dan terbawah |
| Alternate Flow Event | Pada langkah 5, jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan grafik semua segmen secara default | |
| Exceptional Flow Event | Aplikasi tertutup akibat <i>error</i> | |

Tabel B0.4 Use Case Melihat Rentang Nilai Segmen

| UC-5 Melihat Rentang Nilai Segmen | |
|--|--|
| Purpose | Mengetahui rentang nilai variabel LRFM setiap segmen |
| Overview | Dimulai ketika pengguna mengakses aplikasi atau memasuki laman Halaman Visualisasi segmen |
| Actors | Pengguna Aplikasi |
| Pre Condition | Pengguna sudah ada halaman utama |
| Post Condition | Informasi rentang nilai LRFM |
| Typical Couse Event | Actor |
| | System |
| | <ol style="list-style-type: none"> 1. Mengklik ikon  <i>slide bar</i> 3. Mengklik menu “Visualisasi segmen” 5. Mengklik salah satu pilihan segmen 7. Mengklik ikon  pada kotak dengan judul “Rentang Segmen” |
| | <ol style="list-style-type: none"> 2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman visualisasi segmen 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan informasi rentang nilai LRFM tiap segmen |
| Alternate Flow Event | Pada langkah 5, jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan grafik semua segmen secara default |
| Exceptional Flow Event | Aplikasi tertutup akibat <i>error</i> |

Tabel B.5 Use Case Melihat Rataan Variabel LRFM

| UC-6 Melihat Rataan Variabel LRFM Tiap Segmen | | |
|--|--|--|
| Purpose | Mengetahui nilai rata-rata variabel LRFM tiap segmen | |
| Overview | Dimulai ketika pengguna mengakses aplikasi atau memasuki laman Halaman Visualisasi segmen | |
| Actors | Pengguna Aplikasi | |
| Pre Condition | Pengguna sudah ada halaman utama | |
| Post Condition | Informasi rata-rata nilai variabel LRFM setiap segmen | |
| Typical Couse Event | Actor | System |
| | 1. Mengklik ikon  <i>slide bar</i> 3. Mengklik menu “Visualisasi segmen” 5. Mengklik salah satu pilihan segmen | 2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman visualisasi segmen 6. Menampilkan rataan nilai variabel LRFM pilihan segmen |
| Alternate Flow Event | Pada langkah 5, jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan grafik semua segmen secara default | |
| Exceptional Flow Event | Aplikasi tertutup akibat <i>error</i> | |

Tabel B.6 Use Case Melihat Ranking Retailer Setiap Segmen

| UC-7 Melihat Ranking Retailer Setiap Segmen | | |
|--|---|--|
| Purpose | Mengetahui <i>ranking</i> retailer setiap segmen yang berdasarkan nilai CLV yang ditampilkan dalam bentuk tabel | |
| Overview | Dimulai ketika pengguna mengakses aplikasi atau memasuki laman Halaman Visualisasi segmen | |
| Actors | Pengguna Aplikasi | |
| Pre Condition | Pengguna sudah ada halaman utama | |
| Post Condition | Tabel peringkat retailer berdasarkan pilihan segmen | |
| Typical Cause Event | Actor | System |
| | 1. Mengklik ikon  <i>slide bar</i> 3. Mengklik menu “Visualisasi segmen” 5. Mengklik salah satu pilihan segmen 7. Mengklik ikon  pada kotak dengan judul “Peringkat Retailer” | 2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan halaman visualisasi segmen 6. Menampilkan visualisasi sesuai pilihan 8. Menampilkan tabel peringkat retailer sesuai pilihan segmen |
| Alternate Flow Event | Pada langkah 5, jika pengguna tidak mengklik pilihan segmen, maka secara otomatis sistem akan menampilkan grafik semua segmen secara default | |
| Exceptional Flow Event | Aplikasi tertutup akibat <i>error</i> | |

Tabel B.7 Use Case Melihat Tabel Data Segmen Retailer

| UC-8 Melihat Tabel Data Segmen Retailer | | |
|--|--|--|
| Purpose | Mengetahui secara keseluruhan data pelanggan hasil <i>clustering</i> | |
| Overview | Dimulai ketika pengguna mengakses aplikasi atau memasuki laman Tabel Data | |
| Actors | Pengguna Aplikasi | |
| Pre Condition | Pengguna sudah ada halaman utama | |
| Post Condition | Tabel data keseluruhan retailer hasil <i>clustering</i> | |
| Typical Cause Event | Actor | System |
| | 1. Mengklik ikon  <i>slide bar</i> 3. Mengklik menu "Tabel Data" | 2. Menampilkan pilihan menu pada <i>slide bar</i> 4. Menampilkan tabel data segmen retailer |
| Alternate Flow Event | - | |
| Exceptional Flow Event | Aplikasi tertutup akibat <i>error</i> | |

Halaman ini sengaja dikosongkan

LAMPIRAN C

| ID | NAME RET | L | R | F | M |
|-----------|------------------|----------|----------|----------|-----------|
| 103072 | new cell | 321 | 111 | 105 | 16710000 |
| 103073 | mila cell | 216 | 170 | 26 | 13575000 |
| 103473 | eka cell | 332 | 96 | 41 | 17012500 |
| 103478 | top cell | 329 | 110 | 130 | 50070000 |
| 103741 | army cell | 239 | 103 | 28 | 4960000 |
| 103742 | ronggolawe cell | 316 | 99 | 18 | 3250000 |
| 103743 | mitra cell | 304 | 132 | 79 | 19863000 |
| 103745 | jagad cell | 286 | 144 | 10 | 2600000 |
| 103747 | rifqi cell | 313 | 113 | 40 | 11596000 |
| 103841 | sb cell | 294 | 129 | 20 | 7850000 |
| 103842 | aa cell | 341 | 95 | 100 | 350050566 |
| 103843 | Elina Cell | 234 | 108 | 8 | 1305000 |
| 103844 | Rico cell | 334 | 95 | 120 | 59050000 |
| 103850 | fika cell | 259 | 167 | 10 | 1850000 |
| 103904 | lubic cell | 244 | 109 | 8 | 1642000 |
| 103910 | adinda jaya cell | 234 | 108 | 9 | 1463000 |
| 105971 | brow cell | 198 | 230 | 4 | 750000 |
| 105974 | putra cell | 322 | 100 | 79 | 16554000 |
| 106055 | Ucek cell | 308 | 109 | 37 | 6175000 |
| 106308 | SYITA CELL | 332 | 97 | 29 | 485800000 |
| 106374 | TOSARI CELL | 322 | 99 | 45 | 15268000 |
| 106377 | DNA CELL | 318 | 120 | 46 | 55775000 |
| 106496 | banyuanyar cell | 134 | 305 | 12 | 6400000 |
| 109441 | di2k cell | 274 | 120 | 19 | 3050000 |
| 109442 | alvin cell | 335 | 104 | 18 | 266000000 |
| 109444 | kana cell | 337 | 99 | 27 | 71700000 |

C-2

| ID | NAME RET | L | R | F | M |
|-----------|-----------------|----------|----------|----------|-----------|
| 109445 | Dirgahayu cell | 327 | 111 | 55 | 106500000 |
| 109446 | Isho cell | 296 | 130 | 34 | 11700000 |
| 109464 | hasanah cell | 319 | 121 | 40 | 23950000 |
| 109551 | pratama cell | 270 | 95 | 43 | 6773000 |
| 109552 | boger cell | 270 | 95 | 46 | 7074000 |
| 109553 | fifa cell | 265 | 100 | 66 | 10135000 |
| 109558 | ar rohman cell | 265 | 100 | 81 | 12424000 |
| 109564 | viktori cel | 243 | 111 | 19 | 5450000 |
| 109651 | Gmp2 | 322 | 100 | 74 | 11905000 |
| 109716 | hardi cell | 242 | 114 | 16 | 14200000 |
| 109813 | sinar cell | 335 | 95 | 22 | 3660000 |
| 110892 | ardhi cell | 293 | 95 | 85 | 26273486 |
| 110895 | permata cell | 229 | 120 | 31 | 5950000 |
| 111005 | multiphone 1 | 331 | 99 | 49 | 10163000 |
| 111006 | GAYU cell | 291 | 95 | 82 | 25479412 |
| 111028 | hany cell | 293 | 143 | 49 | 21849040 |
| 111063 | cha cha cell | 287 | 99 | 68 | 10500000 |
| 111142 | manyar music 2 | 319 | 97 | 89 | 15045000 |
| 111144 | alexa cell | 327 | 96 | 23 | 4124000 |
| 111146 | ogud cell | 335 | 103 | 43 | 11392000 |
| 111148 | bad cell | 328 | 100 | 21 | 3473000 |
| 111156 | pulsa max cell | 291 | 95 | 81 | 25493262 |
| 111201 | iis cell | 320 | 95 | 50 | 8347000 |
| 111271 | mulia cell | 312 | 95 | 73 | 21465088 |
| 111279 | rasya cell | 333 | 104 | 49 | 13600000 |
| 111412 | FARZ cell | 327 | 95 | 72 | 21733994 |
| 111417 | vanesia cell | 326 | 95 | 20 | 3252000 |
| 111511 | CBAX cell | 335 | 95 | 22 | 3610000 |
| 111512 | Metro cell | 331 | 100 | 19 | 3006000 |

| ID | NAME RET | L | R | F | M |
|-----------|------------------|----------|----------|----------|-----------|
| 111514 | minar bambu cell | 329 | 97 | 20 | 3350000 |
| 111515 | neng cell | 258 | 95 | 163 | 37317181 |
| 111521 | citra cell | 291 | 95 | 76 | 24030952 |
| 111530 | nier cell | 261 | 162 | 127 | 27488000 |
| 111572 | ulin cell | 292 | 95 | 82 | 25463340 |
| 111575 | Asli cell | 342 | 97 | 83 | 28270000 |
| 111578 | noni cell | 279 | 160 | 41 | 6545000 |
| 111639 | indah jaya | 329 | 95 | 20 | 3398000 |
| 117202 | aura cell | 180 | 143 | 7 | 1550000 |
| 117209 | ASKA CELL | 171 | 143 | 7 | 1100000 |
| 117292 | YAYAN cell | 319 | 96 | 42 | 11835000 |
| 117513 | Della cell | 216 | 109 | 10 | 1600000 |
| 117514 | madaniyah cell | 339 | 99 | 74 | 20330000 |
| 117515 | BJT cell | 319 | 96 | 43 | 7289306 |
| 117518 | bbc cell | 329 | 97 | 72 | 20500000 |
| 117519 | bos cell | 339 | 100 | 19 | 52731520 |
| 117520 | jaya sakti1 cell | 324 | 95 | 45 | 7980000 |
| 117561 | alvaro 2 | 330 | 100 | 21 | 3481000 |
| 117565 | sappitu cell | 287 | 152 | 13 | 69702280 |
| 117568 | caapr cell | 344 | 95 | 232 | 130600000 |
| 117698 | soft cell | 321 | 100 | 20 | 3315000 |
| 117700 | joeragan cell | 324 | 97 | 21 | 3494000 |
| 124312 | ardan cell | 344 | 95 | 199 | 64609500 |
| 124313 | naura cell | 226 | 142 | 12 | 1750000 |
| 124314 | putra agung | 331 | 104 | 28 | 4750000 |
| 124316 | ira cell | 331 | 97 | 24 | 3848000 |
| 124318 | taro cell | 344 | 95 | 174 | 93700000 |
| 124546 | RESMI CELL | 323 | 99 | 31 | 9560000 |
| 124548 | NOVA CELL | 308 | 120 | 24 | 4163000 |

LAMPIRAN D

| ID | LNorm | Rnorm_Rev | Fnorm | Mnorm |
|-----------|--------------|------------------|--------------|--------------|
| 103072 | 0.918919 | 0.942446 | 0.430962 | 0.002788 |
| 103073 | 0.564189 | 0.730216 | 0.100418 | 0.002251 |
| 103473 | 0.956081 | 0.996403 | 0.16318 | 0.00284 |
| 103478 | 0.945946 | 0.946043 | 0.535565 | 0.008508 |
| 103741 | 0.641892 | 0.971223 | 0.108787 | 0.000774 |
| 103742 | 0.902027 | 0.985612 | 0.066946 | 0.00048 |
| 103743 | 0.861486 | 0.866906 | 0.322176 | 0.003329 |
| 103745 | 0.800676 | 0.823741 | 0.033473 | 0.000369 |
| 103747 | 0.891892 | 0.935252 | 0.158996 | 0.001911 |
| 103841 | 0.827703 | 0.877698 | 0.075314 | 0.001269 |
| 103842 | 0.986486 | 1 | 0.410042 | 0.059942 |
| 103843 | 0.625 | 0.953237 | 0.025105 | 0.000147 |
| 103844 | 0.962838 | 1 | 0.493724 | 0.010048 |
| 103850 | 0.709459 | 0.741007 | 0.033473 | 0.00024 |
| 103904 | 0.658784 | 0.94964 | 0.025105 | 0.000205 |
| 103910 | 0.625 | 0.953237 | 0.029289 | 0.000174 |
| 105971 | 0.503378 | 0.514388 | 0.008368 | 5.18E-05 |
| 105974 | 0.922297 | 0.982014 | 0.322176 | 0.002761 |
| 106055 | 0.875 | 0.94964 | 0.146444 | 0.000982 |
| 106308 | 0.956081 | 0.992806 | 0.112971 | 0.083217 |
| 106374 | 0.922297 | 0.985612 | 0.179916 | 0.002541 |
| 106377 | 0.908784 | 0.910072 | 0.1841 | 0.009486 |
| 106496 | 0.287162 | 0.244604 | 0.041841 | 0.001021 |
| 109441 | 0.760135 | 0.910072 | 0.07113 | 0.000446 |
| 109442 | 0.966216 | 0.967626 | 0.066946 | 0.045531 |
| | | | | |

D-2

| ID | Lnorm | Rnorm_Rev | Fnorm | Mnorm |
|-----------|--------------|------------------|--------------|--------------|
| 109444 | 0.972973 | 0.985612 | 0.104603 | 0.012217 |
| 109445 | 0.939189 | 0.942446 | 0.221757 | 0.018183 |
| 109446 | 0.834459 | 0.874101 | 0.133891 | 0.001929 |
| 109464 | 0.912162 | 0.906475 | 0.158996 | 0.00403 |
| 109551 | 0.746622 | 1 | 0.171548 | 0.001084 |
| 109552 | 0.746622 | 1 | 0.1841 | 0.001136 |
| 109553 | 0.72973 | 0.982014 | 0.267782 | 0.001661 |
| 109558 | 0.72973 | 0.982014 | 0.330544 | 0.002053 |
| 109564 | 0.655405 | 0.942446 | 0.07113 | 0.000858 |
| 109651 | 0.922297 | 0.982014 | 0.301255 | 0.001964 |
| 109716 | 0.652027 | 0.931655 | 0.058577 | 0.002358 |
| 109813 | 0.966216 | 1 | 0.083682 | 0.000551 |
| 110892 | 0.824324 | 1 | 0.34728 | 0.004428 |
| 110895 | 0.608108 | 0.910072 | 0.121339 | 0.000943 |
| 111005 | 0.952703 | 0.985612 | 0.196653 | 0.001666 |
| 111006 | 0.817568 | 1 | 0.334728 | 0.004292 |
| 111028 | 0.824324 | 0.827338 | 0.196653 | 0.003669 |
| 111063 | 0.804054 | 0.985612 | 0.276151 | 0.001723 |
| 111142 | 0.912162 | 0.992806 | 0.364017 | 0.002503 |
| 111144 | 0.939189 | 0.996403 | 0.087866 | 0.00063 |
| 111146 | 0.966216 | 0.971223 | 0.171548 | 0.001876 |
| 111148 | 0.942568 | 0.982014 | 0.079498 | 0.000519 |
| 111156 | 0.817568 | 1 | 0.330544 | 0.004294 |
| 111201 | 0.915541 | 1 | 0.200837 | 0.001354 |
| 111271 | 0.888514 | 1 | 0.297071 | 0.003604 |
| 111279 | 0.959459 | 0.967626 | 0.196653 | 0.002255 |
| 111412 | 0.939189 | 1 | 0.292887 | 0.00365 |
| 111417 | 0.935811 | 1 | 0.075314 | 0.000481 |
| 111511 | 0.966216 | 1 | 0.083682 | 0.000542 |

| ID | Lnorm | Rnorm_Rev | Fnorm | Mnorm |
|-----------|--------------|------------------|--------------|--------------|
| 111512 | 0.952703 | 0.982014 | 0.07113 | 0.000439 |
| 111514 | 0.945946 | 0.992806 | 0.075314 | 0.000498 |
| 111515 | 0.706081 | 1 | 0.67364 | 0.006321 |
| 111521 | 0.817568 | 1 | 0.309623 | 0.004043 |
| 111530 | 0.716216 | 0.758993 | 0.523013 | 0.004636 |
| 111572 | 0.820946 | 1 | 0.334728 | 0.004289 |
| 111575 | 0.989865 | 0.992806 | 0.338912 | 0.00477 |
| 111578 | 0.777027 | 0.766187 | 0.16318 | 0.001045 |
| 111639 | 0.945946 | 1 | 0.075314 | 0.000506 |
| 117202 | 0.442568 | 0.827338 | 0.020921 | 0.000189 |
| 117209 | 0.412162 | 0.827338 | 0.020921 | 0.000112 |
| 117292 | 0.912162 | 0.996403 | 0.167364 | 0.001952 |
| 117513 | 0.564189 | 0.94964 | 0.033473 | 0.000198 |
| 117514 | 0.97973 | 0.985612 | 0.301255 | 0.003409 |
| 117515 | 0.912162 | 0.996403 | 0.171548 | 0.001173 |
| 117518 | 0.945946 | 0.992806 | 0.292887 | 0.003438 |
| 117519 | 0.97973 | 0.982014 | 0.07113 | 0.008964 |
| 117520 | 0.929054 | 1 | 0.179916 | 0.001291 |
| 117561 | 0.949324 | 0.982014 | 0.079498 | 0.00052 |
| 117565 | 0.804054 | 0.794964 | 0.046025 | 0.011874 |
| 117568 | 0.996622 | 1 | 0.962343 | 0.022315 |
| 117698 | 0.918919 | 0.982014 | 0.075314 | 0.000492 |
| 117700 | 0.929054 | 0.992806 | 0.079498 | 0.000522 |
| 124312 | 0.996622 | 1 | 0.824268 | 0.011001 |
| 124313 | 0.597973 | 0.830935 | 0.041841 | 0.000223 |
| 124314 | 0.952703 | 0.967626 | 0.108787 | 0.000738 |
| 124316 | 0.952703 | 0.992806 | 0.09205 | 0.000583 |
| 124318 | 0.996622 | 1 | 0.719665 | 0.015989 |
| 124546 | 0.925676 | 0.985612 | 0.121339 | 0.001562 |

LAMPIRAN E

| ID | Cluster | LBot | RBot | FBot | MBot | CLV |
|-----------|----------------|-------------|-------------|-------------|-------------|------------|
| 103072 | 2 | 0.101081 | 0.056547 | 0.133598 | 0.001422 | 0.292648 |
| 103073 | 1 | 0.062061 | 0.043813 | 0.03113 | 0.001148 | 0.138151 |
| 103473 | 3 | 0.105169 | 0.059784 | 0.050586 | 0.001448 | 0.216987 |
| 103478 | 2 | 0.104054 | 0.056763 | 0.166025 | 0.004339 | 0.331181 |
| 103741 | 3 | 0.070608 | 0.058273 | 0.033724 | 0.000395 | 0.163 |
| 103742 | 3 | 0.099223 | 0.059137 | 0.020753 | 0.000245 | 0.179358 |
| 103743 | 2 | 0.094764 | 0.052014 | 0.099874 | 0.001698 | 0.24835 |
| 103745 | 3 | 0.088074 | 0.049424 | 0.010377 | 0.000188 | 0.148064 |
| 103747 | 3 | 0.098108 | 0.056115 | 0.049289 | 0.000975 | 0.204487 |
| 103841 | 3 | 0.091047 | 0.052662 | 0.023347 | 0.000647 | 0.167704 |
| 103842 | 2 | 0.108514 | 0.06 | 0.127113 | 0.03057 | 0.326197 |
| 103843 | 1 | 0.06875 | 0.057194 | 0.007782 | 7.49E-05 | 0.133802 |
| 103844 | 2 | 0.105912 | 0.06 | 0.153054 | 0.005124 | 0.324091 |
| 103850 | 1 | 0.078041 | 0.04446 | 0.010377 | 0.000123 | 0.133 |
| 103904 | 3 | 0.072466 | 0.056978 | 0.007782 | 0.000104 | 0.137331 |
| 103910 | 1 | 0.06875 | 0.057194 | 0.009079 | 8.88E-05 | 0.135112 |
| 105971 | 1 | 0.055372 | 0.030863 | 0.002594 | 2.64E-05 | 0.088855 |
| 105974 | 2 | 0.101453 | 0.058921 | 0.099874 | 0.001408 | 0.261656 |
| 106055 | 3 | 0.09625 | 0.056978 | 0.045397 | 0.000501 | 0.199127 |
| 106308 | 3 | 0.105169 | 0.059568 | 0.035021 | 0.042441 | 0.242199 |
| 106374 | 3 | 0.101453 | 0.059137 | 0.055774 | 0.001296 | 0.217659 |
| 106377 | 3 | 0.099966 | 0.054604 | 0.057071 | 0.004838 | 0.21648 |
| 106496 | 1 | 0.031588 | 0.014676 | 0.012971 | 0.00052 | 0.059755 |
| 109441 | 3 | 0.083615 | 0.054604 | 0.02205 | 0.000228 | 0.160497 |
| 109442 | 3 | 0.106284 | 0.058058 | 0.020753 | 0.023221 | 0.208315 |
| 109444 | 3 | 0.107027 | 0.059137 | 0.032427 | 0.00623 | 0.204821 |
| 109445 | 3 | 0.103311 | 0.056547 | 0.068745 | 0.009273 | 0.237876 |
| 109446 | 3 | 0.091791 | 0.052446 | 0.041506 | 0.000984 | 0.186727 |
| 109464 | 3 | 0.100338 | 0.054388 | 0.049289 | 0.002055 | 0.20607 |
| 109551 | 3 | 0.082128 | 0.06 | 0.05318 | 0.000553 | 0.195861 |
| | | | | | | |

E-2

| ID | Cluster | LBot | RBot | FBot | MBot | CLV |
|-----------|----------------|-------------|-------------|-------------|-------------|------------|
| 109552 | 3 | 0.082128 | 0.06 | 0.057071 | 0.000579 | 0.199779 |
| 109553 | 3 | 0.08027 | 0.058921 | 0.083013 | 0.000847 | 0.223051 |
| 109558 | 2 | 0.08027 | 0.058921 | 0.102469 | 0.001047 | 0.242707 |
| 109564 | 3 | 0.072095 | 0.056547 | 0.02205 | 0.000437 | 0.151129 |
| 109651 | 3 | 0.101453 | 0.058921 | 0.093389 | 0.001002 | 0.254765 |
| 109716 | 3 | 0.071723 | 0.055899 | 0.018159 | 0.001203 | 0.146984 |
| 109813 | 3 | 0.106284 | 0.06 | 0.025941 | 0.000281 | 0.192506 |
| 110892 | 2 | 0.090676 | 0.06 | 0.107657 | 0.002258 | 0.260591 |
| 110895 | 1 | 0.066892 | 0.054604 | 0.037615 | 0.000481 | 0.159592 |
| 111005 | 3 | 0.104797 | 0.059137 | 0.060962 | 0.00085 | 0.225746 |
| 111006 | 2 | 0.089932 | 0.06 | 0.103766 | 0.002189 | 0.255887 |
| 111028 | 3 | 0.090676 | 0.04964 | 0.060962 | 0.001871 | 0.20315 |
| 111063 | 3 | 0.088446 | 0.059137 | 0.085607 | 0.000879 | 0.234068 |
| 111142 | 2 | 0.100338 | 0.059568 | 0.112845 | 0.001276 | 0.274028 |
| 111144 | 3 | 0.103311 | 0.059784 | 0.027238 | 0.000321 | 0.190655 |
| 111146 | 3 | 0.106284 | 0.058273 | 0.05318 | 0.000957 | 0.218694 |
| 111148 | 3 | 0.103682 | 0.058921 | 0.024644 | 0.000265 | 0.187512 |
| 111156 | 2 | 0.089932 | 0.06 | 0.102469 | 0.00219 | 0.254591 |
| 111201 | 3 | 0.100709 | 0.06 | 0.062259 | 0.000691 | 0.22366 |
| 111271 | 3 | 0.097736 | 0.06 | 0.092092 | 0.001838 | 0.251666 |
| 111279 | 3 | 0.105541 | 0.058058 | 0.060962 | 0.00115 | 0.22571 |
| 111412 | 3 | 0.103311 | 0.06 | 0.090795 | 0.001861 | 0.255967 |
| 111417 | 3 | 0.102939 | 0.06 | 0.023347 | 0.000245 | 0.186532 |
| 111511 | 3 | 0.106284 | 0.06 | 0.025941 | 0.000276 | 0.192502 |
| 111512 | 3 | 0.104797 | 0.058921 | 0.02205 | 0.000224 | 0.185992 |
| 111514 | 3 | 0.104054 | 0.059568 | 0.023347 | 0.000254 | 0.187223 |
| 111515 | 2 | 0.077669 | 0.06 | 0.208828 | 0.003224 | 0.349721 |
| 111521 | 2 | 0.089932 | 0.06 | 0.095983 | 0.002062 | 0.247978 |
| 111530 | 2 | 0.078784 | 0.04554 | 0.162134 | 0.002364 | 0.288822 |
| 111572 | 2 | 0.090304 | 0.06 | 0.103766 | 0.002187 | 0.256257 |
| 111575 | 2 | 0.108885 | 0.059568 | 0.105063 | 0.002433 | 0.275949 |
| 111578 | 3 | 0.085473 | 0.045971 | 0.050586 | 0.000533 | 0.182563 |
| 111639 | 3 | 0.104054 | 0.06 | 0.023347 | 0.000258 | 0.187659 |
| 117202 | 1 | 0.048682 | 0.04964 | 0.006485 | 9.64E-05 | 0.104904 |

| ID | Cluster | LBot | RBot | FBot | MBot | CLV |
|--------|---------|----------|----------|----------|----------|----------|
| 117209 | 1 | 0.045338 | 0.04964 | 0.006485 | 5.70E-05 | 0.10152 |
| 117292 | 3 | 0.100338 | 0.059784 | 0.051883 | 0.000996 | 0.213001 |
| 117513 | 1 | 0.062061 | 0.056978 | 0.010377 | 0.000101 | 0.129517 |
| 117514 | 2 | 0.10777 | 0.059137 | 0.093389 | 0.001739 | 0.262035 |
| 117515 | 3 | 0.100338 | 0.059784 | 0.05318 | 0.000598 | 0.2139 |
| 117518 | 3 | 0.104054 | 0.059568 | 0.090795 | 0.001753 | 0.256171 |
| 117519 | 3 | 0.10777 | 0.058921 | 0.02205 | 0.004572 | 0.193313 |
| 117520 | 3 | 0.102196 | 0.06 | 0.055774 | 0.000659 | 0.218629 |
| 117561 | 3 | 0.104426 | 0.058921 | 0.024644 | 0.000265 | 0.188256 |
| 117565 | 3 | 0.088446 | 0.047698 | 0.014268 | 0.006056 | 0.156467 |
| 117568 | 2 | 0.109628 | 0.06 | 0.298326 | 0.011381 | 0.479336 |
| 117698 | 3 | 0.101081 | 0.058921 | 0.023347 | 0.000251 | 0.1836 |
| 117700 | 3 | 0.102196 | 0.059568 | 0.024644 | 0.000266 | 0.186675 |
| 124312 | 2 | 0.109628 | 0.06 | 0.255523 | 0.00561 | 0.430762 |
| 124313 | 1 | 0.065777 | 0.049856 | 0.012971 | 0.000114 | 0.128718 |
| 124314 | 3 | 0.104797 | 0.058058 | 0.033724 | 0.000376 | 0.196955 |
| 124316 | 3 | 0.104797 | 0.059568 | 0.028536 | 0.000297 | 0.193199 |
| 124318 | 2 | 0.109628 | 0.06 | 0.223096 | 0.008154 | 0.400879 |
| 124546 | 3 | 0.101824 | 0.059137 | 0.037615 | 0.000797 | 0.199373 |
| 124548 | 3 | 0.09625 | 0.054604 | 0.028536 | 0.000325 | 0.179715 |
| 125800 | 3 | 0.088446 | 0.049424 | 0.011674 | 9.64E-05 | 0.14964 |
| 125820 | 3 | 0.101081 | 0.06 | 0.063556 | 0.000666 | 0.225304 |
| 125910 | 2 | 0.105912 | 0.06 | 0.251632 | 0.002915 | 0.420459 |
| 125931 | 2 | 0.089932 | 0.06 | 0.112845 | 0.002547 | 0.265325 |
| 125932 | 3 | 0.093277 | 0.055683 | 0.08431 | 0.003686 | 0.236956 |
| 125937 | 1 | 0.062061 | 0.043813 | 0.03113 | 0.001148 | 0.138151 |
| 125993 | 2 | 0.089932 | 0.06 | 0.107657 | 0.002283 | 0.259872 |
| 125995 | 1 | 0.063919 | 0.057626 | 0.006485 | 8.76E-05 | 0.128118 |
| 125997 | 1 | 0.049426 | 0.025468 | 0.019456 | 0.000309 | 0.094658 |
| 126001 | 2 | 0.109628 | 0.06 | 0.156946 | 0.006804 | 0.333378 |
| 126002 | 3 | 0.099966 | 0.059568 | 0.024644 | 0.000271 | 0.18445 |
| 126004 | 1 | 0.058345 | 0.057194 | 0.010377 | 0.000101 | 0.126016 |
| 126012 | 3 | 0.100709 | 0.059568 | 0.079121 | 0.000859 | 0.240258 |
| 126015 | 2 | 0.109257 | 0.059784 | 0.120628 | 0.002527 | 0.292196 |

LAMPIRAN F

```

library(shiny)
library(leaflet)
library(plot3Drgl)
library(sqldf)
library(shinydashboard)
library(plotly)

dashboardPage(skin = "green", dashboardHeader(title = "PT.
XYZ Gresik"),
  dashboardSidebar(
    sidebarMenu(
      menuItem("Halaman Utama", tabName =
"home", icon = icon("home")),
      menuItem("Visualisasi Segmen", tabName =
"detail", icon = icon("dashboard")),
      menuItem("Tabel Data", tabName = "tabel", icon =
icon("table"))
    ),
    dashboardBody(
      tabItems(
        tabItem("home",
          fluidRow(
            valueBoxOutput("ket1", width=6),
            valueBoxOutput("ket3", width=6),
            column(width = 12,
              align = "center",
              strong(h3("Visualisasi Segmentasi
Retailer PT. XYZ Gresik")))
          ),
        fluidRow(
          tabBox(
            title = "Jumlah Retailer",
            id = "tabChart1",
            tabPanel("Segmen",
              plotlyOutput("barchart")),

```

```

tabPanel("Area", plotlyOutput("barchartall")),
  tabPanel("Area", plotlyOutput("barchart1")),
  tabPanel("Area", plotlyOutput("barchart2")),
  tabPanel("Area", plotlyOutput("barchart3"))
),
tabBox( title = "% Retailer", side = "right",
  id = "tabChart2",
  tabPanel("Segmen", plotlyOutput("pie")),
  tabPanel("Area", plotlyOutput("pieall")),
  tabPanel("Area 1", plotlyOutput("pie1")),
  tabPanel("Area 2", plotlyOutput("pie2")),
  tabPanel("Area 3", plotlyOutput("pie3"))
)),
tabItem("detail",
  fluidPage(
    column(width=12,
      box(width = NULL,
        background = "olive",
        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("length",width=6),
        infoBoxOutput("recency",width=6)
      ),
      fluidRow(
        infoBoxOutput("frequency",width=6),
        infoBoxOutput("monetary",width=6)
      ),

```

```

box( width = 100,
      title = "Rentang Segmen",
      background = "olive",
      solidHeader = TRUE,
      collapsible = TRUE,
      collapsed = FALSE,
      strong(textOutput('text')),
      verbatimTextOutput('range')
    ),

box( width = 100,
      title = "Retailer Teratas dan Terbawah",
      background = "olive",
      solidHeader = TRUE,
      collapsible = TRUE,
      collapsed = FALSE,
      valueBoxOutput("teratas",width=6),
valueBoxOutput("terbawah",width=6)
    ),

fluidRow(
  column(width=12,
    box(width = 800,
      title = "Peringkat Retailer",
      solidHeader = TRUE,
      collapsible = TRUE,
      collapsed = FALSE,
      dataTableOutput("clv")
    )
  )
),
fluidRow(
  tabBox(
    title = "Box-Plot",
    # The id lets us use input$tabset1 on the
server to find the current tab
    id = "tabChart3",
    #tabPanel("CLV",
plotlyOutput("boxclv")),

```

