

Menganalisa Pola Pembelian Pelanggan Menggunakan Metode *Data Mining*: *Customer Segmentation* dan *Market Basket Analysis* Dengan Algoritme K-Means dan Eclat

Dewi Kumala Sari¹, I Kadek Dwi Nuryana²

^{1,2} S1 Sistem Informasi, Universitas Negeri Surabaya

¹dewi.17051214006@unesa.ac.id

²dwinuryana@unesa.ac.id

Abstrak— *Data mining* merupakan proses pencarian pola atau informasi yang menarik pada suatu data menggunakan teknik atau metode tertentu[1]. Dengan menerapkan *data mining* toko swalayan dapat mengolah data transaksinya menjadi informasi yang berharga guna meningkatkan penjualan toko. Dua metode *data mining* yang sering digunakan dalam mengelola data transaksi yang besar yaitu metode *customer segmentation* dan *market basket analysis*. Metode *customer segmentation* dapat mengelompokkan pelanggan yang memiliki kesamaan perilaku menjadi beberapa cluster[2]. Sedangkan metode *market basket analysis* dapat mengolah data transaksi tiap *cluster* menjadi beberapa aturan yang kuat, dimana aturan ini dapat digunakan untuk mengidentifikasi peluang penjualan silang (*cross-selling*) dan peluang penjualan tambahan (*up-selling*)[3].

Dalam penelitian ini, data yang digunakan diperoleh dari CV. Aloha Ponorogo pada tahun 2022 sebanyak 997 data pelanggan dan 22.359 data transaksi. Data tersebut diproses menggunakan metode *customer segmentation* dengan algoritme k-means melalui bahasa pemrograman R dan menghasilkan 5 *cluster* berdasarkan kesamaan sifat atau karakteristik pelanggannya. Dan selanjutnya untuk tiap-tiap kelompok *cluster* akan dianalisa menggunakan metode *market basket analysis* dengan algoritme eclat, batas nilai *minimal support* sebesar 3.5% dan nilai *minimal confidence* sebesar 10% dari kelima *cluster* tersebut. Untuk hasil yang diperoleh tiap *cluster* memiliki aturan asosiasi yang berbeda beda, pada *cluster* pertama menghasilkan 8 aturan asosiasi, *cluster* kedua menghasilkan 4 aturan asosiasi, *cluster* ketiga menghasilkan 2 aturan asosiasi, *cluster* keempat menghasilkan 10 aturan asosiasi, dan *cluster* kelima menghasilkan 18 aturan asosiasi.

Output dari hasil penelitian kedua metode tersebut berupa rekomendasi penataan produk dengan tujuan untuk meningkatkan penjualan (*up-selling*) dan penjualan silang (*cross-selling*) dalam pemaketan produk dimana produk sering dibeli secara bersamaan.

Kata Kunci— *Data Mining*, *Customer Segmentation*, *Market Basket Analysis*, *K-Means*, *Eclat*, *R*.

I. PENDAHULUAN

Pada era digital ini, bisnis dan teknologi terus mengalami perkembangan yang pesat. Para pelaku bisnis dihadapkan berbagai tantangan untuk mengembangkan bisnisnya, tantangan terbesar dari pelaku bisnis yaitu kurang pengelolaan dari banyaknya data yang dihasilkan pada seluruh transaksi pelanggan. *Data mining* menjadi kunci untuk memanfaatkan data ini guna memahami perilaku pelanggan, merancang strategi pemasaran dan meningkatkan pengalaman berbelanja baik secara *online* maupun *offline*. Metode *customer segmentation* dan *market basket analysis* dapat membantu pemilik usaha untuk mengetahui preferensi pelanggan dan pola pembelian pelanggan, berdasarkan data transaksi mereka.

Setiap pelanggan memiliki pola, kebutuhan dan perilaku belanja yang berbeda-beda. Membagi kelompok pelanggan kedalam beberapa segmen dapat mempermudah pemilik usaha dalam menganalisa pola pembelian pelanggan. Dengan metode *customer segmentation*, pelanggan akan dibagi menjadi 5 cluster/kelompok yang memiliki karakteristik yang sama. Selanjutnya, dari kelompok-kelompok pelanggan tersebut akan dianalisa pola pembelian produknya berdasarkan nilai *asosiasi rules*, yaitu dimana kemungkinan barang dibeli secara bersamaan pada satu transaksi.

Dengan mengelola dan menganalisa data transaksi pelanggan, akan didapatkan beberapa informasi yang berguna untuk mengidentifikasi hubungan atau keterikatan antar produk yang berguna untuk mengoptimalkan penempatan produk dalam toko serta merancang strategi *cross-selling* dan *up-selling* yang efektif[4]. Pemilik usaha juga memiliki potensi peluang dalam meningkatkan penjualan dengan menargetkan promosi ataupun rekomendasi produk yang tepat kepada setiap kelompok pelanggan.

Metode *customer segmentation* dan *market basket analysis* dapat diproses secara bersamaan secara cepat dan akurat

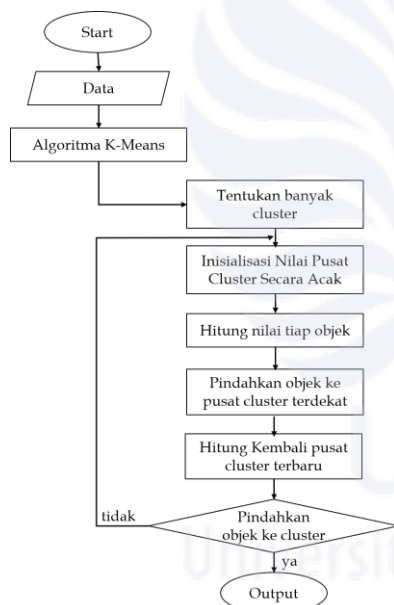
menggunakan bahasa pemrograman R. Bahasa pemrograman R merupakan bahasa yang populer untuk menganalisa data dan memiliki berbagai macam paket dan fungsi yang mendukung dalam pengimplementasian metode *customer segmentation* dan *market basket analysis*. Dengan menggunakan bahasa R, peneliti dapat melakukan proses analisa data secara mudah, efektif dan cepat[5].

II. METODOLOGI

Sumber dalam penelitian ini berasal dari jurnal-jurnal, buku, artikel, maupun website yang berkaitan dengan permasalahan yang ada. Data yang akan digunakan dalam perhitungan penelitian ini yaitu berupa data pelanggan dan data transaksi penjualan produk pada tahun 2022 di CV Aloha Ponorogo.

A. Metode Customer Segmentation menggunakan Algoritme K-Means

Sebelum menerapkan metode *customer segmentation*, berikut diagram *flowchart customer segmentation* menggunakan algoritme k-means pada Gbr.1:

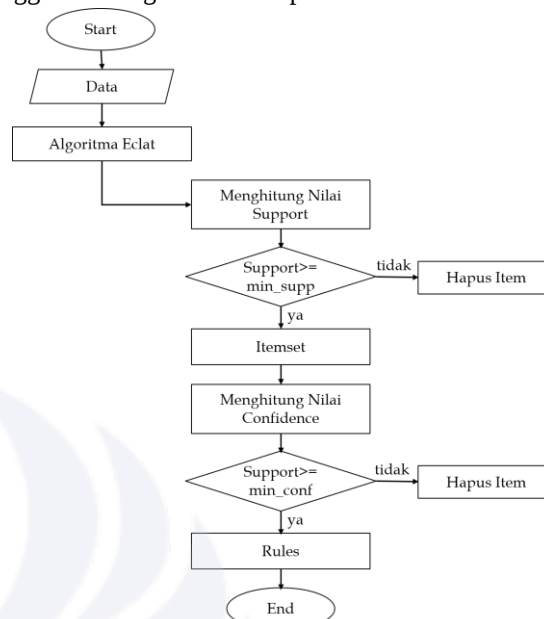


Gbr.1 Flowchat Customer Segmentation dengan Algoritme K-Means

Dalam penerapan metode *customer segmentation* data yang diperlukan yaitu data pelanggan yang berisi nama pelanggan, jenis kelamin, usia, nilai belanja, dan beberapa kriteria lainnya.

B. Metode Market Basket Analysis menggunakan Algoritme Eclat

Sebelum menerapkan metode *market basket analysis*, berikut diagram *flowchart market basket analysis* menggunakan algoritme eclat pada Gbr.2:



Gbr.2 Flowchart Market Basket Analysis Dengan Algoritme Eclat

Dalam penerapan metode *market basket analysis*, data yang diperlukan yaitu berupa data transaksi pelanggan yang berisi id transaksi dan list pembelian produk tiap satu transaksi.

III. IMPLEMENTASI, HASIL DAN PEMBAHASAN

Data yang diperoleh dari CV Aloha Ponorogo pada tahun 2022 yaitu data pelanggan sebanyak 995 pelanggan dan data transaksi sebanyak 22.359 data dari 4.589 transaksi dan 258 item produk. Berikut cuplikan isi data pelanggan dan isi data transaksi pada Tabel I dan Tabel II:

TABEL I
DATA PELANGGAN CV ALOHA PONOROGO PADA TAHUN 2022

no	idCust	namaCust	jk	usia	tnilai
1	C293	Budi	L	24	1650
2	C2659	Watik	P	31	2000
3	C3684	Komari	L	49	2175
4	A15	Antok	L	37	2575
5	F5869	Andri	L	35	2675
6	C475	Sutar	L	61	4150
7	C083	Jarwati	P	65	4250
8	C543	Kasenan	L	19	5950
9	ANG324	Angel	P	49	6000
10	AKB789	Akbar	L	39	7000
11	C535	Eko	L	55	7680
12	EGEG	Roni	L	61	8395
13	C583	Miatun	P	45	8960

14	A005	Abdul	L	52	11700
....					
995	RI01	Rianto	L	46	18245010

TABEL II
DATA TRANSAKSI PENJUALAN CV ALOHA
PONOROGO PADA TAHUN 2022

No	Id Transaksi	Item Produk	Id Cust	Nama Cust
1	2207160001	Mie Beruang	KI	Wito
2	2207160001	Milo	KI	Wito
3	2207160002	Desaku Kunyit	ERW423	Erwin
4	2207160002	Desaku Tambar	ERW423	Erwin
5	2207160002	Kratingdaeng	ERW423	Erwin
6	2207160002	Nabati	ERW423	Erwin
7	2207160002	Nutrisari	ERW423	Erwin
8	2207160002	Padi Mas	ERW423	Erwin
9	2207160003	Kaki Tiga	C236	Watik
10	2207160004	Gula	C038	Didik
14	2207160005	Nuvo	A23	Ami
....				
22359	2208160051	Sun Kara	C522	Imron

Dari data-data tersebut akan dianalisa menggunakan metode *customer segmentation* dan metode *market basket analysis*.

Berikut tahapan untuk menganalisa data transaksi CV Aloha Ponorogo pada tahun 2022 menggunakan metode *customer segmentation* dan metode *market basket analysis*:

A. Pre-processing Data

Baik data pelanggan maupun data transaksi dilakukan tahapan *pre-prosesing* data agar dalam pengolahan data tidak ditemukan kesalahan perhitungan atau mempengaruhi hasil perhitungan. Berikut Gbr.3 kode program menghitung nilai *missing value* file pelanggan.txt pada *software* R Studio dengan bahasa pemrograman R:

```
library(dplyr)
pelanggan <-
read.csv("C:\\skripsi\\pelanggan.txt", sep="\t")

pelanggan[pelanggan == ""] <- NA
anyNA(pelanggan)
is.na(pelanggan)%>% colSums()
```

Gbr.3 Kode program *pre-processing* data pelanggan bahasa pemrograman R

Hasil dari code program diatas yaitu pada Gbr.4 berikut:

```
> anyNA(pelanggan)
[1] FALSE
> is.na(pelanggan)%>% colSums()
kodecust namacust jk usia
0 0 0 0
tnilai
0
```

Gbr.4 Hasil Pengecekan *Missing Value*

Pada data pelanggan tersebut tidak ditemukannya data kosong/ *missing value*, maka data dapat langsung digunakan. Langkah selanjutnya yaitu menghitung nilai *missing value* pada data transaksi.txt. Berikut Gbr.5 kode program untuk proses tersebut:

```
library(dplyr)
transaksi <-
read.csv("C:\\skripsi\\transaksi.txt", sep="\t")

transaksi[transaksi == ""] <- NA
anyNA(transaksi)
is.na(transaksi)%>% colSums()
```

Gbr.5 Kode Program *pre-processing* data transaksi bahasa pemrograman R

Hasil dari kode diatas yaitu pada Gbr.6 berikut:

```
> transaksi[transaksi == ""] <- NA
> anyNA(transaksi)
[1] TRUE
> is.na(transaksi)%>% colSums()
id idTrans namaBarang idCust namaCust
0 0 0 385 385
```

Gbr.6 Hasil Mengecek Nilai *Missing Value*

Hasil dari kode program pada Gbr.6, pada file transaksi.txt ditemukan 385 data kosong pada kolom idCust dan namaCust. Agar tidak terjadi kesalahan perhitungan dalam penerapana metode *customer segmentation* maupun *market basket analysis*, field yang mengandung data kosong / *missing value* dapat dihilangkan. Berikut pada Gbr.7 kode program untuk menghilangkan data yang terdapat *missing value* dan pada Gbr.8 hasil dari kode tersebut:

```
trans <- na.omit(transaksi)
trans
```

Gbr.7 Kode Program Menghilangkan Data yang Terdapat Nilai *Missing Value*

```
> trans <- na.omit(transaksi)
> trans
id idTrans namaBarang idCust namaCust
1 1 2207160001 MIE BERUANG KI WITO
2 1 2207160001 MILO KI WITO
3 2 2207160002 DESAKU KUNYIT ERW423 ERWIN
4 2 2207160002 DESAKU TUMBAR ERW423 ERWIN
5 2 2207160002 KRATINGDAENG ERW423 ERWIN
6 2 2207160002 NABATI ERW423 ERWIN
7 2 2207160002 NUTRISARI ERW423 ERWIN
8 2 2207160002 PADI MAS ERW423 ERWIN
9 3 2207160003 KAKI TIGA C236 WATIK
10 4 2207160004 GULA C038 DIDIK
11 5 2207160005 TOP COFFE C623 ARIF
12 6 2207160006 SOKLIN C502 ANA
13 7 2207160007 MG FORTUNE A23 AMI
14 7 2207160007 NUVO A23 AMI
15 8 2207160008 BEAR BRAND C5780 ADAM
16 9 2207160009 NUTRISARI C7412 LUTFI
17 10 2207160010 SUSU BENDERA HUD987 HUDA
18 10 2207160010 SUSU BENDERA HUD987 HUDA
19 10 2207160010 TOP COFFE HUD987 HUDA
```

```
182 45 2207160045 GAM H345 HALIFAH
183 45 2207160045 GIV H345 HALIFAH
184 45 2207160045 ROSE BERAS H345 HALIFAH
185 45 2207160045 SUNLIGHT H345 HALIFAH
186 45 2207160045 T SEGITIGA H345 HALIFAH
187 46 2207160046 BD BERAS C685 PARTI
188 46 2207160046 GULA C685 PARTI
189 46 2207160046 MICIN SASA C685 PARTI
190 46 2207160046 NABATI C685 PARTI
191 46 2207160046 SUNLIGHT C685 PARTI
192 47 2207160047 MG FORTUNE H5266 IRWAN
193 48 2207160048 CAKRA 1KG C254 AGUS
194 48 2207160048 GAM C254 AGUS
195 48 2207160048 KRATINGDAENG C254 AGUS
196 48 2207160048 MIE WALIWIS C254 AGUS
197 48 2207160048 MWON C254 AGUS
198 48 2207160048 NUVO C254 AGUS
199 49 2207160049 GULA C467 SUNDARI
200 49 2207160049 ROSE BERAS C467 SUNDARI
[ reached 'max' / getoption("max.print") -- omitted 21774 rows ]
```

Gbr.8 Data Setelah Melakukan Proses *Cleaning*

Pada proses tersebut terdapat 385 baris dari data transaksi.txt yang dihilangkan, dari data sebelumnya 22.359 data dan sekarang menjadi 21.774 data yang akan diolah menggunakan metode *market basket analysis*.

Tahapan selanjutnya yaitu *upload* data dari data yang sudah dibersihkan kedalam file transaksi91.txt dengan kode program sebagai berikut:

```
write.table(trans,
"C:\\skripsi\\transaksi91.txt", fileEncoding =
"UTF-8", sep="\t", quote = FALSE)
```

Gbr.9 Kode Program *Download* Data Transaksi

File yang sudah di-*cleaning* siap digunakan pada proses selanjutnya.

B. Menerapkan Metode *Customer Segmentation* Dengan Bahasa Pemrograman R Pada *Software* R Studio

Setelah melalui tahap *pre-processing*, selanjutnya data dapat diolah menggunakan metode *customer segmentation* dengan algoritme k-means bahasa pemrograman R. Data-data yang digunakan dalam perhitungan metode *customer segmentation* yaitu jenis kelamin, usia, dan total nilai belanja pada file pelanggan.txt. Berikut tahapan untuk metode *customer segmentation* di *software* R Studio:

1) Mengkonversikan data kategorik menjadi data numerik

Pada *field* jenis kelamin isi data masih berupa data kategorik (laki-laki dan perempuan), untuk menerapkan metode *customer segmentation* seluruh data yang diperlukan harus berupa data numerik. Berikut kode program untuk merubah data kategorik menjadi data numerik menggunakan bahasa pemrograman R:

```
pelanggan <-
read.csv("C:\\skripsi\\pelanggan.txt",
sep="\t")

pelanggan_matrix <-
data.matrix(pelanggan[c("jk")])
pelanggan <- data.frame(pelanggan,
pelanggan_matrix)
```

```
head(pelanggan)
```

Gbr.10 Kode Program Mengkonversikan Data Kategorik

```
> pelanggan_matrix <- data.matrix(pelanggan[c("jk")])
> pelanggan <- data.frame(pelanggan, pelanggan_matrix)
> head(pelanggan)
kodecust namacust jk usia tnilai jk.1
1 C293 BUDI L 24 1650 1
2 C2659 WATIK P 31 2000 2
3 C3684 KOMARI L 49 2175 1
4 A15 ANTOK L 37 2575 1
5 F5869 ANDRI L 35 2675 1
6 C475 SUTAR L 61 4150 1
```

Gbr.11 Hasil Mengkonversikan Data Kategorik Ke Bentuk Data Numerik

Pada Gbr.11 terdapat *field* jk dan jk.1, *field* jk berupa data kategorik dan *field* jk.1 yaitu data numerik hasil dari *field* jk yang telah dikonversikan.

2) Mensinkronisasi jenis kelamin dan menormalisasi nilai belanja

Langkah berikutnya yaitu mensinkronisasi jenis kelamin, menormalisasi nilai belanja, mempersiapkan data yang diperlukan dan melakukan *feature scaling* pada data pelanggan yang akan digunakan. *Feature scaling* dilakukan agar tidak terjadi hasil perhitungan jarak yang sangat jauh antara jenis kelamin, usia, dan nilai belanja. Berikut Rumus dari *feature scaling*, yang berupa rumus *standardization*:

$$X_{stand} = \frac{x - \text{mean}(x)}{\text{standard deviation}(x)} \quad [6]$$

Keterangan:

X_{stand} = Nilai *standard*

x = Banyak data

$\text{mean}(x)$ = Nilai rata-rata data x

$\text{standard deviation}(x)$ = Nilai *standard deviation* data x

Berikut pada Gbr.12 dan Gbr.13 kode program dan hasil dari proses tersebut:

```
Jenis.Kelamin <-
unique(pelanggan[c("jk", "jk.1")])
pelanggan$tnilai <- pelanggan$tnilai/1
field_yang_digunakan =
pelanggan[,c("jk.1", "usia", "tnilai")]
pelanggan_standardize <-
scale(field_yang_digunakan)
apply(pelanggan_standardize, 2, mean)
apply(pelanggan_standardize, 2, sd)

head(field_yang_digunakan)
```

Gbr.12 Kode Program *Field* Yang Akan Digunakan Pada Proses *Customer Segmentation*

```
> Jenis.Kelamin <- unique(peleanggan[c("jk","jk.1")])
> pelanggan$tnilai <- pelanggan$tnilai/1
> Field_yang_digunakan = pelanggan[,c("jk.1", "usia", "tnilai")]
> pelanggan_standardize <- scale(field_yang_digunakan)
> apply(pelanggan_standardize,2,mean)
      jk.1      usia      tnilai
1.642525e-16 2.103123e-16 1.025579e-17
> apply(pelanggan_standardize,2,sd)
      jk.1      usia      tnilai
1.1089205 0.9880954 0.6213828
> head(pelanggan_standardize)
      jk.1      usia      tnilai
[1,] -1.1089205 -1.3124822 -0.6226053
[2,]  0.9008716 -0.8772378 -0.6224341
[3,] -1.1089205  0.2419621 -0.6223486
[4,] -1.1089205 -0.5041711 -0.6221530
[5,] -1.1089205 -0.6285267 -0.6221041
[6,] -1.1089205  0.9880954 -0.6213828
```

Gbr.13 Hasil Kode Program Dari Field Yang Akan Digunakan Yang Telah Di Feature Scaling

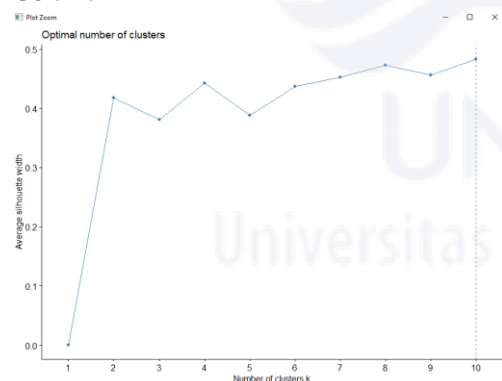
3) Menentukan jumlah cluster optimal

Dalam bahasa pemrograman R terdapat beberapa metode untuk menentukan *cluster* yang optimal, berikut 3 metode diantaranya untuk menentukan *cluster* yang optimal yaitu metode *silhouette coefficient*, *elbow/wss*, dan *gap statistic*. Berikut pada Gbr.14 kode program untuk menentukan jumlah *cluster* yang optimal menggunakan bahasa pemrograman R:

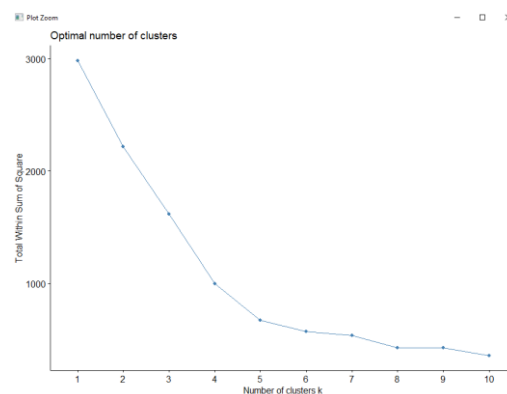
```
fviz_nbclust(pelanggan_standardize,FUNcluster
= kmeans,method = "silhouette")
fviz_nbclust(pelanggan_standardize,FUNcluster
= kmeans,method = "wss")
fviz_nbclust(pelanggan_standardize,FUNcluster
= kmeans,method = "gap_stat")
```

Gbr.14 Kode Program Menentukan Cluster Optimal Menggunakan 3 Metode

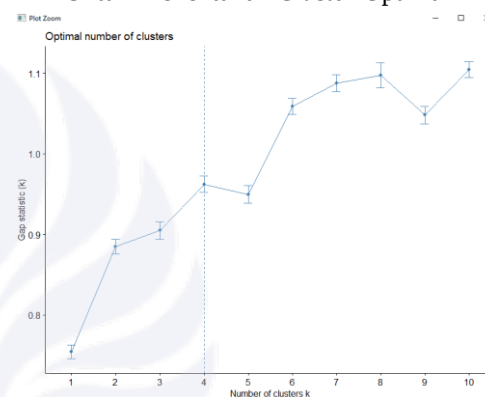
Hasil dari kode program pada Gbr.14 akan ditampilkan berupa gambar grafik pada Gbr.15, Gbr.16, dan Gbr.17.



Gbr.15 Hasil Diagram Grafik Metode Silhouette Untuk Menentukan Cluster Optimal



Gbr.16 Hasil Diagram Grafik Metode Elbow/WSS Untuk Menentukan Cluster Optimal



Gbr.17 Hasil Diagram Grafik Metode Gap Statistic Untuk Menentukan Cluster Optimal

Pada Gbr.15 menggunakan metode *silhouette* jumlah *cluster* yang optimal yaitu 10 *cluster*, Pada Gbr.16 menggunakan metode *elbow/wss* jumlah *cluster* yang optimal berdasarkan sudut siku yang dihasilkan mempertimbangkan nilai SSE yang rendah dan jumlah *cluster* yang tidak terlalu banyak yaitu 5 *cluster*, sedangkan Pada Gbr.17 menggunakan metode *gap statistic* jumlah *cluster* yang optimal yaitu 4 *cluster*.

Menurut hasil perhitungan dari ketiga metode tersebut dalam menentukan jumlah *cluster* yang optimal, maka pada penelitian ini akan menggunakan perhitungan 5 *cluster* sesuai hasil perhitungan menggunakan metode *elbow/wss*. Hal ini diputuskan berdasarkan hasil nilai median antara ketiga metode tersebut, yaitu 4, 5, dan 10.

4) Menjalankan algoritme k-means

Setelah menentukan *cluster* yang optimal, proses selanjutnya yaitu menjalankan metode customer segmentation menggunakan algoritme k-means pada bahasa pemrograman R dengan jumlah *cluster* yang akan dibuat yaitu 5 *cluster*.

Berikut pada Gbr.18 kode program algoritme k-means:

```
kmeans_aloha <-
eclust(pelanggan_standardize, stand =
TRUE, FUNcluster = "kmeans", k=5, graph = F)
kmeans_aloha$cluster
pelanggan$cluster <- kmeans_aloha$cluster
pelanggan
fviz_cluster(kmeans_aloha)
```

Gbr.18

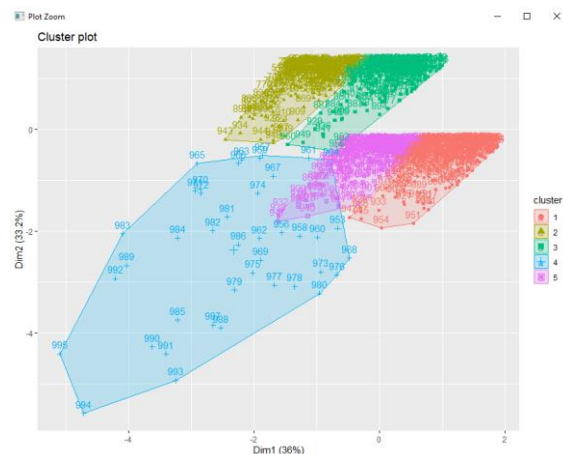
Hasil dari kode program pada Gbr.18 yaitu akan ditampilkan pada Gbr.19, Gbr.20, dan Gbr.21 berikut:

```
> kmeans_aloha <- eclust(pelanggan_standardize, stand = TRUE, FUNcluster = "kmeans", k=5, graph = F)
> kmeans_aloha$cluster
[1] 3 1 2 3 3 2 5 3 3 3 2 2 1 2 2 5 3 1 1 5 2 2 5 1 5 1 2 1 2 2 5 5 1 1 5 5 5 1 1 3 5 5 1 1 2 1 1
[48] 1 1 2 3 1 5 1 5 1 3 2 3 1 3 3 3 5 1 5 1 5 5 5 3 3 2 5 1 2 1 1 1 1 3 2 5 5 1 1 5 1 1 3 1 1 2
[95] 1 1 1 3 1 2 3 2 3 1 5 1 1 1 3 2 2 1 1 3 1 3 1 5 1 1 3 3 5 1 2 5 5 1 1 5 2 2
[142] 5 5 5 3 2 3 3 5 5 3 3 3 2 2 2 1 5 5 3 5 5 1 1 3 3 1 1 1 2 1 5 2 3 5 5 2 2 1 5 3
[189] 1 1 2 3 2 5 2 1 2 2 3 1 3 5 1 3 5 5 5 5 2 3 5 1 5 3 3 3 5 1 5 1 1 3 2 2 3 1
[236] 3 1 1 2 1 3 1 2 3 1 5 1 5 3 1 1 1 1 3 3 1 2 1 5 1 2 1 1 1 1 1 1 1 2 2 3
[283] 5 5 5 5 2 5 2 2 1 1 2 2 2 2 1 1 2 2 1 5 1 3 3 3 1 2 2 1 1 3 5 5 2 2 2 3 5 5 1 5
[330] 5 5 1 5 5 5 5 5 1 1 3 3 3 5 5 2 2 1 5 5 1 5 1 3 5 2 1 1 3 2 3 1 1 2 1 3 5 5 1
[377] 1 1 3 1 3 1 5 1 1 3 3 1 2 5 5 5 2 1 1 3 1 3 1 1 5 4 1 1 1 5 2 2 1 1
[424] 3 1 2 5 3 2 2 2 2 1 1 2 5 5 3 3 5 1 3 2 2 1 1 5 5 5 1 3 1 1 2 5 2 3 3 1 1 1
[471] 5 5 2 1 1 2 1 2 1 1 5 2 3 1 1 5 2 2 2 5 5 5 5 1 5 3 3 1 1 5 2 3 5 2 2 2 2
[518] 5 5 2 1 2 5 1 1 2 1 1 2 1 1 1 2 5 5 5 1 1 1 2 2 1 1 2 2 1 1 2 2 1 1 1 3 5 2
[565] 5 5 2 1 2 5 5 5 1 2 2 1 2 5 2 2 1 5 5 5 1 1 1 2 2 5 5 1 2 2 1 1 1 1 1 3 5 2
[612] 2 2 3 3 1 2 5 1 1 5 2 3 2 5 1 1 3 1 3 3 3 3 5 5 5 1 5 5 5 5 2 1 2 5 1 1
[659] 2 2 3 3 2 2 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[706] 2 2 3 3 2 2 1 1 5 3 3 1 2 5 5 5 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[753] 3 1 1 2 1 2 5 5 5 5 1 2 5 5 2 1 1 3 1 1 3 3 5 1 5 5 3 3 1 1 1 5 2 5 5 1 5 3 2 3
[800] 2 5 5 5 1 1 1 5 1 1 2 5 2 5 5 5 5 4 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[847] 3 3 5 2 2 1 1 5 5 1 3 3 5 5 3 5 5 1 1 2 2 1 3 1 1 1 1 3 3 3 1 1 5 5 1 1 5 5
[894] 3 5 2 2 2 5 5 1 1 2 5 5 3 2 2 5 5 5 5 5 1 1 1 5 5 1 5 2 5 5 5 1 1 1 5 1 2 3 2 2 5
[941] 3 1 2 2 1 2 5 2 3 1 1 3 4 1 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
[988] 4 4 4 4 4 4 4
```

Gbr.19 Hasil Cluster Algoritme K-means

```
> pelanggan$cluster <- kmeans_aloha$cluster
> pelanggan
  kodecust  namacust  jk  usia  tn1lai  jk.1 cluster
1      C293      BUDI  L   24   1650    1    3
2      C2659     WATIK  P   31   2000    2    1
3      C3684     KOMARI L   49   2175    1    2
4       A15      ANOK  L   37   2575    1    3
5      F5869     ANDRI  L   35   2675    1    3
6      C475      SUTAR  L   61   4150    1    2
7      C083      JARWATI P   65   4250    2    5
8      C543      KASENAN L   19   5950    1    3
9      ANG324     ANGEL  P   49   6000    2    5
10     AKB789     AKBAR  L   39   7000    1    3
11     C535      EKO    L   55   7680    1    2
12     EGEG      RONI   L   61   8395    1    2
13     C583      MIATUN  P   45   8960    2    1
14     A005      ABDUL  L   52   11700   1    2
15     C5556     ANSOR  L   73   12150   1    2
16     C068      HASNA  P   64   13235   2    5
17     L5241     JONI   L   24   13235    1    3
18     J21498     IKAH   P   40   13800    2    1
19     C083      SUTINI  P   26   14475    2    1
20     REW68     ERNAWATI P   61   14500    2    5
130     J4586     RIKA   P   53   83170   2    5
131     C318      IMA    P   20   85285   2    1
132     C088      JUARI  L   35   87000    1    3
133     O1LL      NIZAM  L   61   87200    1    2
134     C578      SALIMI  P   62   87735   2    5
135     C1546     RUSMINI P   73   87750   2    5
136     SK865     SUKINI  P   67   88525   2    5
137     C5684     MARIYAM P   34   88825   2    1
138     C256      ANING  P   44   89550   2    1
139     K2563     ENI    P   66   91300   2    5
140     C5601     BENI   L   51   91470    1    2
141     HY56      TOIRIN  L   57   92205    1    2
142     M234      MARTIAH P   47   92350    2    5
[reached 'max' / getopton("max.print") -- omitted 853 rows]
```

Gbr.20 Data Pelanggan setelah di cluster
Menggunakan Algoritme K-Means



Gbr.21 Grafik Hasil Pengelompokan 5 Cluster
Menggunakan Algoritme K-Means

Dari perhitungan data diatas dapat disimpulkan bahwa dari 995 data transaksi yang dianalisa menggunakan metode *elbow/wss* menghasilkan kelompok cluster pelanggan yang memiliki kesamaan perilaku sebanyak 5 cluster, dimana anggota cluster pertama sebanyak 217 cluster, anggota cluster kedua sebanyak 209 cluster, anggota cluster ketiga sebanyak 295 cluster, anggota cluster keempat sebanyak 233 cluster, dan anggota cluster kelima sebanyak 41 cluster. Berikut penjelasan dari hasil cluster tersebut:

- Cluster pertama yaitu didominasi oleh perempuan dengan rentang usia ± 30 tahun dan memiliki rata-rata total belanja sebesar ± 937 ribu rupiah
- Cluster kedua yaitu didominasi oleh laki-laki dengan rentang usia ± 60 tahun dan memiliki rata-rata total belanja sebesar ± 975 ribu rupiah
- Cluster ketiga yaitu didominasi oleh laki-laki dengan rentang usia ± 30 tahun dan memiliki rata-rata total belanja sebesar ± 876 ribu rupiah
- Cluster keempat yaitu campuran baik laki-laki maupun perempuan dengan rentang usia ± 50 tahun dan memiliki rata-rata total belanja sebesar $\pm 9,12$ juta rupiah
- Cluster kelima yaitu didominasi oleh perempuan dengan rentang usia ± 60 tahun dan memiliki rata-rata total belanja sebesar ± 962 ribu rupiah

C. Menerapkan Metode Market Basket Analysis Dengan Bahasa Pemrograman R Pada Software R Studio

Dari hasil proses *customer segmentation* menghasilkan 5 kelompok/cluster. Setelah data pelanggan dibagi menjadi beberapa cluster, selanjutnya data untuk tiap-tiap cluster akan dianalisa menggunakan metode *market basket analysis*. Berikut pada Gbr.22 kode program untuk

menghitung nilai *market basket analysis* pada anggota *cluster* pertama menggunakan algoritme eclat dengan bahasa pemrograman R pada *software* RStudio dan nilai *minimal support* sebesar 3,5% dan nilai *minimal confidence* sebesar 10%:

```
library(arules)
library(arulesViz)
transaksi <-
read.transactions(file="C:\\skripsi\\cluster1.txt",
format="single", sep="\t", cols=c(1,2),
skip=1)

freq.items <- eclat(transaksi,
parameter=list(supp=0.035, minlen=2, maxlen=10))
inspect(sort(freq.items, by = "support"))
freq.rules <- ruleInduction(freq.items,
transaksi, confidence=0.1)
inspect(sort(freq.rules, by = "confidence"))

plot(subset(freq.items), method="graph")
```

Gbr.22 Kode Program Algoritme Eclat Dengan Bahasa Pemrograman R Cluster Pertama

Berikut pada Gbr.23 dan Gbr.24 untuk hasil dari kode program diatas:

```
R 4.12.1 ~/>
> freq.items <- eclat(transaksi, parameter=list(supp=0.035, minlen=2, maxlen=10))
eclat

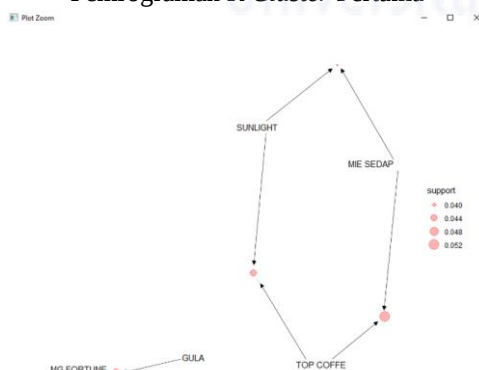
parameter specification:
tidlists support minlen maxlen target ext
FALSE 0.035 2 10 frequent itemsets TRUE

algorithmic control:
sparse sort verbose
7 -2 TRUE

Absolute minimum support count: 24

create itemset ...
set transactions ... [219 item(s), 1276 transaction(s)] done [0.00s].
sorting and recoding items ... [32 item(s)] done [0.00s].
creating sparse bit matrix ... [32 row(s), 1276 column(s)] done [0.00s].
writing ... [4 set(s)] done [0.00s].
Creating S4 object ... done [0.00s].
> inspect(sort(freq.items, by = "support"))
items support count
[1] {MIE SEDAP, TOP COFFE} 0.05250784 67
[2] {SUNLIGHT, TOP COFFE} 0.04388715 56
[3] {GULA, MG FORTUNE} 0.04153603 53
[4] {MIE SEDAP, SUNLIGHT} 0.03918495 50
> freq.rules <- ruleInduction(freq.items, transaksi, confidence=0.1)
> inspect(sort(freq.rules, by = "confidence"))
lhs rhs support confidence lift itemset
[1] {SUNLIGHT} => {TOP COFFE} 0.04388715 0.3835616 2.353003 3
[2] {SUNLIGHT} => {MIE SEDAP} 0.03918495 0.3424658 1.883562 2
[3] {TOP COFFE} => {MIE SEDAP} 0.05250784 0.3221154 1.771635 4
[4] {MIE SEDAP} => {TOP COFFE} 0.05250784 0.2887931 1.771635 4
[5] {TOP COFFE} => {SUNLIGHT} 0.04388715 0.2692308 2.353003 3
[6] {GULA} => {MG FORTUNE} 0.04153603 0.2465116 1.182514 1
[7] {MIE SEDAP} => {SUNLIGHT} 0.03918495 0.2155172 1.883562 2
[8] {MG FORTUNE} => {GULA} 0.04153603 0.1992481 1.182514 1
```

Gbr.23 Hasil Algoritme Eclat Menggunakan Bahasa Pemrograman R Cluster Pertama



Gbr.24 Hasil Grafik Algoritme Eclat Pada Data Transaksi Cluster Pertama

Dari hasil perhitungan data transaksi menggunakan algoritme eclat pada data transaksi cluster pertama menghasilkan informasi bahwa dari 217 pelanggan terdapat 1.276 transaksi dengan nilai *minimal support* sebesar 3.5% dan nilai *minimal confidence* 10% menghasilkan 8 aturan asosiasi yang kuat.

Selanjutnya mengolah data transaksi *cluster* kedua, untuk mengolah data transaksi *cluster* kedua, *cluster* ketiga, *cluster* keempat dan *cluster* kelima menggunakan metode *market basket analysis* dengan algoritme eclat, kode program pada *software* RStudio sama seperti pada gambar Gbr.22 hanya perlu mengganti nama file yang akan dieksekusi saja, misalns cluster2.txt untuk anggota transaksi pelanggan *cluster* kedua. Berikut untuk hasil eksekusi kode program *cluster* kedua, *cluster* ketiga, *cluster* keempat, dan *cluster* kelima:

```
> freq.items <- eclat(transaksi, parameter=list(supp=0.035, minlen=2, maxlen=10))
eclat

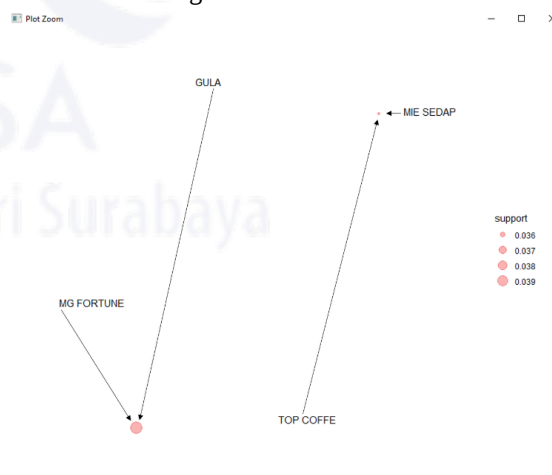
parameter specification:
tidlists support minlen maxlen target ext
FALSE 0.035 2 10 frequent itemsets TRUE

algorithmic control:
sparse sort verbose
7 -2 TRUE

Absolute minimum support count: 24

create itemset ...
set transactions ... [198 item(s), 701 transaction(s)] done [0.00s].
sorting and recoding items ... [29 item(s)] done [0.00s].
creating sparse bit matrix ... [29 row(s), 701 column(s)] done [0.00s].
writing ... [2 set(s)] done [0.00s].
Creating S4 object ... done [0.00s].
> inspect(sort(freq.items, by = "support"))
items support count
[1] {GULA, MG FORTUNE} 0.03994294 28
[2] {MIE SEDAP, TOP COFFE} 0.03566334 25
> freq.rules <- ruleInduction(freq.items, transaksi, confidence=0.1)
> inspect(sort(freq.rules, by = "confidence"))
lhs rhs support confidence lift itemset
[1] {MIE SEDAP} => {TOP COFFE} 0.03566334 0.2976190 2.398057 2
[2] {TOP COFFE} => {MIE SEDAP} 0.03566334 0.2873563 2.398057 2
[3] {GULA} => {MG FORTUNE} 0.03994294 0.2137405 1.012379 1
[4] {MG FORTUNE} => {GULA} 0.03994294 0.1891892 1.012379 1
```

Gbr.25 Hasil Algoritme Eclat Menggunakan Bahasa Pemrograman R Cluster Kedua



Gbr.26 Hasil Grafik Algoritme Eclat Pada Data Transaksi Cluster Kedua

Dari hasil perhitungan data transaksi menggunakan algoritme eclat pada data transaksi *cluster* kedua menghasilkan informasi bahwa dari 209 pelanggan

terdapat 701 transaksi dengan nilai *minimal support* sebesar 3.5% dan nilai *minimal confidence* 10% menghasilkan 4 aturan asosiasi yang kuat.

```
> freq.items <- eclat(transaksi, parameter=list(supp=0.035, minlen=2, maxlen=10))
Eclat

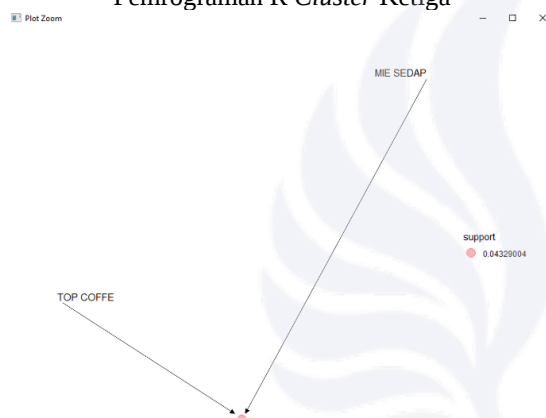
parameter specification:
tidlists support minlen maxlen target ext
FALSE 0.035 2 10 frequent itemsets TRUE

algorithmic control:
sparse sort verbose
7 -2 TRUE

Absolute minimum support count: 24

create itemset ...
set transactions ... [196 item(s), 693 transaction(s)] done [0.00s].
sorting and recoding items ... [29 item(s)] done [0.00s].
creating sparse bit matrix ... [29 row(s), 693 column(s)] done [0.00s].
writing ... [1 set(s)] done [0.00s].
creating s4 object ... done [0.00s].
> inspect(sort(freq.items, by = "support"))
items support count
[1] {MIE SEDAP, TOP COFFE} 0.04329004 30
> freq.rules <- ruleInduction(freq.items, transaksi, confidence=0.1)
> inspect(sort(freq.rules, by = "confidence"))
lhs rhs support confidence lift itemset
[1] {MIE SEDAP} => {TOP COFFE} 0.04329004 0.3797468 2.860484 1
[2] {TOP COFFE} => {MIE SEDAP} 0.04329004 0.3260870 2.860484 1
```

Gbr.27 Hasil Algoritme Eclat Menggunakan Bahasa Pemrograman R Cluster Ketiga



Gbr.28 Hasil Grafik Algoritme Eclat Pada Data Transaksi Cluster Ketiga

Dari hasil perhitungan data transaksi menggunakan algoritme eclat pada data transaksi *cluster* ketiga menghasilkan informasi bahwa dari 295 pelanggan terdapat 693 transaksi dengan nilai *minimal support* sebesar 3.5% dan nilai *minimal confidence* 10% menghasilkan 2 aturan asosiasi yang kuat.

```
> freq.items <- eclat(transaksi, parameter=list(supp=0.035, minlen=2, maxlen=10))
Eclat

parameter specification:
tidlists support minlen maxlen target ext
FALSE 0.035 2 10 frequent itemsets TRUE

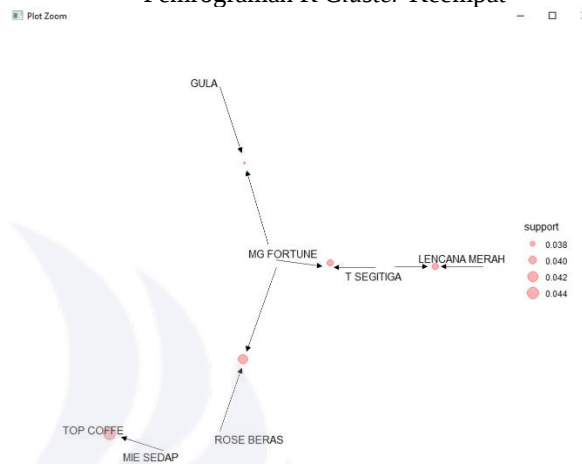
algorithmic control:
sparse sort verbose
7 -2 TRUE

Absolute minimum support count: 25

create itemset ...
set transactions ... [203 item(s), 726 transaction(s)] done [0.00s].
sorting and recoding items ... [36 item(s)] done [0.00s].
creating sparse bit matrix ... [36 row(s), 726 column(s)] done [0.00s].
writing ... [5 set(s)] done [0.00s].
creating s4 object ... done [0.00s].
> inspect(sort(freq.items, by = "support"))
items support count
[1] {MIE SEDAP, TOP COFFE} 0.04545455 33
[2] {MG FORTUNE, ROSE BERAS} 0.04132231 30
[3] {LENCANA MERAH, T SEGITIGA} 0.03856749 28
[4] {MG FORTUNE, T SEGITIGA} 0.03856749 28
[5] {GULA, MG FORTUNE} 0.03719008 27
```

```
> freq.rules <- ruleInduction(freq.items, transaksi, confidence=0.1)
> inspect(sort(freq.rules, by = "confidence"))
lhs rhs support confidence lift itemset
[1] {LENCANA MERAH} => {T SEGITIGA} 0.03856749 0.4057971 3.8260870 1
[2] {T SEGITIGA} => {LENCANA MERAH} 0.03856749 0.3636364 3.8260870 1
[3] {T SEGITIGA} => {MG FORTUNE} 0.03856749 0.3636364 1.3821990 2
[4] {TOP COFFE} => {MIE SEDAP} 0.04545455 0.3402062 2.4948454 5
[5] {MIE SEDAP} => {TOP COFFE} 0.04545455 0.3333333 2.4948454 5
[6] {ROSE BERAS} => {MG FORTUNE} 0.04132231 0.3260870 1.2394719 3
[7] {GULA} => {MG FORTUNE} 0.03719008 0.2213115 0.8412153 4
[8] {MG FORTUNE} => {ROSE BERAS} 0.04132231 0.1570681 1.2394719 3
[9] {MG FORTUNE} => {T SEGITIGA} 0.03856749 0.1465969 1.3821990 2
[10] {MG FORTUNE} => {GULA} 0.03719008 0.1413613 0.8412153 4
```

Gbr.29 Hasil Algoritme Eclat Menggunakan Bahasa Pemrograman R Cluster Keempat



Gbr.30 Hasil Grafik Algoritme Eclat Pada Data Transaksi Cluster Keempat

Dari hasil perhitungan data transaksi menggunakan algoritme eclat pada data transaksi *cluster* keempat menghasilkan informasi bahwa dari 233 pelanggan terdapat 726 transaksi dengan nilai *minimal support* sebesar 3.5% dan nilai *minimal confidence* 10% menghasilkan 10 aturan asosiasi yang kuat.

```
> freq.items <- eclat(transaksi, parameter=list(supp=0.035, minlen=2, maxlen=10))
Eclat

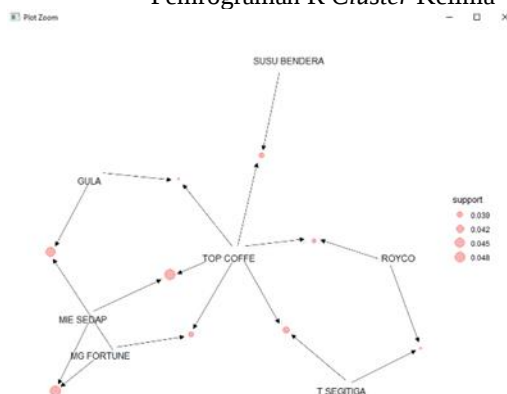
parameter specification:
tidlists support minlen maxlen target ext
FALSE 0.035 2 10 frequent itemsets TRUE

algorithmic control:
sparse sort verbose
7 -2 TRUE

Absolute minimum support count: 35

create itemset ...
set transactions ... [210 item(s), 1022 transaction(s)] done [0.00s].
sorting and recoding items ... [33 item(s)] done [0.00s].
creating sparse bit matrix ... [33 row(s), 1022 column(s)] done [0.00s].
writing ... [9 set(s)] done [0.00s].
creating s4 object ... done [0.00s].
> inspect(sort(freq.items, by = "support"))
items support count
[1] {MIE SEDAP, TOP COFFE} 0.04892368 50
[2] {MG FORTUNE, MIE SEDAP} 0.04892368 50
[3] {GULA, MG FORTUNE} 0.04598826 47
[4] {T SEGITIGA, TOP COFFE} 0.03913894 40
[5] {SUSU BENDERA, TOP COFFE} 0.03816047 39
[6] {MG FORTUNE, TOP COFFE} 0.03816047 39
[7] {ROYCO, TOP COFFE} 0.03718200 38
[8] {ROYCO, T SEGITIGA} 0.03620352 37
[9] {GULA, TOP COFFE} 0.03620352 37
> freq.rules <- ruleInduction(freq.items, transaksi, confidence=0.1)
> inspect(sort(freq.rules, by = "confidence"))
lhs rhs support confidence lift itemset
[1] {SUSU BENDERA} => {TOP COFFE} 0.03816047 0.4239130 2.578804 1
[2] {T SEGITIGA} => {TOP COFFE} 0.03913894 0.3603604 2.192192 2
[3] {T SEGITIGA} => {ROYCO} 0.03620352 0.3333333 2.682415 3
[4] {MIE SEDAP} => {TOP COFFE} 0.04892368 0.3048780 1.854675 7
[5] {MIE SEDAP} => {MG FORTUNE} 0.04892368 0.3048780 1.372623 8
[6] {ROYCO} => {TOP COFFE} 0.03718200 0.2992126 1.820210 6
[7] {TOP COFFE} => {MIE SEDAP} 0.04892368 0.2976190 1.854675 7
[8] {ROYCO} => {T SEGITIGA} 0.03620352 0.2913386 2.682415 3
[9] {GULA} => {MG FORTUNE} 0.04598826 0.2435233 1.096391 5
[10] {TOP COFFE} => {T SEGITIGA} 0.03913894 0.2380952 2.192192 2
[11] {TOP COFFE} => {SUSU BENDERA} 0.03816047 0.2321429 2.578804 1
[12] {TOP COFFE} => {MG FORTUNE} 0.03816047 0.2321429 1.045154 9
[13] {TOP COFFE} => {ROYCO} 0.03718200 0.2261905 1.820210 6
[14] {MG FORTUNE} => {MIE SEDAP} 0.04892368 0.2202643 1.372623 8
[15] {TOP COFFE} => {GULA} 0.03620352 0.2202381 1.166235 4
[16] {MG FORTUNE} => {GULA} 0.04598826 0.2070485 1.096391 5
[17] {GULA} => {TOP COFFE} 0.03620352 0.1917098 1.166235 4
[18] {MG FORTUNE} => {TOP COFFE} 0.03816047 0.1718062 1.045154 9
```

Gbr.31 Hasil Algoritme Eclat Menggunakan Bahasa Pemrograman R Cluster Kelima



Gbr.32 Hasil Grafik Algoritme Eclat Pada Data Transaksi Cluster Kelima

Dari hasil perhitungan data transaksi menggunakan algoritme eclat pada data transaksi *cluster* kelima menghasilkan informasi bahwa dari 41 pelanggan terdapat 1.022 transaksi dengan nilai *minimal support* sebesar 3.5% dan nilai *minimal confidence* 10% mengasilkan 18 aturan asosiasi yang kuat.

IV. KESIMPULAN

. Dari hasil analisa penelitian ini dapat diambil beberapa kesimpulan yaitu, sebagai berikut:

- A. Data hasil transaksi swalayan CV. Aloha Ponorogo diolah menggunakan metode *customer segmentation* dengan algoritme k-means dan metode *market basket analysis* dengan algoritme eclat yaitu langkah pertama mengelompokkan pelanggan sesuai kesamaan sifat maupun karakteristiknya menggunakan algoritme k-means terlebih dahulu. Menurut perhitungan metode *elbow* jumlah *cluster* pelanggan yang optimal dalam penelitian ini yaitu berjumlah 5 *cluster*. Selanjutnya, dari kelima *cluster* pelanggan tersebut untuk tiap-tiap *cluster* akan dianalisa menggunakan metode *market basket analysis* dengan algoritme eclat. Dari hasil analisa akan dipeoleh aturan-aturan asosiasi yang kuat untuk tiap-tiap *cluster*-nya. Dan output dari penelitian ini yaitu hasil rekomendasi pemaketan atau penempatan produk dengan tujuan untuk menambah penjualan (*up-selling*) dan penjualan silang (*cross-selling*) menurut hasil analisa.
- B. Hasil menganalisa data pelanggan yang diperoleh dari CV. Aloha Ponorogo pada tahun 2022 yang dianalisa menggunakan metode *customer segmentation* dengan algoritme k-means melalui bahasa pemrograman R menghasilkan 5 *cluster* berdasarkan kesamaan sifat dan karakteristik pelanggannya dengan jumlah kelompok

cluster pertama sebanyak 217 pelanggan, kelompok *cluster* kedua sebanyak 209 pelanggan, kelompok *cluster* ketiga sebanyak 295 pelanggan, kelompok *cluster* keempat sebanyak 233 pelanggan, dan kelompok *cluster* kelima sebanyak 41 pelanggan.

- C. Hasil analisa *market basket analysis* menggunakan algoritme eclat dari 5 kelompok / *cluster* pelanggan yang telah dibagi menggunakan metode *customer segmentation* dengan algoritme k-means ditentukan nilai *minimal support* sebesar 3.5% dan nilai *minimal confidence* sebesar 10% dapat disimpulkan bahwa tiap *cluster* menghasilkan aturan asosiasi yang berbeda beda, yaitu:

- 1) Pada kelompok *cluster* pertama, dari 1.276 data transaksi menghasilkan 8 aturan asosiasi
- 2) Pada kelompok *cluster* kedua, dari 701 data transaksi menghasilkan 4 aturan asosiasi
- 3) Pada kelompok *cluster* ketiga, dari 693 data transaksi menghasilkan 2 aturan asosiasi
- 4) Pada kelompok *cluster* keempat, dari 726 data transaksi menghasilkan 10 aturan asosiasi
- 5) Pada kelompok *cluster* kelima, dari 1.022 data transaksi menghasilkan 18 aturan asosiasi

- D. Dari hasil penelitian kedua metode tersebut dapat direkomendasikan hasil penataan produk dengan tujuan untuk menambah penjualan (*up-selling*) dan penjualan silang (*cross-selling*) yaitu sebaagi berikut:

- 1) Pada kelompok *cluster* pertama, dapat menempatkan atau memaketkan produk gula dengan mg fortune secara berdekatan dan mie sedap, sunlight, dengan top coffe secara berdekatan
- 2) Pada kelompok *cluster* kedua, data menempatkan gula dengan mg fortune secara berdekatan dan mie sedap dengan top coffe secara berdekatan
- 3) Pada kelompok *cluster* ketiga, dapat menempatkan mie sedap dengan top coffe secara berdekatan
- 4) Pada kelompok *cluster* keempat, dapat menempatkan gula dengan mie sedap secara berdekatan dan gula, mg fortune, t segitiga, lencana merah, dengan rose beras secara berdekatan
- 5) Pada kelompok *cluster* kelima, dapat menempatkan susu bendera, gula, mg fortune, t segitiga, top coffe, royco, dengan mie sedap secara berdekatan

V. SARAN

Dengan terbatasnya waktu, terdapat masih banyak sekali kekurangan dari penelitian ini. Saran untuk kedepannya semoga pengembang dapat mencoba mengolah data

menggunakan algoritme yang berbeda agar mengetahui perbandingan hasil dengan algoritme yang lain ataupun menerapkan bahasa program R pada penelitian ini kedalam bentuk aplikasi berbahasa pemrograman R.

REFERENSI

- [1] Han, Jiawei. Kamber, Micheline. Pei, Jian. (2012). *Data Mining Concepts and Techniques. Third Edition. USA: Morgan Kaufmann Publishers is an imprint of Elsevier.*J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.
- [2] Al Arisyi, Imam. Nuryana, IKD. Permadi, GS. (2019). *Sistem Informasi Pengelompokan Kelas Diniyah Santi Pada Pondok Pesantren Kreatif AlMuhsinin Dengan Metode K-Means*. Jurnal Ilmiah Inovasi Teknologi Informasi. Vol. 4 Hal. 48-56.
- [3] Haristyarini, Raniar. Yustanti, Wiyli. (2022). *Penerapan Metode Maeket Basket Analysis Dengan Algoritma Eclat dan Predeksi dengan Artifical Neural Neteork pada Data Transaksi Penjualan*. JEISBI (Journal of Emerging Information Systems and Business Intelligence). Vol. 02 No. 03 Hal. 21-29.
- [4] Putri, TBI. Sujatmiko, Bambang. Andriani, Anita. (2022). *Implementasi Algoritma Apriori Untuk Menentukan Strategi Pemasaran*. Jurnal Ilmiah Inovasi Teknologi Informasi. Vol.7 Hal. 41-48.
- [5] Sussolaikah, Kelik. (2021). *Market Basket Analysis for Determination of Consumer Behavior at XYZ Stores Using R Programming*. Advance Sustainable Science, Engineering and Technology (ASSET). Vol. 3 No.2.
- [6] Kukuh, R.. (2019) *Melakukan Feature Scaling pada Dataset*. Diakses 10 Oktober 2023 dari <https://medium.com/@rkukuh/melakukan-feature-scaling-pada-dataset-229531bb08de>

