

## Pengelompokan Pelanggan Toko Kerajinan Menggunakan K-Means dengan Model RFM dan LRFM

I Kadek Ari Surya<sup>a1</sup>, Made Agung Raharja<sup>a2</sup>, I Komang Ari Mogi<sup>a3</sup>, Agus Muliantara<sup>a4</sup>, I Gede Arta Wibawa<sup>a5</sup>, I Gusti Ngurah Anom Cahyadi Putra<sup>a6</sup>

<sup>a</sup>Department of Informatics, Faculty of Mathematics and Natural Sciences, Udayana University  
South Kuta, Badung, Bali, Indonesia

<sup>1</sup>ikd.arisurya@gmail.com

<sup>2</sup>made.agung@unud.ac.id

<sup>3</sup>arimogi@unud.ac.id

<sup>4</sup>muliantara@unud.ac.id

<sup>5</sup>gede.arta@unud.ac.id

<sup>6</sup>anom.cp@unud.ac.id

### Abstract

*Customer groups in a shop are important to identify to determine sales strategy. The research was conducted at one of the handicraft shops in Bali which was affected by intense market competition. The purpose of this research is to get the best customer cluster and data model. The stages in this research start from preprocessing customer data to generate Recency, Frequency, Monetary (RFM) and Length, Recency, Frequency, Monetary (LRFM) models, then clustering with K-Means and evaluating cluster quality with Silhouette Coefficient (SC). The results showed the RFM model produces an SC value of 0.545 (medium structure) when  $k = 2$  and the LRFM model with the largest SC value at  $k = 3$  with a value of 0.415 (weak structure). SC values in both models tend to increase as the percentage of data increases. Cluster 1 has 817 members with the last transaction taking a long time, but has below average orders and monetary. Cluster 2 has 158 members with the last transaction at the most recent time, and with above average orders and monetary.*

**Keywords:** Clustering, Customer, Craft, K-Means, RFM, LRFM, Silhouette Coefficient

### 1. Pendahuluan

Mengenali karakteristik kelompok pelanggan dalam suatu bidang bisnis menjadi hal yang penting untuk dapat mempertahankan pelanggan. Karakteristik pelanggan dapat dijadikan acuan dalam menentukan strategi pemasaran dalam menjaga aliran bisnis tetap terjaga pada suatu perusahaan. Hal ini perlu dipertimbangkan oleh berbagai bidang usaha mengingat ketatnya persaingan pasar pada era sekarang.

Industri kerajinan merupakan bidang usaha dengan perkembangan inovasi yang begitu pesat dan persaingan pasar yang ketat. Bidang industri kerajinan terhitung sebagai salah satu bidang yang berperan besar terhadap perekonomian masyarakat. Pemulihan ekonomi daerah harus didukung oleh peran aktif masyarakat dan pemerintah agar dapat menghasilkan kondisi perekonomian yang lebih baik [1]. Bali merupakan daerah di Indonesia yang aktif memproduksi produk kerajinan. Namun beberapa tahun belakangan ini dari data Dinas Perdagangan dan Perindustrian (Disperindag) Provinsi Bali pada tahun 2019 terjadi penurunan permintaan export sebesar 13% dan tahun 2020 sebesar 29.44% [2] [3]. Kondisi ini berdampak pada industri kecil lainnya salah satunya toko kerajinan oleh2bali.com. Toko ini memiliki pencatatan data pelanggan dan order yang dapat diolah dengan data mining yakni teknik *clustering* untuk dapat membantu mengenali karakteristik kelompok pelanggan.

Salah satu metode yang populer digunakan untuk mengelompokkan data adalah K-Means. Penelitian pada kasus *customer segmentation* menggunakan perbandingan K-Means, K-Medoids, dan DBSCAN menunjukkan hasil bahwa algoritma K-Means memiliki akurasi paling baik [4]. Metode K-Means sering digunakan karena bersifat dinamis pada data tersebar serta mudah diterapkan dan diinterpretasikan

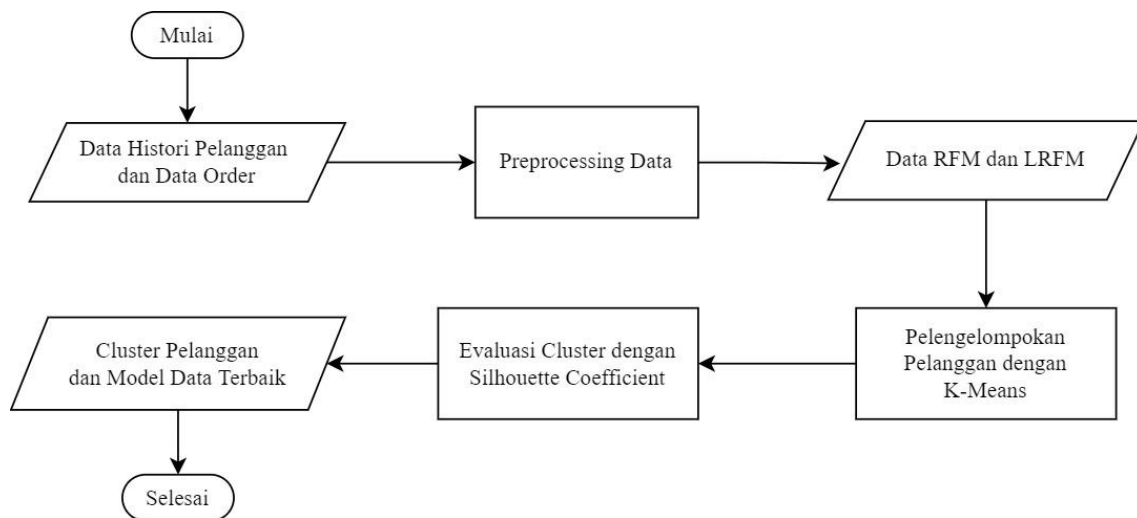
[5]. Dalam mengimplementasikan teknik *clustering* menggunakan metode tertentu seperti K-Means dapat menerapkan beberapa model data.

RFM adalah model data yang dapat digunakan untuk menganalisis perilaku pelanggan, seperti seberapa baru pelanggan melakukan pembelian (*Recency*), seberapa sering pelanggan membeli (*Frequency*), dan berapa banyak uang yang dihabiskan pelanggan dalam melakukan transaksi (*Monetary*) [4]. Model RFM telah banyak digunakan dalam dunia pemasaran terutama untuk segmentasi pasar [6]. Selain model RFM terdapat juga model LRFM. Model LRFM adalah model yang berasal dari pengembangan RFM dengan menambahkan variabel *Length*. Model LRFM merupakan metode analisis nilai pelanggan yang diterapkan pada segmentasi pelanggan [7]. Model LRFM membedakan antara pelanggan yang mempunyai hubungan jangka panjang atau jangka pendek dalam suatu perusahaan [8].

Berdasarkan penjelasan di atas pada penelitian ini bertujuan untuk mengetahui cluster pelanggan dengan menggunakan metode K-Means dan mengetahui model data terbaik dengan membandingkan hasil clustering dengan model data RFM dan LRFM pada proses clustering pelanggan di toko kerajinan tempat studi kasus.

## 2. Metode Penelitian

Proses penelitian dimulai dari pengumpulan data, *preprocessing* data, proses pengelompokan pelanggan menggunakan K-Means, dan perbandingan perhitungan *silhouette coefficient* digunakan dalam menentukan kualitas cluster hasil dari model RFM dan LRFM. Tahapan penelitian pada dilihat pada *flowchart* yang ditunjukkan pada Gambar 1.



**Gambar 1.** Desain Umum Penelitian

### 2.1 Pengumpulan Data

Data pada penelitian menggunakan data sekunder yang diperoleh dari toko kerajinan oleh2bali.com yang berlokasi di Denpasar, Bali. Data yang didapatkan adalah data histori pelanggan dan data order dari tahun 2014-2020. Terdapat 975 data pelanggan dan 2463 data order. Format data berupa *comma separated values* (.csv). Data histori pelanggan terdiri dari atribut *id\_customer*, *name*, *last\_active*, *email*, *orders*, *total\_spend*, *aov*, *country*, *city*, *region*. Sedangkan data order terdiri atribut *id\_order*, *id\_customer*, *name*, *status*, *date*, *total*. Berikut ini adalah atribut dari data yang didapatkan yang ditunjukkan pada Tabel 1 dan Tabel 2:

**Tabel 1.** Deskripsi Data Histori Pelanggan

Nama Field	Tipe Data	Keterangan
<i>id_customer</i>	<i>string</i>	Id unik dari pelanggan.
<i>name</i>	<i>string</i>	Nama pelanggan.

<i>last_active</i>	<i>datetime</i>	Tanggal terakhir pelanggan melakukan pemesanan barang.
<i>email</i>	<i>string</i>	Alamat email pelanggan
<i>order</i>	<i>integer</i>	Total order yang sudah terbayarkan oleh pelanggan.
<i>total_spend</i>	<i>integer</i>	Total uang yang dihabiskan oleh pelanggan untuk total order produk.
<i>aov</i>	<i>integer</i>	<i>Average Order Value</i> adalah rata-rata pembelian pelanggan.
<i>country</i>	<i>string</i>	Negara asal dari pelanggan.
<i>city</i>	<i>string</i>	Kota asal dari pelanggan.
<i>region</i>	<i>string</i>	Provinsi asal dari pelanggan.

**Table 2.** Deskripsi Data Order

Nama Field	Tipe Data	Keterangan
<i>id_order</i>	<i>string</i>	Id unik dari setiap order.
<i>id_customer</i>	<i>string</i>	Id unik dari customer.
<i>name</i>	<i>string</i>	Nama pelanggan yang melakukan order.
<i>status</i>	<i>string</i>	Status order yang menunjukkan sudah completed atau belum completed order.
<i>date</i>	<i>date</i>	Tanggal order dilakukan.
<i>total</i>	<i>integer</i>	Total uang yang harus dibayarkan untuk melakukan setiap order.

## 2.2 Preprocessing Data

*Preprocessing* data merupakan proses perbaikan untuk mengolah data mentah menjadi format data yang dibutuhkan pada penelitian. Pada penelitian ini *preprocessing* data dimulai dari melakukan *data cleaning*, *data selection* dan *data transformation*. Pada *data cleaning* dilakukan penghapusan *record* data yang memiliki nilai *null* atau *missing value*. Kemudian pada *data selection* dipilih beberapa atribut data yang dibutuhkan pada penelitian terutama untuk memperoleh model data RFM dan LRFM. Berikut ini adalah penjelasan mengenai atribut *recency* (R), *frequency* (F), *monetary* (M), dan *length* (L).

### a. *Recency* (R)

*Recency* merupakan jarak waktu antara waktu penelitian dengan waktu terakhir pelanggan melakukan transaksi. Pelanggan yang paling baru memiliki nilai *recency* yang kecil dan berpotensi dalam melakukan pembelian lagi jika dibandingkan pelanggan dengan waktu terakhir transaksi yang cukup lama. *Recency* dapat dicari dengan menggunakan persamaan 1.

$$R_{(i)} = TR - TL_{(i)} \quad (1)$$

Dimana:

$R_{(i)}$  : *recency* pada pelanggan i

$TR$  : tanggal penelitian

$TL_{(i)}$  : tanggal terakhir pelanggan i melakukan transaksi

### b. *Frequency* (F)

*Frequency* merupakan total jumlah transaksi yang dilakukan pada jangka waktu yang ditentukan. Nilai *frequency* yang tinggi menandakan tingginya tingkat loyalitas pelanggan. *Frequency* dapat dicari dengan menggunakan persamaan 2.

$$F_{(i)} = \sum_{t=TA}^{t=TR} fr_{(t)} \quad (2)$$

Dimana:

$F_{(i)}$  : *frequency* pada pelanggan  $i$

$t$  : waktu dalam tanggal

$TA$  : tanggal awal yang ditentukan

$TR$  : tanggal penelitian

$fr_{(t)}$  : jumlah transaksi yang dilakukan pelanggan  $i$  pada setiap waktu  $t$

c. *Monetary* (M)

*Monetary* merupakan nilai total uang yang dihabiskan pelanggan pada periode tertentu dalam melakukan transaksi. Nilai *monetary* yang tinggi menunjukkan bahwa kontribusi pelanggan terbilang besar terhadap perusahaan. *Monetary* dapat dicari dengan menggunakan persamaan 3.

$$M_{(i)} = \sum_{t=TA}^{t=TR} mr_{(t)} \quad (3)$$

Dimana:

$M_{(i)}$  : *monetary* pada pelanggan  $i$

$t$  : waktu dalam tanggal

$TA$  : tanggal awal yang ditentukan

$TR$  : tanggal penelitian

$mr_{(t)}$  : jumlah uang yang dibayarkan pelanggan  $i$  untuk setiap transaksi pada waktu  $t$

d. *Length* (L)

*Length* menyatakan jarak antara waktu transaksi terakhir pada periode tertentu dengan waktu awal pelanggan melakukan transaksi. *Length* dapat dicari dengan menggunakan persamaan 4.

$$L_{(i)} = TL_{(i)} - TF_{(i)} \quad (4)$$

Dimana:

$L_{(i)}$  : *length* pada pelanggan  $i$

$TL_{(i)}$  : tanggal terakhir pelanggan  $i$  melakukan transaksi

$TF_{(i)}$  : tanggal pertama kali pelanggan  $i$  melakukan transaksi

Berdasarkan penjelasan di atas model data RFM dan LRFM dapat dihitung dari ketersediaan data penelitian. Data *recency* didapatkan dari tanggal penelitian dikurangi dengan tanggal *last\_active*, data *frequency* bisa didapatkan dari atribut *order* dan *monetary* bisa didapatkan dari atribut *total\_spend*. Sedangkan untuk atribut data LRFM (*length*, *recency*, *frequency*, *monetary*) yakni atribut *length* nilainya dapat dicari dari atribut *date* pada dataset order saat pertama kali pelanggan melakukan transaksi dikurangi dengan tanggal *last\_active* pada dataset histori pelanggan.

*Preprocessing* dilanjutkan pada proses *data transformation*. Transformasi data dilakukan dengan menyamakan rentang data dari 0 – 1 dengan metode *Min Max Normalization*. Berikut ini merupakan hasil normalisasi data yang ditunjukkan pada Tabel 3 dan Tabel 4.

**Tabel 3.** Normalisasi Data RFM

Id	R	F	M
1	0.00000	1.00000	0.28903
2	0.00590	0.33333	0.10008
3	0.01417	0.16667	0.02802
973	0.89335	0.00000	0.14812
974	0.90437	0.16667	0.28423
975	0.91814	0.33333	0.02482

**Tabel 4.** Normalisasi Data LRFM

Id	L	R	F	M
1	0.70769	0.00000	1.00000	0.28903
2	0.14615	0.00590	0.33333	0.10008
3	0.31154	0.01417	0.16667	0.02802
973	0.00000	0.89335	0.00000	0.14812
974	0.11923	0.90437	0.16667	0.28423
975	0.20769	0.91814	0.33333	0.02482

### 2.3 Pengelompokan Pelanggan dengan K-Means

K-Means merupakan metode *clustering* non hirarki dengan membagi data menjadi satu atau lebih kelompok data (cluster) [9]. Metode K-Means memiliki konsep yang sederhana dan umum digunakan. Selain itu K-Means dapat melakukan pengelompokan data dalam jumlah besar dengan waktu komputasi yang terbilang cepat serta efisien. Berikutnya dijelaskan mengenai tahapan dari metode *clustering* menggunakan algoritma K-Means:

- Tentukan jumlah cluster sebanyak k cluster dan dataset yang akan dikelompokkan.
- Dipilih *initial centroid* sebanyak k data. Pada penelitian ini *initial centroid* dipilih secara random.
- Mencari jarak pada setiap objek data dengan masing-masing *centroid*. Perhitungan jarak dapat dilakukan menggunakan persamaan ukuran jarak *Euclidean Distance* dengan persamaan:

$$d(x,y) = |x - y| = \sqrt{\sum_{i=1}^n d(x_i - y_i)^2} \quad (5)$$

- Mengalokasikan masing-masing objek data ke cluster dengan jarak paling kecil atau minimum.
- Tentukan centroid cluster yang baru dengan menghitung rata-rata dari semua data yang terdapat di dalam cluster tersebut.
- Lakukan langkah c, d, dan e sampai tidak ada perubahan objek data dalam suatu cluster.

### 2.4 Evaluasi Cluster dengan Silhouette Coefficient

*Silhouette Coefficient* diperkenalkan oleh Rousseeuw pada tahun 1987. *Silhouette Coefficient* digunakan untuk mengetahui kualitas dari cluster data yang dihasilkan oleh algoritma *clustering* dengan mengkombinasikan konsep *cohesion* dan *separation*. Tahap perhitungannya sebagai berikut [10]:

- Menghitung rata-rata jarak suatu objek i ke semua objek selain objek tersebut dalam suatu cluster yang dilambangkan dengan a(i) :

$$a_i = \frac{1}{|A|-1} \sum_j \epsilon_{A,j \neq i} d(i,j) \quad (6)$$

Dengan A adalah cluster tempat objek tersebut berada, d(i,j) merupakan jarak antar objek pada cluster yang sama.

- Kemudian cari rata-rata jarak dari objek data tersebut dengan objek lain dalam cluster yang berbeda, kemudian pilih nilai rata-rata jarak yang terkecil :

$$b(i) = (D(i, C)) \quad (7)$$

- Setelah nilai a(i) dan b(i) didapatkan, selanjutnya dapat dihitung *silhouette coefficient* dari objek ke i :

$$s(i) = \frac{b(i)-a(i)}{(a(i),b(i))} \quad (8)$$

Berdasarkan kategori nilai rata-rata *silhouette coefficient* (SC) menurut Rousseeuw, hasil perhitungan nilai SC yang nilainya mendekati 1, menandakan bahwa kualitas cluster yang dihasilkan semakin bagus. Kategori nilai SC berdasarkan pendapat Rousseeuw dapat dilihat sebagai berikut:

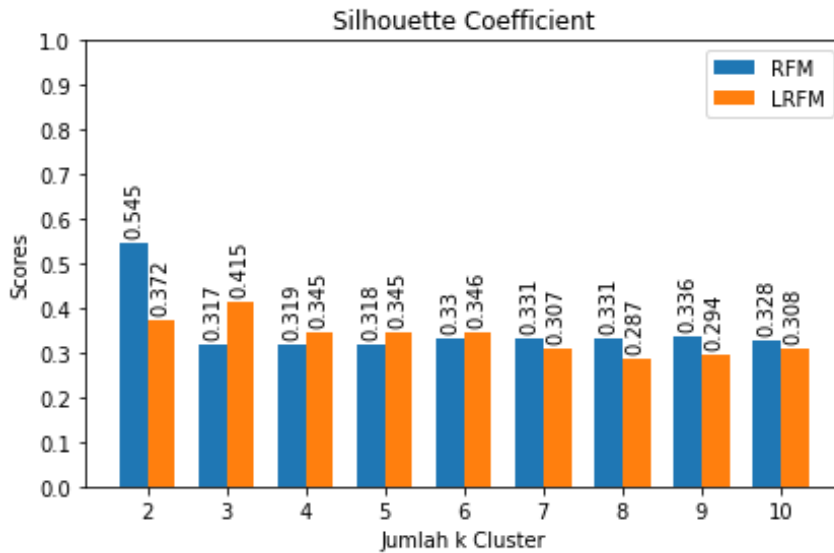
- SC ≤ 0,25 (*no structure*)
- 0,25 < SC ≤ 0,5 (*weak structure*)
- 0,5 < SC ≤ 0,7 (*medium structure*)
- 0,7 < SC ≤ 1 (*strong structure*)

### 3. Hasil dan Pembahasan

Pada bagian ini menjelaskan mengenai perbandingan *silhouette coefficient* yang dihasilkan dari model data RFM dan LRFM. Selain itu juga dijelaskan mengenai hasil pengelompokan terbaik dari setiap model data beserta karakteristiknya.

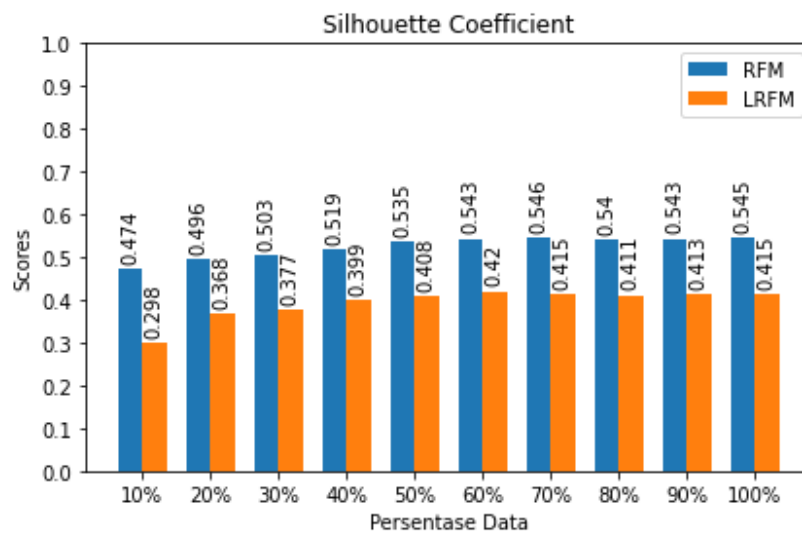
#### 3.1. Perbandingan Silhouette Coefficient

Perbandingan *silhouette coefficient* (SC) model RFM dan LRFM bertujuan untuk mengetahui model data yang menghasilkan hasil *clustering* terbaik. Proses *clustering* dilakukan pada jumlah cluster  $k = 2$  sampai  $k = 10$  untuk mengetahui nilai SC terbaik pada setiap model.



**Gambar 2.** *Silhouette Coefficient* pada Model RFM dan LRFM

Pada Gambar 2 terlihat nilai SC terbesar dihasilkan oleh model RFM pada saat jumlah  $k = 2$  dengan nilai 0.545. Seiring dengan pertambahan jumlah  $k$  cluster nilai SC pada model RFM mengalami penurunan saat  $k = 3$  dan perubahan nilai SC setelahnya tidak terlalu besar. Sedangkan model LRFM hanya menghasilkan nilai SC terbesar 0.415 pada saat jumlah  $k = 3$ . Berikutnya juga dilakukan perbandingan nilai SC pada variasi persentase data.



**Gambar 3.** *Silhouette Coefficient* pada Variasi Persentase Data Model RFM dan LRFM

Perbandingan SC juga dilakukan menggunakan variasi persentase data RFM dan LRFM dari 10% hingga 100% dengan jumlah  $k$  cluster terbaik pada masing-masing model. Pengujian ini dilakukan

untuk mengenali pengaruh dari penggunaan variasi persentase data terhadap nilai SC yang dihasilkan. Dari grafik yang ditampilkan pada Gambar 3 terlihat bahwa model data RFM selalu menghasilkan skor SC lebih besar dibandingkan dengan model data LRFM. Seiring dengan bertambahnya variasi persentase data skor terbesar SC yang dihasilkan pada model data RFM dan LRFM sama-sama cenderung meningkat walaupun tidak terlalu besar.

**Tabel 5.** Keanggotaan Cluster pada Model RFM Berdasarkan *Silhouette Coefficient*

Cluster Pelanggan	<i>Strong Structure</i>	<i>Medium Structure</i>	<i>Weak Structure</i>	<i>No Structure</i>	Total
Cluster 1	504	267	30	16	817
Cluster 2	0	97	40	21	158
Total	504	364	70	37	

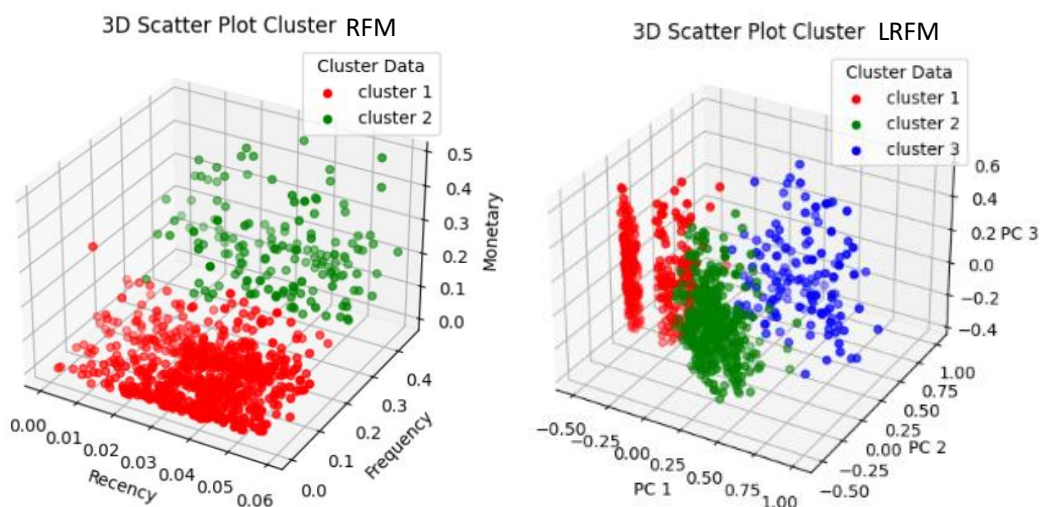
**Tabel 6.** Keanggotaan Cluster pada Model LRFM Berdasarkan *Silhouette Coefficient*

Cluster Pelanggan	<i>Strong Structure</i>	<i>Medium Structure</i>	<i>Weak Structure</i>	<i>No Structure</i>	Total
Cluster 1	0	237	133	69	439
Cluster 2	0	172	154	65	391
Cluster 3	0	0	79	66	145
Total	0	409	366	200	

Pada Tabel 5 dan Tabel 6 merupakan keanggotaan cluster pada model RFM dan LRFM berdasarkan pada nilai SC terbaik pada masing-masing model data. Model RFM menghasilkan cluster pelanggan yang lebih baik dibandingkan menggunakan model LRFM pada toko kerajinan tempat studi kasus karena nilai SC terbesar yang dihasilkan model RFM sebesar 0.545. Selain itu dari struktur keanggotaan cluster, model RFM memiliki *strong structure* sebanyak 504 anggota dan hanya memiliki *weak structure* sebanyak 37 anggota. Dibandingkan dengan hasil clustering terbaik LRFM tidak memiliki anggota *strong structure* tetapi banyak memiliki *no structure* yakni 200 anggota.

### 3.2. Perbandingan Hasil Cluster Terbaik

Berikut ini merupakan visualisasi dari hasil *clustering* terbaik pada setiap model data. Visualisasi ditampilkan melalui *scatterplot*.



**Gambar 4.** *Scatterplot* Cluster Pelanggan dengan Model RFM dan LRFM

Pada Gambar 4 visualisasi cluster pelanggan pada model RFM ditampilkan melalui *3D scatterplot* dengan sumbu *recency*, *frequency*, dan *monetary*. Pada model RFM pengelompokan pelanggan menjadi dua cluster juga terlihat cukup baik karena tidak banyak data anggota cluster yang tumpang

tindih. Sedangkan pada model LRFM memiliki empat atribut data dan untuk dapat menampilkan cluster pelanggan pada *3D scatterplot* data model LRFM perlu dilakukan reduksi dimensi. Reduksi dimensi yang digunakan adalah *Principal Component Analysis* (PCA). Visualisasi cluster pelanggan model LRFM juga terlihat cukup baik namun pengelompokan data pelanggan pada model RFM terlihat lebih baik dibandingkan LRFM. Sehingga berdasarkan pengujian sebelumnya dan visualisasi cluster hasil *clustering* terbaik dihasilkan oleh model RFM. Masing-masing cluster terbaik yang dihasilkan model RFM memiliki karakteristiknya masing-masing.

a. Cluster 1

Cluster 1 memiliki total 817 pelanggan. Berdasarkan *centroid* Cluster 1 pelanggan melakukan pembelian terakhir dalam waktu yang sudah lama, dan memiliki intensitas order dan *monetary* yang di bawah rata-rata. Kemudian berdasarkan nilai aktual dari setiap atribut data pada Cluster 1 adalah pelanggan memiliki rata-rata nilai *recency* sebesar 1492, *frequency* sebesar 2, dan *monetary* sebesar Rp. 65.712. Pelanggan mancanegara dari cluster ini sebanyak 1.5% yang berasal dari Canada, Germany, Prancis, Hongkong, Amerika, Singapore, Malaysia dan Jepang. Berdasarkan tiga wilayah *region* paling banyak berasal dari daerah Jawa Barat, DKI Jakarta, Jawa Timur dengan dominan waktu order pada bulan Oktober, Agustus, Juli yang tepatnya pada pertengahan bulan sekitar jam 11.00 hingga 14.00 siang.

b. Cluster 2

Cluster 2 memiliki total 158 pelanggan. Berdasarkan *centroid* Cluster 2 pelanggan melakukan pembelian terakhir pada waktu terkini, dan memiliki intensitas order dan *monetary* yang di atas rata-rata. Kemudian berdasarkan nilai aktual dari setiap atribut data pada Cluster 2 adalah pelanggan memiliki rata-rata nilai *recency* sebesar 1490, *frequency* sebesar 6, dan *monetary* sebesar Rp. 277.943. Pelanggan mancanegara dari cluster ini sebanyak 3.5% yang berasal dari Malaysia, Canada, dan Hungary. Berdasarkan tiga wilayah *region* paling banyak berasal dari daerah DKI Jakarta, Jawa Barat, Banten dengan dominan waktu order pada bulan Oktober, November, Desember yang tepatnya pada pertengahan bulan sekitar jam 8.00 hingga 10.00 pagi.

#### 4. Kesimpulan

Kesimpulan dari penelitian ini mendapatkan hasil bahwa model RFM menghasilkan kelompok pelanggan terbaik pada toko kerajinan tempat studi kasus dengan nilai *silhouette coefficient* sebesar 0.545 (*medium structure*) saat  $k = 2$  dibandingkan dengan model LRFM yang hanya 0.415 saat  $k = 3$ . Nilai *silhouette coefficient* yang dihasilkan pada setiap model cenderung meningkat ketika terjadi penambahan persentase data. Kemudian terdapat dua kelompok pelanggan terbaik yakni Cluster 1 memiliki anggota sebanyak 817 dengan melakukan pembelian terakhir dalam waktu yang cukup lama, dan memiliki intensitas order dan *monetary* yang di bawah rata-rata. Cluster 2 memiliki anggota sebanyak 158 dengan melakukan transaksi terakhir pada waktu terkini, dan memiliki intensitas order dan *monetary* yang di atas rata-rata.

#### Referensi

- [1] M. A. Raharja, I. D. M. B. A. Darmawan, D. P. E. Nilakusumawati, and I. W. Supriana, "Analysis of membership function in implementation of adaptive neuro fuzzy inference system (ANFIS) method for inflation prediction," *J. Phys. Conf. Ser.*, vol. 1722, no. 1, 2021, doi: 10.1088/1742-6596/1722/1/012005.
- [2] Disperindag Bali, "Laporan Kinerja Instansi Pemerintah (LKjIP) Dinas Perdagangan dan Perindustrian Provinsi Bali 2019," *Dinas Perdagangan dan Perindustrian Provinsi Bali*, 2019, [Online]. Available: <https://disperindag.baliprov.go.id/lkjip-2019-dinas-perindustrian-dan-perdagangan-provinsi-bali/>
- [3] Disperindag Bali, "Laporan Kinerja Instansi Pemerintah (LKjIP) Dinas Perdagangan dan Perindustrian Provinsi Bali 2020," *Dinas Perdagangan dan Perindustrian Provinsi Bali*, 2020, [Online]. Available: <https://disperindag.baliprov.go.id/wp-content/uploads/2021/03/LKjIP-Disperindag-2020.pdf>
- [4] R. W. Sembiring Brahmana, F. A. Mohammed, and K. Chairuang, "Customer Segmentation Based on RFM Model Using K-Means, K-Medoids, and DBSCAN Methods," *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 11, no. 1, 2020, doi: 10.24843/lkjiti.2020.v11.i01.p04.
- [5] A. A. A. Rospricilia, Tita Ayu, Syurfah Ayu Ithriah, "Segmentasi Pelanggan Menggunakan



- Metode K-Means Clustering Berdasarkan Model RFM,” *J. Inform. dan Sist. Inf.*, vol. 1, no. 3, pp. 699–708, 2020.
- [6] B. E. Adiana, I. Soesanti, and A. E. Permanasari, “Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model Dan Teknik Clustering,” *J. Terap. Teknol. Inf.*, vol. 2, no. 1, pp. 23–32, 2018, doi: 10.21460/jutei.2018.21.76.
- [7] S. Monalisa, P. Nadya, and R. Novita, “Analysis for Customer Lifetime Value Categorization with RFM Model,” in *Procedia Computer Science*, 2019, vol. 161, pp. 834–840. doi: 10.1016/j.procs.2019.11.190.
- [8] R. Rahmadiani, A. Dhini, and E. Laoh, “Estimating Customer Lifetime Value using LRFM Model in Pharmaceutical and Medical Device Distribution Company,” in *2020 International Conference on ICT for Smart Society (ICISS)*, 2020, pp. 1–5. doi: 10.1109/ICISS50791.2020.9307592.
- [9] M. A. Raharja and I. W. Supriana, “Analisis Klasifikasi Tingkat Kesehatan Lembaga Perkreditan Desa (Lpd) Menggunakan Metode K-Means Clustering,” *J. Teknol. Inf. dan Komput.*, vol. 5, no. 1, pp. 83–90, 2019, doi: 10.36002/jutik.v5i1.641.

*This page is intentionally left blank*