

## BAB IV

### HASIL DAN PEMBAHASAN

Pada penelitian ini data produk penjualan terdiri dari 18 Unit Series yang ada pada PT. Tunas Dwipa Matra, data diproses dengan membuang data yang tidak konsisten, memperbaiki kesalahan data, kemudian data ditransformasikan untuk mengubah data dari bentuk asalnya kedalam bentuk yang cocok untuk dikelompokkan.

#### 4.1 Pengumpulan Data

Sumber data utama yang digunakan dalam penelitian ini adalah data produk transaksi penjualan pada PT. Dwipa Matra dengan jumlah keseluruhan data ada 7594 dari tahun 2021-2023. Yang mana telah dilakukan *cleaning* data, dan menjadi 7078 data. Data yang diterima dalam bentuk tabel pada *Microsoft excel* sehingga mempermudah untuk dilakukan pembersihan data atau memfilter data, seperti pada gambar 4.1. Data ini di dapatkan melalui perizinan kepada pihak PT. Dwipa Matra.

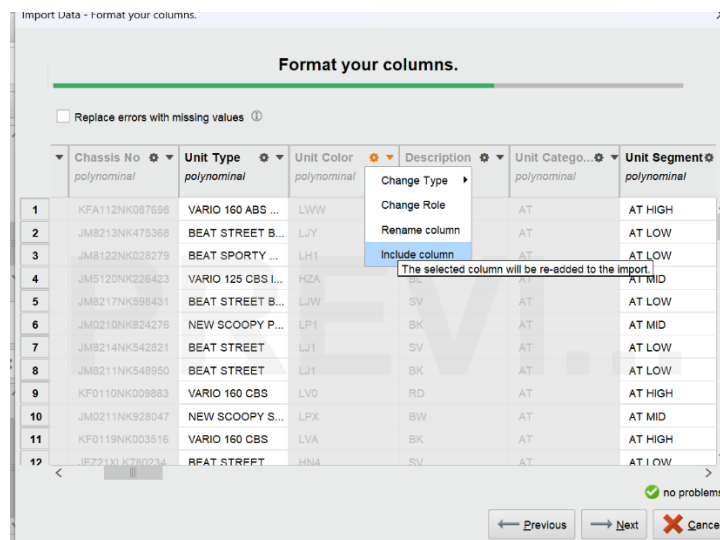
No	Dealer	Unit Type	Unit Color	Description	Unit Category	Unit Segment	Unit Series	Jabatan	Jenis Penjualan	Encry C	Filcoy Name	Nama Customer	No KIP	h
1	PT TOM (Sekampung)	VARIO 160 CBS BLAS	LWH	WH	AT	AT HIGH	VARIO SERIES	Sales Counter	Credit	FC0000000	PT. Summit Oto F	Uswatun Khasanah	180754711920006	18071315163400
2	PT TOM (Sekampung)	BEAT STREET BLASTE	LJY	BK	AT	AT LOW	BEAT STREET SERIES	Salesman	Credit	FC0000000	PT. Summit Oto F	YUNIA PUSPITA SARI	180711489694001	180755960103000
3	PT TOM (Sekampung)	BEAT SPORTY CBS	LXL	PH	AT	AT LOW	BEAT SPORTY SERIES	Sales Counter	Cash	-	-	Barudin	180752629568002	180754664656000
4	PT TOM (Sekampung)	VARIO 125 CBS ISS	H2A	BL	AT	AT MID	VARIO 125 SERIES	Salesman	Cash	-	-	Andrianto	180752929977003	180752929977000
5	PT TOM (Sekampung)	BEAT STREET BLASTE	LJW	SV	AT	AT LOW	BEAT STREET SERIES	Salesman	Credit	FC0000000	PT. Summit Oto F	Sumarno	18075012031750011	180750146311002
6	PT TOM (Sekampung)	NEW SCOOPI PREST	LPI	BK	AT	AT MID	SCOOPI SERIES	Sales Counter	Cash	-	-	Nazrin	180752325678004	1807519616558000
7	PT TOM (Sekampung)	BEAT STREET	L11	SV	AT	AT LOW	BEAT STREET SERIES	Sales Counter	Cash	-	-	Sunari	180750220789001	180754489111000
8	PT TOM (Sekampung)	BEAT STREET	L11	BK	AT	AT LOW	BEAT STREET SERIES	Sales Counter	Cash	-	-	Heri Susanto	180750220789001	180754646464000
9	PT TOM (Sekampung)	VARIO 160 CBS	LVD	RD	AT	AT HIGH	VARIO SERIES	Salesman	Cash	-	-	Siti Aminah	180750481277002	180750481277000
10	PT TOM (Sekampung)	NEW SCOOPI STYLUS	LFX	BW	AT	AT MID	SCOOPI SERIES	Koordinator Sale	Cash	-	-	Kholis Roihanah	1807509451295003	180713484014000
11	PT TOM (Sekampung)	VARIO 160 CBS	LVA	BK	AT	AT HIGH	VARIO SERIES	Sales Counter	Cash	-	-	Sunaryo	1807509394795002	180713484014000
12	PT TOM (Sekampung)	BEAT STREET CBS	HN4	SV	AT	AT LOW	BEAT STREET SERIES	Sales Counter	Cash	-	-	Sumartini	180751029961001	180713484014001
13	PT TOM (Sekampung)	VARIO 125 CBS ISS	SRDA	WH	AT	AT MID	VARIO 125 SERIES	Koordinator Sale	Cash	-	-	Widodo	180751251080002	180769998986794
14	PT TOM (Sekampung)	VARIO 120	HNDD	PH	AT	AT HIGH	VARIO 120 SERIES	Sales Counter	Cash	-	-	Tuyun Ningih	180750491790004	180750752530000
15	PT TOM (Sekampung)	BEAT SPORTY CBS ISS	LXC	BL	AT	AT LOW	BEAT SPORTY SERIES	Sales Counter	Cash	-	-	Nika Waritiana	180752556930003	180752556930000
16	PT TOM (Sekampung)	VARIO 160 CBS	LVI0	RD	AT	AT HIGH	VARIO SERIES	Sales Counter	Cash	-	-	Selvia Dwi Cahyanti	1807504640101003	1807504640101000
17	PT TOM (Sekampung)	VARIO 125 CBS ISS	H2A	BL	AT	AT MID	VARIO 125 SERIES	Team Leader Pa	Credit	FC0000000	PT. Federal Intern	Edi WIBOWO	180751043897001	180751043897000
18	PT TOM (Sekampung)	BEAT STREET BLASTE	LJY	SV	AT	AT LOW	BEAT STREET SERIES	Koordinator Sale	Cash	-	-	Utari Noer Kartika Dewi	18075159111910008	1807510930317000
19	PT TOM (Sekampung)	VARIO 160 CBS	LVD	BK	AT	AT HIGH	VARIO SERIES	Salesman	Credit	FC0000000	PT. Summit Oto F	Filliwanto	18075116110850006	18075116110850000
20	PT TOM (Sekampung)	BEAT SPORTY CBS	LXL	WH	AT	AT LOW	BEAT SPORTY SERIES	Koordinator Sale	Credit	FC000003	PT. Mega Central	FIP PARTI	18075052112900005	1807505960103000
21	PT TOM (Sekampung)	VARIO 160 CBS BLAS	LJW	BK	AT	AT HIGH	VARIO SERIES	Salesman	Credit	FC000003	PT. Mega Central	FIPONBAH	18075045037800200	1807505960103000
22	PT TOM (Sekampung)	VARIO 160 CBS BLAS	LJW	RD	AT	AT HIGH	VARIO SERIES	Salesman	Cash	-	-	Tia	1807504109920002	1807505960103000
23	PT TOM (Sekampung)	VARIO 125 CBS ISS	H2A	WH	AT	AT MID	VARIO 125 SERIES	Salesman	Cash	-	-	Junimo	1807503029920006	1807503029920000
24	PT TOM (Sekampung)	BEAT SPORTY CBS ISS	LXL	LS	AT	AT LOW	BEAT SPORTY SERIES	Salesman	Credit	FC0000000	PT. Adira Dinamika	USWATUN KHASANAH	18052864202950002	1807505960103000
25	PT TOM (Sekampung)	BEAT SPORTY CBS ISS	LXT	BK	AT	AT LOW	BEAT SPORTY SERIES	Sales Counter	Cash	-	-	Yati	18075146238563002	18075146238563000
26	PT TOM (Sekampung)	BEAT STREET	L11	SV	AT	AT LOW	BEAT STREET SERIES	Salesman	Cash	-	-	Sujono	18075096565630011	18075096565630000
27	PT TOM (Sekampung)	BEAT SPORTY CBS	LXL	WH	AT	AT LOW	BEAT SPORTY SERIES	Sales Counter	Cash	-	-	Sugyo Muhruf	18075112404550011	18075112404550000
28	PT TOM (Sekampung)	BEAT STREET BLASTE	LJY	SV	AT	AT LOW	BEAT STREET SERIES	Salesman	Cash	-	-	Dwi Firmansyah	1807520311950011	180769669599000
29	PT TOM (Sekampung)	BEAT SPORTY CBS ISS	LXT	BK	AT	AT LOW	BEAT SPORTY SERIES	Salesman	Credit	FC0000000	PT. Summit Oto F	Desyiana	18075046937820005	1807504678848000
30	PT TOM (Sekampung)	BEAT SPORTY CBS	LXL	WH	AT	AT LOW	BEAT SPORTY SERIES	Salesman	Cash	-	-	Ilwanan	180751171948001	180751171948000

Gambar 4.1 Data Transaksi

#### 4.2 Seleksi Data

Data transaksi di saring terlebih dahulu dan diambil beberapa Atribut dari tabel untuk dianalisis, gunakan Tools Rapid Miner lalu klik blank process, klik import data yang akan digunakan dalam bentuk file

excel, csv atau sumber data yang akan di seleksi, pilih Atribut yang akan digunakan dengan cara mengexclude column, selanjutnya klik next dan klik salah satu file agar dapat menyimpan data yang telah di exclude, dan klik finish seperti gambar dibawah yang Terdapat enam variabel yang digunakan yaitu, Unit Series, jenis penjualan, jenis kelamin, pendidikan, kecamatan, pekerjaan. Pengambilan data berdasarkan variabel yang digunakan dan penyeleksian data menggunakan *selection data* yang ada pada fitur *Rapid Miner*.



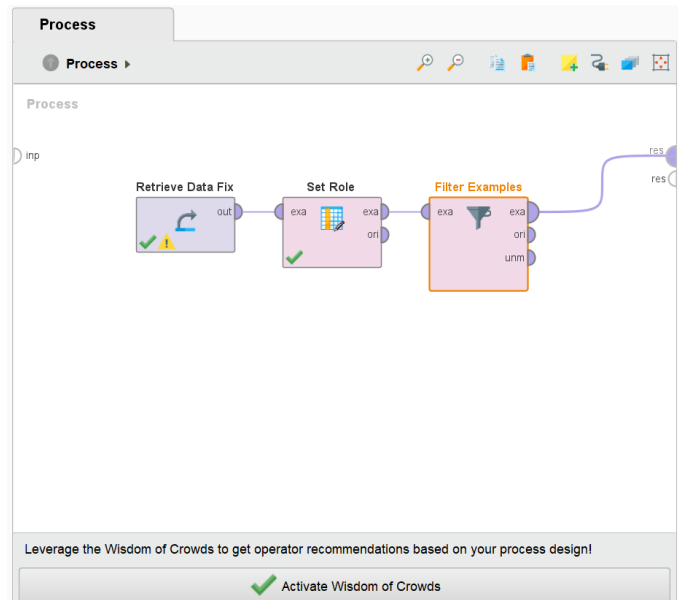
Gambar 4.2 Seleksi Data

### 4.3 Pembersihan Data (*Data Cleaning*)

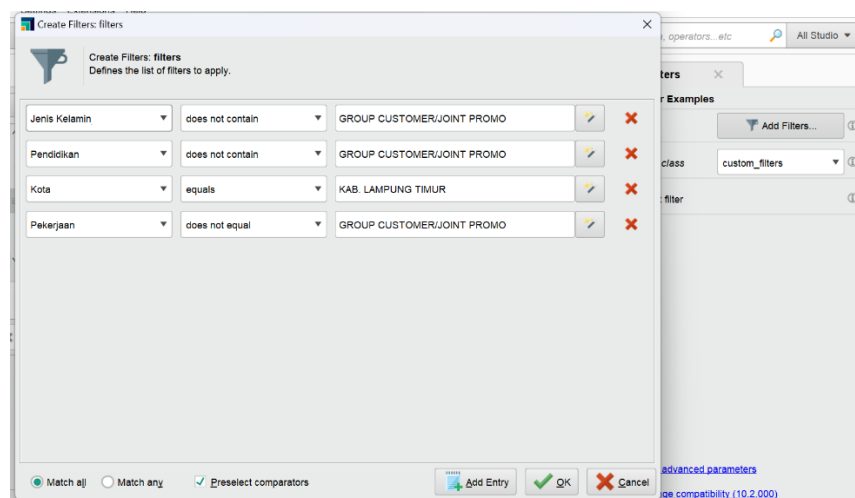
Proses mengidentifikasi dan mengoreksi atau membuang data yang salah dari dataset. Data yang dibersihkan berupa data yang tidak diperlukan, atau format yang dapat mengganggu proses analisis data selanjutnya.

Dengan menggunakan “Pencarian Filter” untuk menemukan dan menghapus data yang tidak diperlukan, pergi ke Tools Rapid Miner lalu pilih “Blank Process”, masukkan file data excel yang telah di seleksi, lalu cari dalam operator “Filter Examples”, hubungkan operasi “Filter Examples” dengan data yang akan di filter, klik parameters sebelah kanan atas dan add Atribut/ data yang ingin di filter selanjutnya klik

“Run” dibagian atas dari Tools Rapid Miner. Seperti pada gambar 4.3 dan 4.4.



Gambar 4.3 Operator Tools



Gambar 4. 4 Data Yang Tidak Diperlukan

#### 4.4 Transformasi Data

Pada tahap ini dilakukan proses perubahan data, agar data dapat diolah dengan algoritma *k-means clustering*. Data yang *non-numeric* dilakukan proses inisiasi ke dalam bentuk *numeric*. Namun jika data yang dimiliki sudah dalam bentuk *numeric* maka tidak perlu dilakukan inisiasi. tentukan angka tiap data untuk merubah data *non-numeric* ke

dalam bentuk *numeric*, lalu gunakan fungsi dengan cara menulis rumus “*Index-Match*” untuk mencocokkan nilai dari satu daerah data ke daerah data lainnya, Variabel yang digunakan pada penelitian ini adalah Unit Series, jenis penjualan, jenis kelamin, pendidikan, kecamatan, pekerjaan. Proses inisiasi sebagai berikut.

Unit Series	Kode Jenis Penjualan	Kode Jenis	Kode pendidikan	Kode kecamatan	Kode pekerjaan
VARIO SERIES	1	1	1	1	1
BEAT STREET SERIES	1	1	2	2	1
BEAT SPORTY SERIES	2	1	2	3	2
VARIO 125 SERIES	2	2	2	1	3
BEAT STREET SERIES	1	2	2	1	3
SCOOPY SERIES	2	2	2	4	3
BEAT STREET SERIES	2	2	2	1	2
BEAT STREET SERIES	2	2	2	3	3
VARIO SERIES	2	1	2	1	1
SCOOPY SERIES	2	1	2	3	3
VARIO SERIES	2	2	2	3	2
BEAT STREET SERIES	2	1	2	1	2
VARIO 125 SERIES	2	2	2	5	2
VARIO 150 SERIES	2	1	2	1	1
BEAT SPORTY SERIES	2	1	2	1	1
VARIO SERIES	2	1	2	6	4
VARIO 125 SERIES	1	2	2	7	4
BEAT STREET SERIES	2	1	2	8	1
VARIO SERIES	1	2	2	2	3
BEAT SPORTY SERIES	1	1	2	1	1
VARIO SERIES	1	1	2	1	1
VARIO SERIES	2	1	2	1	3
VARIO 125 SERIES	2	2	2	1	3
BEAT SPORTY SERIES	1	1	2	1	1
BEAT SPORTY SERIES	2	1	2	9	1
BEAT STREET SERIES	2	2	3	1	5
BEAT SPORTY SERIES	2	2	2	2	2
BEAT STREET SERIES	2	2	2	1	2
BEAT SPORTY SERIES	1	1	2	3	1
BEAT SPORTY CBS	2	2	2	1	3

Gambar 4.5 Transformasi Data

#### a) Unit Series

Variabel yang berisikan jenis produk yang dijual dan dikeluarkan oleh PT. Tunas Dwipa Matra dengan jumlah 7078 data, Unit Series digunakan untuk menentukan tingkatan penjualan produk tersebut, seperti gambar dibawah ini.

Unit Series
VARIO SERIES
BEAT STREET SERIES
BEAT SPORTY SERIES
VARIO 125 SERIES
BEAT STREET SERIES
SCOOPY SERIES
BEAT STREET SERIES
BEAT STREET SERIES
VARIO SERIES
SCOOPY SERIES
VARIO SERIES
BEAT STREET SERIES
VARIO 125 SERIES
VARIO 150 SERIES
BEAT SPORTY SERIES
VARIO SERIES
VARIO 125 SERIES
BEAT STREET SERIES
VARIO SERIES
BEAT SPORTY SERIES
VARIO SERIES
VARIO SERIES
VARIO 125 SERIES
BEAT SPORTY SERIES
BEAT SPORTY SERIES
BEAT STREET SERIES
BEAT SPORTY SERIES
BEAT STREET SERIES
BEAT SPORTY SERIES
BEAT STREET SERIES
BEAT SPORTY SERIES
BEAT SPORTY CBS

Gambar 4.6 Variabel Unit Series

## b) Jenis Penjualan

Data pada variabel jenis penjualan ini diinisiasikan berdasarkan nilai/ angka yang telah ditentukan dan dan berapa banyak yang menggunakan transaksi pembelian melalui cash sebanyak 4439 data dan credit sebanyak 2639 data, seperti pada gambar dibawah ini.

Jenis Penjualan	jumlah	inisal
Credit	2639	1
Cash	4439	2

Gambar 4.7 Jenis Penjualan

## c) Pendidikan

Data pada variabel pendidikan ini diinisiasikan berdasarkan jumlah frekuensi yang terkecil hingga terbesar dengan niali yang telah ditentukan. Dan memiliki jumlah data yang berbeda seperti, SLTP/SMP memiliki 200 data, SLTA/SMA 6641 data, Akademi/Diploma 91 data, Sarjana 135 data, SD 5 data, Pasca Sarjana 3 data dan Tidak Tamat SD memiliki 3 data.

Pendidikan	jumlah	inisial
SLTP/SMP	200	1
SLTA/SMU	6641	2
Akademi/Diploma	91	3
Sarjana	135	4
SD	5	5
Pasca Sarjana	3	6
Tidak Tamat SD	3	7

Gambar 4.8 Pendidikan

## d) Kabupaten

Data yang digunakan dalam Variabel Kabupaten, dengan menggunakan Kab. Lampung Timur untuk menentukan Visualisasi berbentuk maps di wilayah tersebut, seperti gambar di bawah.

Kabupaten
KAB. LAMPUNG TIMUR

Gambar 4. 9 Variabel Kabupaten

## e) Kecamatan

Varibel Kecamatan merupakan alamat dari data konsumen yang berada di wilayah Kab. Lampung Timur dengan berbagai macam kecamatan seperti, Sekampung memiliki 3047 data, Margatiga 1315 data, Batanghari 1179 data, Way Bungur 46 data, Batanghari Nuban 146 data, Pekalongan 125 data, Jabung 14 data, Bandar Sribawono 12 data, Bumi Agung 432 data, Raman Utara 79 data, Sukadana 216 data, Sekampung Udik 191 data, Labuhan Ratu 26 data, Way Jepara 24 data, Purbolinggo 81 data, Metro Kibang 80 data, Pasir Sakti 4 data, Waway Karya 3 data, Marga Sekampung 17 data, Braja Sebah 13 data, Labuhan Maringgai 22 data, Mataram Baru 6 data, dan Gunung Pelindung

memiliki 1 data dengan inisial yang telah ditentukan seperti gambar dibawah.

kecamatan	jumlah	inisial
SEKAMPUNG	3047	1
MARGATIGA	1315	2
BATANGHARI	1179	3
WAY BUNGUR	46	4
BATANGHARI NUBAN	146	5
PEKALONGAN	125	6
JABUNG	14	7
BANDAR SRIBHAWON	12	8
BUMI AGUNG	432	9
RAMAN UTARA	79	10
SUKADANA	216	11
SEKAMPUNG UDIK	191	12
LABUHAN RATU	26	13
WAY JEPARA	23	14
PURBOLINGGO	81	15
METRO KIBANG	80	16
PASIR SAKTI	4	17
WAWAY KARYA	3	18
MARGA SEKAMPUNG	17	19
BRAJA SELEBAH	13	20
LABUHAN MARINGGA	22	21
MATARAM BARU	6	22
GUNUNG PELINDUNG	1	23

Gambar 4. 10 Variabel Kecamatan

f) Pekerjaan

Variabel pekerjaan merupakan pekerjaan yang dilakukan konsumen sehari-hari, dengan adanya tabel tersebut agar dapat membantu sales dalam melakukan pemilihan produk yang cocok untuk setiap konsumen yang memiliki pekerjaan seperti, Ibu Rumah Tangga 663 data, Petani 3804 data, Lain-Lain 612 data, Mahasiswa/Pelajar 383 data, Pegawai Negeri 84 data, TNI/Polri 13 data, Guru/Dosen 75 data, Jasa/Wiraswasta/Pedagang 137 data, Industri- Pegawai Swasta 9 data, Perdagangan

(Retail) Wiraswasta 500 data, Jasa-Pegawai Swasta 26 data, Pertambangan-Wiraswasta 10 data, Pertanian/Perkebunan 62 data, Kontruksi-Wiraswasta 4 data, Kontruksi Pegawai Swasta 6 data, Industri-Wiraswasta/Pedagang 5 data, Perdagangan (Retail) 564 data, Ojek 39 data, Nelayan 48 data, dan Pertanian/Perkebunan memiliki 30 data.

Pekerjaan	jumlah	inisial
Ibu Rumah Tangga	663	1
Petani	3804	2
Lain-lain	612	3
Mahasiswa/Pelajar	383	4
Pegawai Negeri	84	5
TNI/Polri	13	6
Guru/Dosen	75	7
Jasa - Wiraswasta / Pedagan	137	8
Industri - Pegawai Swasta	9	9
Perdagangan (Retail) - Wiras	500	10
Jasa - Pegawai Swasta	26	11
Pertambangan - Wiraswasta	10	12
Pertanian / Perkebunan / Ke	62	13
Konstruksi - Wiraswasta / Pe	4	14
Konstruksi - Pegawai Swasta	6	15
Industri - Wiraswasta / Peda	5	16
Perdagangan (Retail) - Pegav	564	17
Ojek	39	18
Nelayan	48	19
Pertanian / Perkebunan / Ke	30	20

Gambar 4. 11 Pekerjaan

#### 4.5 Menghitung menggunakan tahapan-tahapan Algoritma *K-means*

Ada beberapa tahapan untuk mendapatkan hasil perhitungan dengan menggunakan *k-means*

##### a) Penentuan *Cluster* Awal

dalam menentukan  $n$  buah pusat *cluster* awal dilakukan pembangkitan bilangan random yang merepresentasikan urutan data *input*. Pusat awal *cluster* didapatkan dari data sendiri bukan dengan menentukan titik baru, yaitu dengan random pusat awal dari data. C1 merupakan tingkat tertinggi dari penjualan wilayah tertentu, C2 merupakan tingkat sedang dari penjualan wilayah tertentu, dan C3 merupakan tingkat terendah dari penjualan wilayah tertentu.



Guna mempermudah melakukan perhitungan, maka jenis penjualan di inialisasikan menjadi JP, jenis kelamin menjadi JK, pendidikan menjadi Pend, kecamatan menjadi Kec, dan pekerjaan menjadi Pek. Berikut adalah tabel dari data *random*.

Cluster	JP	JK	Pend	Kec	Pek
c1	1	1	1	1	1
c2	1	1	2	2	1
c3	2	1	2	3	2

Gambar 4. 12 Penentuan Cluster Awal

b) Penentuan jarak dengan pusat *cluster*

Untuk mengukur jarak antar data dengan *cluster* digunakan *euclidian distance*, algoritma perhitungan jarak data dengan pusat *cluster*.

berikut ini merupakan proses perhitungan jarak *Centroid* masing masing *cluster*.

Jarak data ke-1 ke pusat *cluster*

C1

$$\sqrt{(1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2}$$

$$= 0$$

C2

$$\sqrt{(1-1)^2 + (1-1)^2 + (2-1)^2 + (2-1)^2 + (1-1)^2}$$

$$= 1.41$$

C3

$$\sqrt{(2-1)^2 + (1-1)^2 + (2-1)^2 + (3-1)^2 + (2-1)^2}$$

$$= 2.44$$

Jarak Data ke-2 ke pusat *cluster*

C1

$$\sqrt{(1-1)^2 + (1-1)^2 + (2-1)^2 + (2-1)^2 + (1-1)^2}$$

$$= 1.41$$

C2

$$\sqrt{(1-1)^2 + (1-1)^2 + (2-2)^2 + (2-2)^2 + (1-1)^2}$$

$$= 0$$

C3

$$\sqrt{(1-2)^2 + (1-1)^2 + (2-2)^2 + (2-3)^2 + (2-1)^2}$$

$$= 2.44$$

Tabel 4. 1 Cluster Awal

ID	C1	C2	C3	Nilai minimum	Termasuk dalam Cluster =
1	0	1.41421	2.44949	0	C1
2	1.41421	0	1.41421	0	C2
3	2.64575	1.73205	0	0	C3
4	2.64575	2.64575	2.44949	2.449489743	C3
5	2.44949	2.44949	2.44949	2.449489743	C1
6	4	3.16228	1.73205	1.732050808	C3
7	2	2	2.23607	2	C1
8	3.31662	2.64575	1.41421	1.414213562	C3
9	1.41421	1.41421	2.23607	1.414213562	C1
10	3.16228	2.44949	1	1	C3
11	2.82843	2	1	1	C3
12	1.73205	1.73205	2	1.732050808	C1
13	4.47214	3.4641	2.23607	2.236067977	C3
14	1.41421	1.41421	2.23607	1.414213562	C1
15	1.41421	1.41421	2.23607	1.414213562	C1
16	6	5.09902	3.60555	3.605551275	C3
17	6.85565	5.91608	4.58258	4.582575695	C3
18	7.14143	6.08276	5.09902	5.099019514	C3
19	2.64575	2.23607	1.73205	1.732050808	C3
20	1	1	2.23607	1	C1
21	1	1	2.23607	1	C1
22	2.44949	2.44949	2.23607	2.236067977	C3
23	2.64575	2.64575	2.44949	2.449489743	C3
24	1	1	2.23607	1	C1
25	8.12404	7.07107	6.08276	6.08276253	C3
26	4.69042	4.47214	3.87298	3.872983346	C3
27	2.23607	1.73205	1.41421	1.414213562	C3
28	2	2	2.23607	2	C1

Tabel diatas adalah hasil perhitungan jarak setiap data dengan masing-masing *cluster*. Jarak yang terlihat antar data dan *Centroid* di setiap *cluster*, jika jarak dua titik semakin dekat, maka semakin dekat pula kesamaan antara kedua titik tersebut. Data pertama mempunyai jarak paling dekat dengan *Centroid* kelompok 1 dibandingkan dengan *Centroid* dari kelompok lain, maka bisa disimpulkan data pertama mempunyai karakteristik paling dekat dengan kelompok 1 dibanding dengan kelompok lain sehingga data pertama dimasukkan kedalam kelompok 1. Pengelompokan data lainnya seperti pada tabel diatas.

Dari data diatas dapat disimpulkan bahwa data yang masuk ke masing-masing cluster di Iterasi ke-1 dapat dilihat pada tabel dibawah ini.

ID	JP	JK	Pend	Kec	Pek
1	1	1	1	1	1
5	1	2	2	1	3
7	2	2	2	1	2
9	2	1	2	1	1
12	2	1	2	1	2
14	2	1	2	1	1
15	2	1	2	1	1
20	1	1	2	1	1
21	1	1	2	1	1
24	1	1	2	1	1
28	2	2	2	1	2

Gambar 4. 13 Kelompok data C1 pada Iterasi ke-1

Pada iterasi ke-1 dapat dilihat bahwa data yang termasuk dalam kelompok C1 terdapat 11 data, yaitu data dengan id pada gambar 4.13

ID	JP	JK	Pend	Kec	Pek
2	1	1	2	2	1
29	1	1	2	3	1

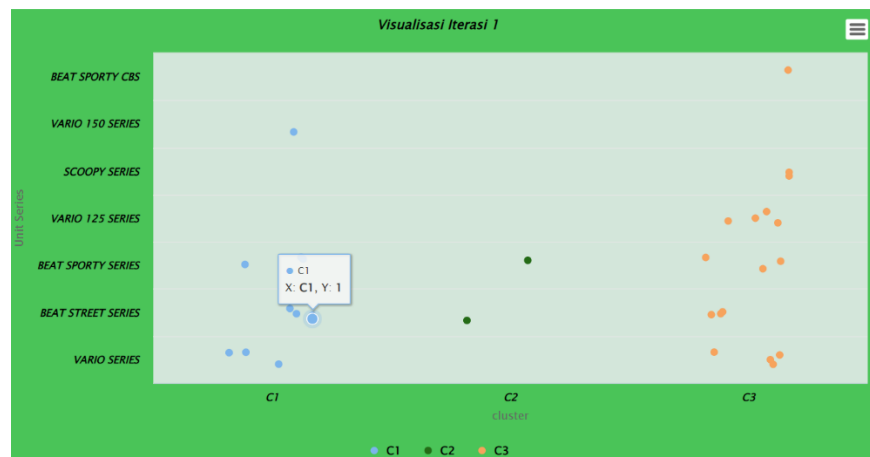
Gambar 4. 14 Kelompok data C2 pada Iterasi ke-1

Pada iterasi ke-2 dapat dilihat bahwa data yang termasuk dalam kelompok C2 terdapat 2 data, yang dapat dilihat pada gambar 4.14

ID	JP	JK	Pend	Kec	Pek
3	2	1	2	3	2
4	2	2	2	1	3
6	2	2	2	4	3
8	2	2	2	3	3
10	2	1	2	3	3
11	2	2	2	3	2
13	2	2	2	5	2
16	2	1	2	6	4
17	1	2	2	7	4
18	2	1	2	8	1
19	1	2	2	2	3
22	2	1	2	1	3
23	2	2	2	1	3
25	2	1	2	9	1
26	2	2	3	1	5
27	2	2	2	2	2
30	2	2	2	1	3

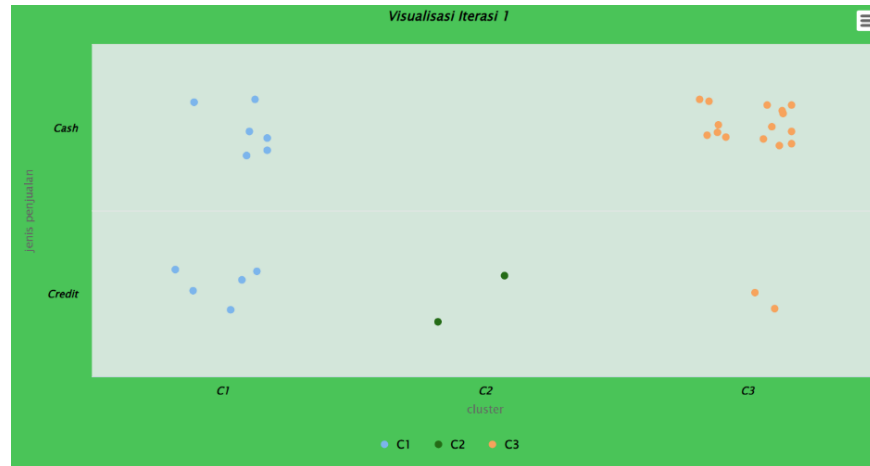
Gambar 4. 15 Kelompok data C3 pada Iterasi ke-1

Dan data yang termasuk dalam C3 pada iterasi ke-1 yaitu berjumlah 17 data, dengan data yang dapat dilihat pada gambar 4.15



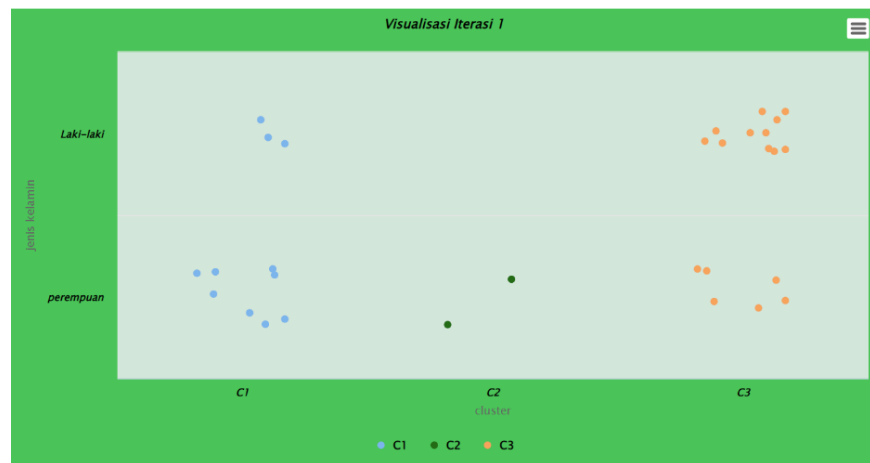
Gambar 4. 16 Visualisasi Unit Series Iterasi 1

Pada iterasi ke-1, dengan Atribut Unit Series data C1 ada pada jumlah 36.66%, C2 dengan jumlah 6.66% dan C3 ada pada jumlah 56.66%



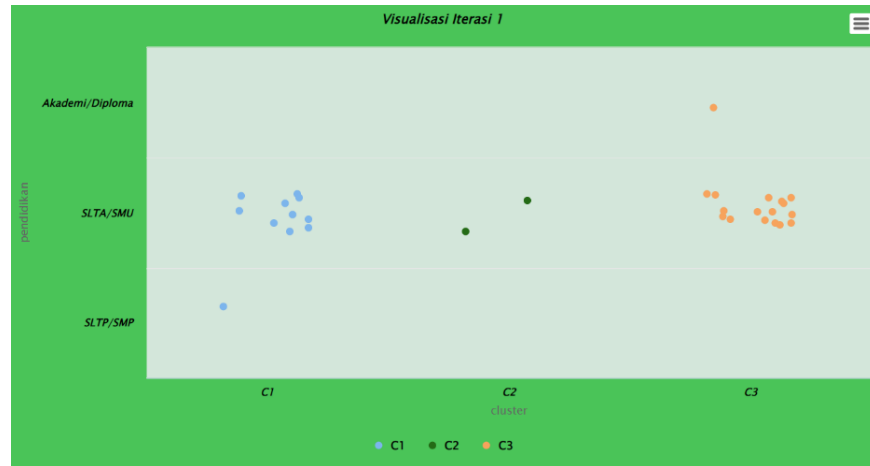
Gambar 4. 17 Visualisasi Jenis Penjualan Iterasi 1

Pada gambar 4.17 dapat dilihat dari gambar diatas bahwa data terbanyak ada pada C3 dengan jumlah 56.66, C1 dengan jumlah 36.66% dan C2 dengan jumlah 6.66%.



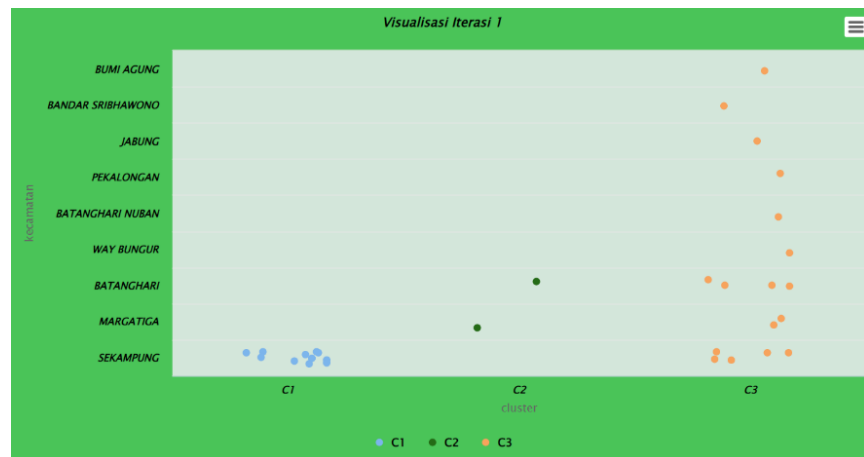
Gambar 4. 18 Visualisasi Jenis Kelamin Iterasi 1

Pada gambar 4.18 dapat dilihat dari gambar diatas bahwa penjualan terbanyak pada C3 56.66%, C1 dengan jumlah 36.66% dan C2 berjumlah 6.66%.



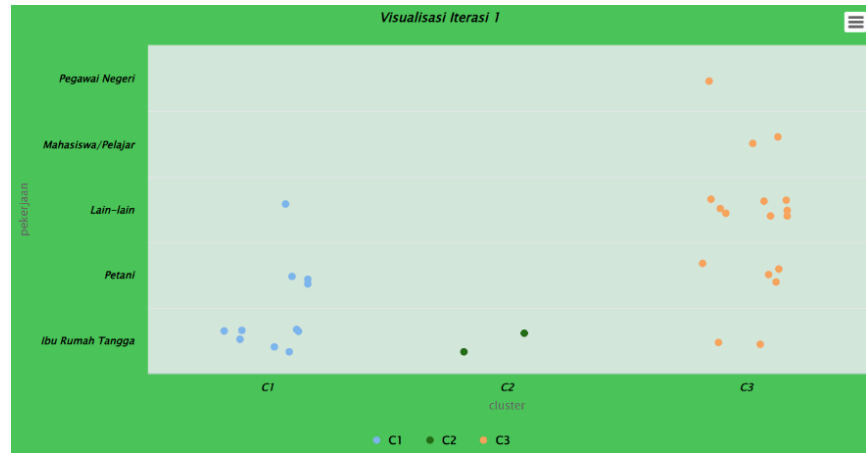
Gambar 4. 19 Visualisasi Pendidikan Iterasi 1

Pada gambar 4.19 dapat dilihat bahwa pendidikan terbanyak ada pada C3 56.66%, C1 dengan jumlah 36.66% dan C2 berjumlah 6.66%.



Gambar 4. 20 Visualisasi Kecamatan Iterasi 1

Pada gambar 4.20 dapat dilihat bahwa penjualan terbanyak terdapat pada C1 dengan jumlah 36.66%, C2 dengan jumlah 6.66% dan C3 dengan jumlah 56,66%.



Gambar 4. 21 Visualisasi Pekerjaan Iterasi 1

Pada gambar 4.21 dapat dilihat pada gambar diatas bahwa C1 dengan jumlah 36.66% , C2 dengan jumlah 6.66% dan C3 56.66%

### c) Menentukan titik *Centroid* baru

Menentukan titik *Centroid* baru (*Centroid* Iterasi ke-2)

Cluster	JP	JK	Pend	Kec	Pek
C1	1,545455	1,272727	1,909091	1	1,454545
C2	1	1	2	2,5	1
C3	1,882353	1,647059	2,058824	3,529412	2,764706

Gambar 4. 22 titik *Centroid* baru Iterasi ke-2

Jarak antara pusat data *Centroid* dengan masing-masing data, menentukan titik pusat data centroid dapat dicari dengan mencari rata – rata dari masing-masing iterasi terhadap masing-masing Atribut

ID	C1	C2	C3
1	1,18531	1,802776	3,439542
2	1,259673	0,5	2,579531
3	2,141676	1,5	1,140631
4	1,769811	2,872281	2,568104
5	1,795311	2,692582	2,712906
6	3,483135	2,872281	0,647059
7	1,020452	2,291288	2,669189
8	2,670624	2,5	0,69102
9	0,704179	1,802776	3,15406
10	2,584122	2,291288	0,878423
11	2,245289	1,802776	1,003454
12	0,766014	2,061553	2,723726
13	4,128114	3,041381	1,699786
14	0,704179	1,802776	3,15406
15	0,704179	1,802776	3,15406
16	5,636364	4,716991	2,840025
17	6,581341	5,5	3,804932
18	7,03533	5,59017	4,851426
19	2,055028	2,291288	1,816876
20	0,766014	1,5	3,273042
21	0,766014	1,5	3,273042
22	1,636364	2,692582	2,624743
23	1,769811	2,872281	2,568104
24	0,766014	1,5	3,273042
25	8,030932	6,576473	5,785975
26	3,807344	4,609772	3,524015
27	1,428749	1,802776	1,750927
28	1,020452	2,291288	2,669189
29	2,141676	0,5	2,143631
30	1,769811	2,872281	2,568104

Gambar 4. 23 jarak data dengan *Centroid* (Iterasi ke-2)

Pada gambar 4.23 dapat dilihat jarak antara masing-masing data dengan masing-masing centroid.



ID	JP	JK	Pend	Kec	Pek
1	1	1	1	1	1
4	2	2	2	1	3
5	1	2	2	1	3
7	2	2	2	1	2
9	2	1	2	1	1
12	2	1	2	1	2
14	2	1	2	1	1
15	2	1	2	1	1
20	1	1	2	1	1
21	1	1	2	1	1
22	2	1	2	1	3
23	2	2	2	1	3
24	1	1	2	1	1
27	2	2	2	2	2
28	2	2	2	1	2
30	2	2	2	1	3

Gambar 4. 24 kelompok data pada C1 di Iterasi ke-2

Pada gambar 4.24 dapat dilihat bahwa data yang termasuk pada C1 di iterasi ke-2 berjumlah 16 data

ID	JP	JK	Pend	Kec	Pek
2	1	1	2	2	1
29	1	1	2	3	1

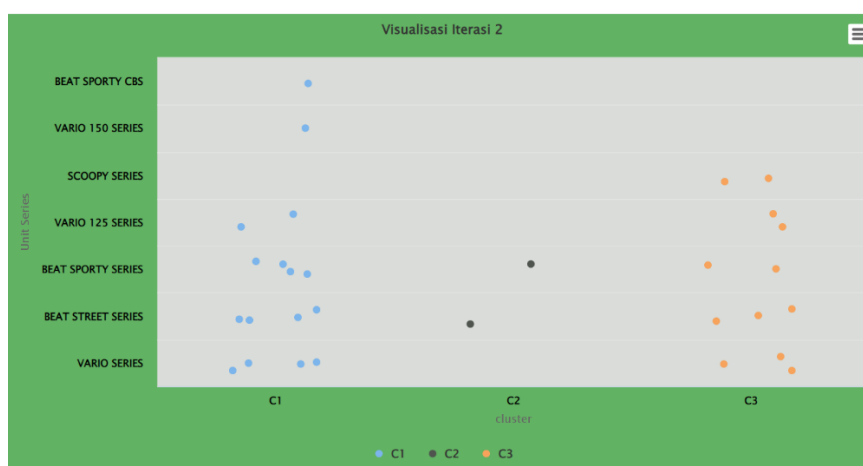
Gambar 4. 25 kelompok data pada C2 di Iterasi ke-2

Pada gambar 4.25 dapat dilihat data yang termasuk pada C2 di iterasi ke-2 berjumlah 2 data.

ID	JP	JK	Pend	Kec	Pek
3	2	1	2	3	2
6	2	2	2	4	3
8	2	2	2	3	3
10	2	1	2	3	3
11	2	2	2	3	2
13	2	2	2	5	2
16	2	1	2	6	4
17	1	2	2	7	4
18	2	1	2	8	1
19	1	2	2	2	3
25	2	1	2	9	1
26	2	2	3	1	5

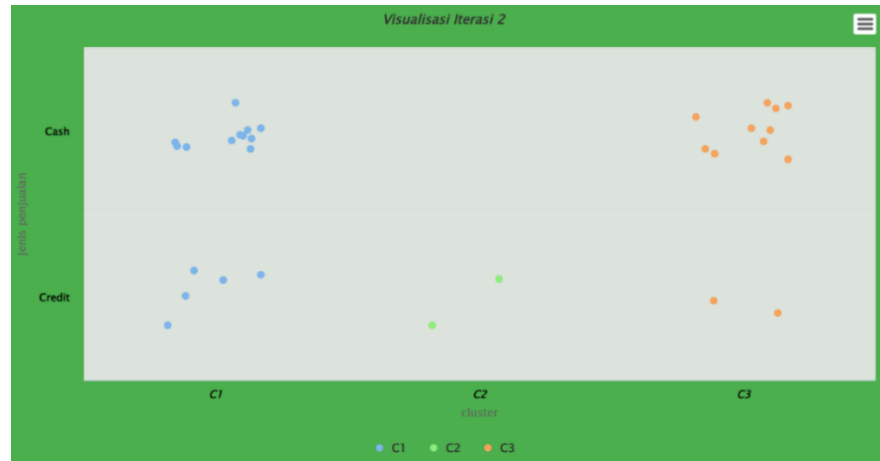
Gambar 4. 26 kelompok data pada C3 di Iterasi ke-2

Pada gambar 4.26 dapat dilihat bahwa data yang termasuk pada C3 di iterasi ke-2 berjumlah 12 data



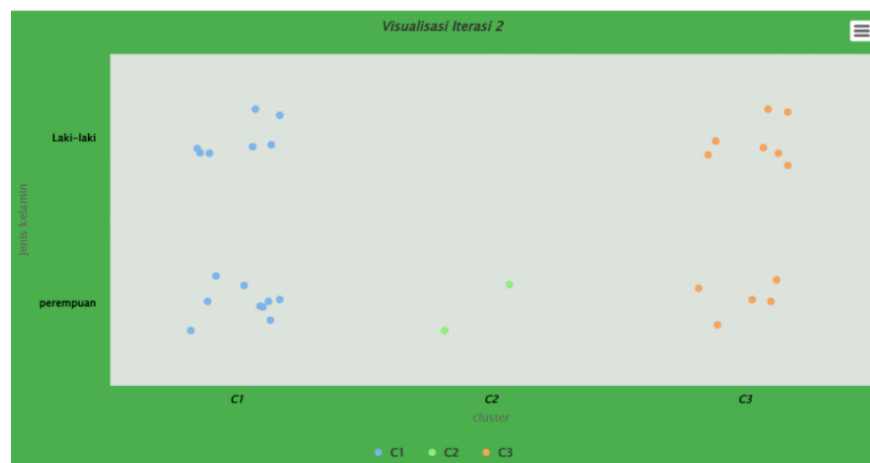
Gambar 4. 27 Visualisasi Unit Series Iterasi 2

Pada gambar 4.27 dapat dilihat bahwa data pada Unit Series C1 dengan jumlah 53.33%, C2 dengan jumlah 6.66% dan C3 dengan jumlah 40%.



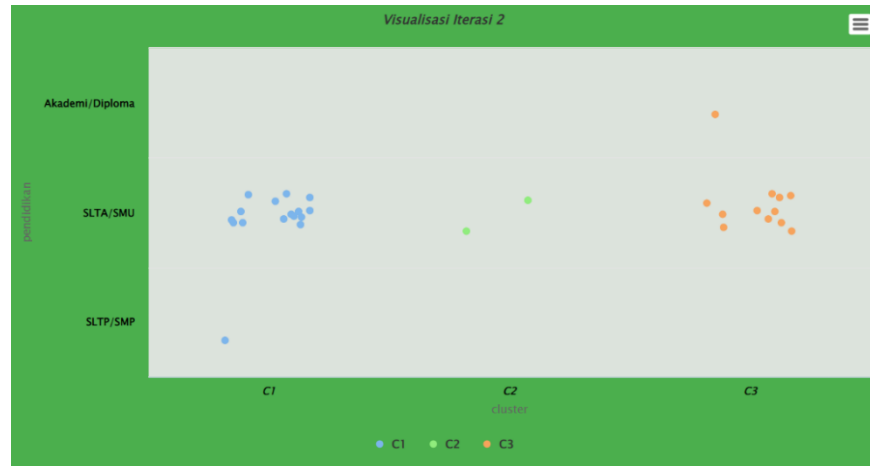
Gambar 4. 28 Visualisasi Jenis Penjualan Iterasi 2

Pada gambar 4.28 dapat dilihat bahwa pada titik data terbanyak terdapat di C1 dengan jumlah 53.33%, C3 dengan 40% dan terkecil pada C2 dengan jumlah 6.66%.



Gambar 4. 29 Visualisasi Jenis Kelamin Iterasi 2

Pada gambar 4.29 dapat dilihat bahwa pembagian titik data pada C1 sejumlah 53.33%, C3 sejumlah 40% dan C2 sebanyak 6.66%.



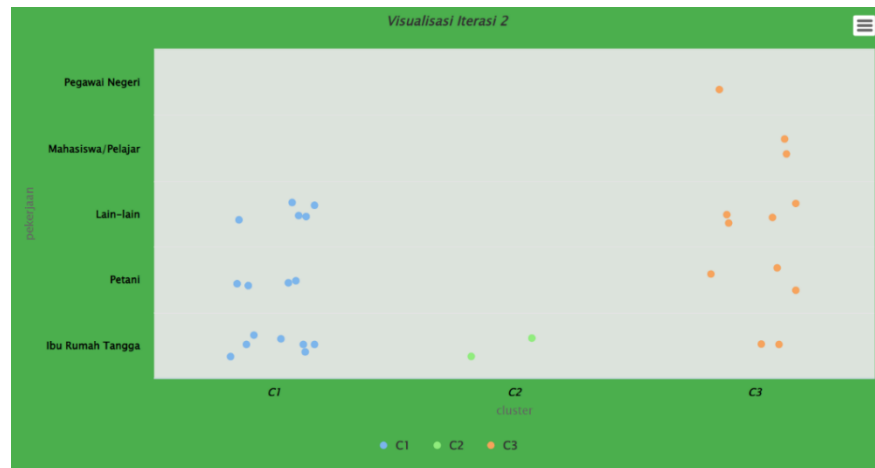
Gambar 4. 30 Visualisasi Pendidikan Iterasi 2

Pada gambar 4.30 dapat dilihat bahwa pembagian titik data pada C1 sejumlah 53.33%, C3 sejumlah 40% dan C2 sebanyak 6.66%.



Gambar 4. 31 Visualisasi Kecamatan Iterasi 2

Pada gambar 4.31 dapat dilihat bahwa pembagian titik data pada C1 sejumlah 53.33%, C3 sejumlah 40% dan C2 sebanyak 6.66%.



Gambar 4. 32 Visualisasi Pekerjaan Iterasi 2

Pada gambar 4.32 dapat dilihat bahwa pembagian titik data pada C1 sejumlah 53.33%, C3 sejumlah 40% dan C2 sebanyak 6.66%.

Cluster	JP	JK	Pend	Kec	Pek
C1	1,6875	1,4375	1,9375	1,0625	1,875
C2	1	1	2	2,5	1
C3	1,833333	1,583333	2,083333	4,5	2,75

Gambar 4. 33 *Centroid* baru di Iterasi ke-3

Pada gambar 4.33 dapat dilihat centroid baru yang akan digunakan pada perhitungan di iterasi ke-3, centroid tersebut didapat dari rata-rata cluster pada masing-masing Atribut di iterasi ke-2.

Id	C1	C2	C3
1	1,520691	1,802776	4,185789
2	1,520691	0,5	3,217789
3	2,015564	1,5	1,785357
4	1,299038	2,872281	3,538479
5	1,436141	2,692582	3,63146
6	3,211308	2,872281	0,721688
7	0,661438	2,291288	3,608439
8	2,331845	2,5	1,587713
9	1,030776	1,802776	3,960745
10	2,304886	2,291288	1,63936
11	2,046338	1,802776	1,738054
12	0,559017	2,061553	3,63146
13	3,99218	3,041381	1,010363
14	1,030776	1,802776	3,960745
15	1,030776	1,802776	3,960745
16	5,402546	4,716991	2,046338
17	6,36887	5,5	2,947457
18	7,01338	5,59017	3,960745
19	1,713914	2,291288	2,680951
20	1,198958	1,5	4,044029
21	1,198958	1,5	4,044029
22	1,25	2,692582	3,561952
23	1,299038	2,872281	3,538479
24	1,198958	1,5	4,044029
25	8,003905	6,576473	4,866981
26	3,363406	4,609772	4,284176
27	1,145644	1,802776	2,649686
28	0,661438	2,291288	3,608439
29	2,277608	0,5	2,520747
30	1,299038	2,872281	3,538479

Gambar 4. 34 jarak antara data dengan *Centroid* Iterasi 3

Terlihat pada gambar 4.34 yaitu merupakan jarak antara masing-masing data dengan centroid pada iterasi ke-3.

ID	JP	JK	Penc	Kec	Pek
1	1	1	1	1	1
4	2	2	2	1	3
5	1	2	2	1	3
7	2	2	2	1	2
9	2	1	2	1	1
12	2	1	2	1	2
14	2	1	2	1	1
15	2	1	2	1	1
19	1	2	2	2	3
20	1	1	2	1	1
21	1	1	2	1	1
22	2	1	2	1	3
23	2	2	2	1	3
24	1	1	2	1	1
26	2	2	3	1	5
27	2	2	2	2	2
28	2	2	2	1	2
30	2	2	2	1	3

Gambar 4. 35 kelompok data C1 di Iterasi ke-3

Pada gambar 4.35, dapat dilihat terdapat 18 data pada kelompok C1.

ID	JP	JK	Pen	Kec	Pek
2	1	1	2	2	1
3	2	1	2	3	2
29	1	1	2	3	1

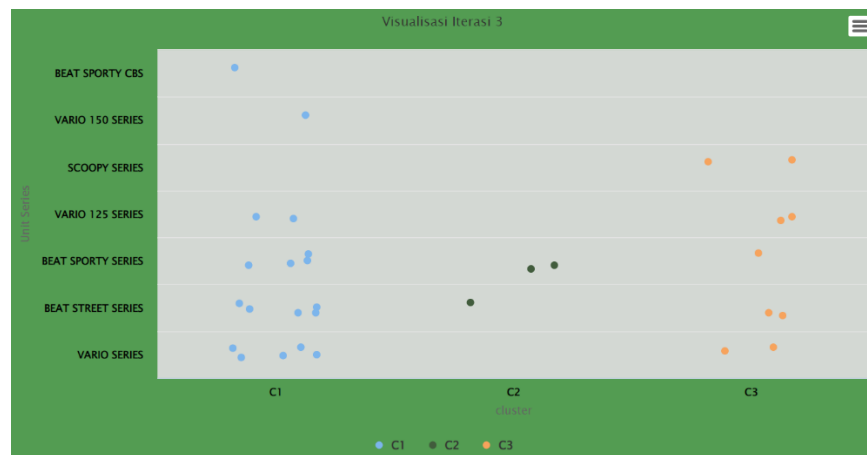
Gambar 4. 36 kelompok data C2 di Iterasi ke-3

Pada gambar 4.36 dapat dilihat bahwa hanya terdapat 3 data pada kelompok C2 di iterasi ke-3.

ID	JP	JK	Pen	Kec	Pek
6	2	2	2	4	3
8	2	2	2	3	3
10	2	1	2	3	3
11	2	2	2	3	2
13	2	2	2	5	2
16	2	1	2	6	4
17	1	2	2	7	4
18	2	1	2	8	1
25	2	1	2	9	1

Gambar 4. 37 kelompok data C3 di Iterasi ke-3

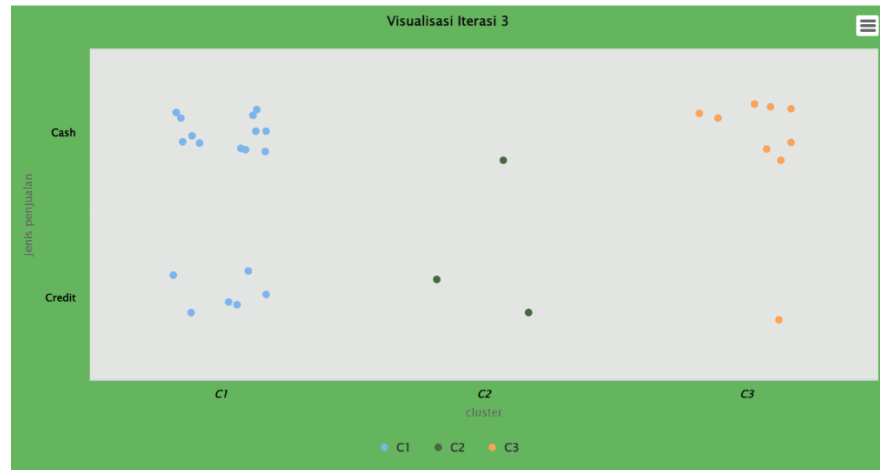
Pada gambar 4.37 dapat dilihat bahwa terdapat 9 data pada C2 di iterasi ke-3.



Gambar 4. 38 Visualisasi Unit Series Iterasi 3

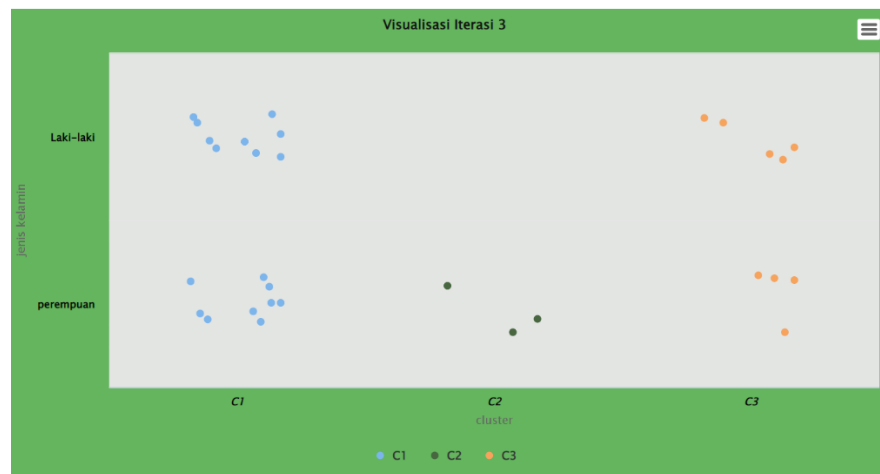
Pada gambar 4.38 dapat dilihat bahwa data di C1 memiliki jumlah tertinggi sebesar 60%, C3 memiliki jumlah sedang 30% dan C2 dengan jumlah 10%.





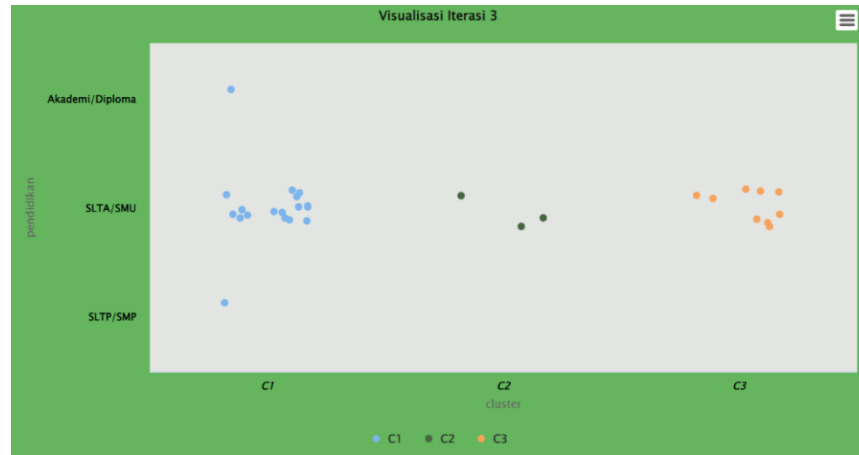
Gambar 4. 39 Visualisasi Jenis Penjualan Iterasi 3

Pada gambar 4.39 dapat dilihat bahwa data di C1 memiliki jumlah tertinggi sebesar 60%, C3 memiliki jumlah sedang 30% dan C3 dengan jumlah 10%.



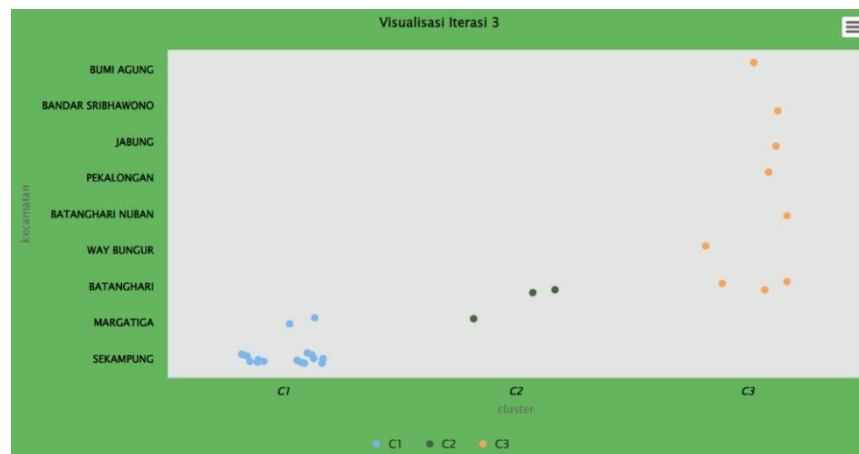
Gambar 4. 40 Visualisasi Jenis Kelamin Iterasi 3

pada gambar 4.40 dapat dilihat bahwa data di C1 memiliki jumlah tertinggi sebesar 60%, C3 memiliki jumlah sedang 30% dan C3 dengan jumlah 10%.



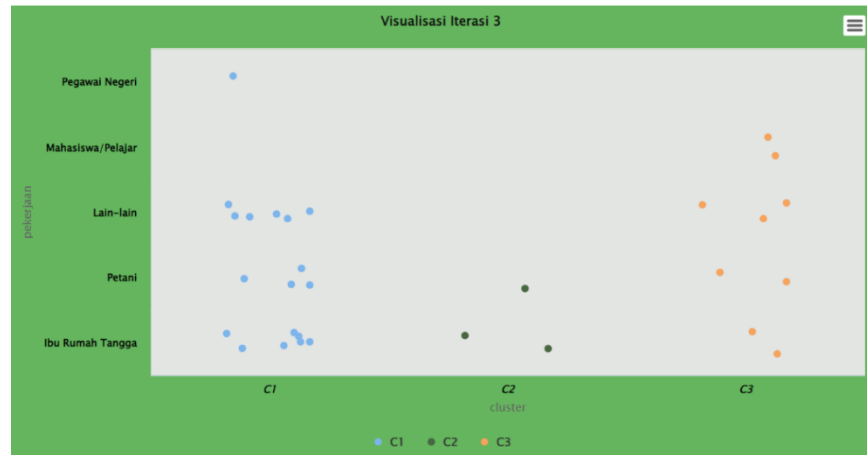
Gambar 4. 41 Visualisasi Pendidikan Iterasi 3

Pada gambar 4.41 dapat dilihat bahwa data di C1 memiliki jumlah tertinggi sebesar 60%, C3 memiliki jumlah sedang 30% dan C3 dengan jumlah 10%.



Gambar 4. 42 Visualisasi Kecamatan Iterasi 3

Pada gambar 4.42 dapat dilihat bahwa data di C1 memiliki jumlah tertinggi sebesar 60%, C3 memiliki jumlah sedang 30% dan C3 dengan jumlah 10%.



Gambar 4. 43 Visualisasi Pekerjaan Iterasi 3

Pada gambar 4.43 dapat dilihat bahwa data di C1 memiliki jumlah tertinggi sebesar 60%, C3 memiliki jumlah sedang 30% dan C3 dengan jumlah 10%.

*Centroid* pada Iterasi ke-4

Cluster	JP	JK	Pend	Kec	Pek
C1	1,666667	1,5	2	1,111111	2,111111
C2	1,333333	1	2	2,666667	1,333333
C3	1,888889	1,555556	2	5,333333	2,555556

Gambar 4. 44 Centroid baru di Iterasi ke-4

Pada gambar 4.44 menentukan centroid baru guna menentukan apakah masih ada perubahan atau tidak terhadap centroid sebelumnya, yaitu dengan mencari rata – rata dari masing-masing cluster terhadap masing-masing Atribut.

id	C1	C2	C3
1	1,715039	2	4,826624
2	1,64898	0,816497	3,82487
3	1,985285	1	2,464564
4	1,078694	2,645751	4,38009
5	1,223484	2,581989	4,467993
6	3,081706	2,44949	1,478237
7	0,62113	2,160247	4,392755
8	2,172357	2,081666	2,41906
9	1,268079	1,825742	4,638806
10	2,172357	1,825742	2,441918
11	1,985285	1,414214	2,441918
12	0,62113	1,914854	4,405384
13	3,936612	2,708013	0,793492
14	1,268079	1,825742	4,638806
15	1,268079	1,825742	4,638806
16	5,275438	4,320494	1,688743
17	6,240301	5,196152	2,41906
18	7,003747	5,385165	3,138766
19	1,508208	2,081666	3,506608
20	1,393326	1,732051	4,721895
21	1,393326	1,732051	4,721895
22	1,078694	2,44949	4,392755
23	1,078694	2,645751	4,38009
24	1,393326	1,732051	4,721895
25	7,989383	6,377042	4,023082
26	3,117553	4,320494	5,095386
27	1,078694	1,527525	3,410224
28	0,62113	2,160247	4,392755
29	2,34455	0,57735	2,993821
30	1,078694	2,645751	4,38009

Gambar 4. 45 jarak antar data dengan *Centroid* di Iterasi ke-4

Pada gambar 4.45 merupakan jarak antara data dengan centroid di iterasi ke-4, pencarian jarak ini merupakan cara untuk menentukan data masuk kedalam cluster mana.

ID	JP	JK	Pend	Kec	Pek
1	1	1	1	1	1
4	2	2	2	1	3
5	1	2	2	1	3
7	2	2	2	1	2
9	2	1	2	1	1
12	2	1	2	1	2
14	2	1	2	1	1
15	2	1	2	1	1
19	1	2	2	2	3
20	1	1	2	1	1
21	1	1	2	1	1
22	2	1	2	1	3
23	2	2	2	1	3
24	1	1	2	1	1
26	2	2	3	1	5
27	2	2	2	2	2
28	2	2	2	1	2
30	2	2	2	1	3

Gambar 4. 46 kelompok data C1 di Iterasi ke-4

Pada gambar 4.46 dapat dilihat bahwa data yang termasuk pada C1 iterasi ke-4 sebanyak 18 data

ID	JP	JK	Pend	Kec	Pek
2	1	1	2	2	1
3	2	1	2	3	2
8	2	2	2	3	3
10	2	1	2	3	3
11	2	2	2	3	2
29	1	1	2	3	1

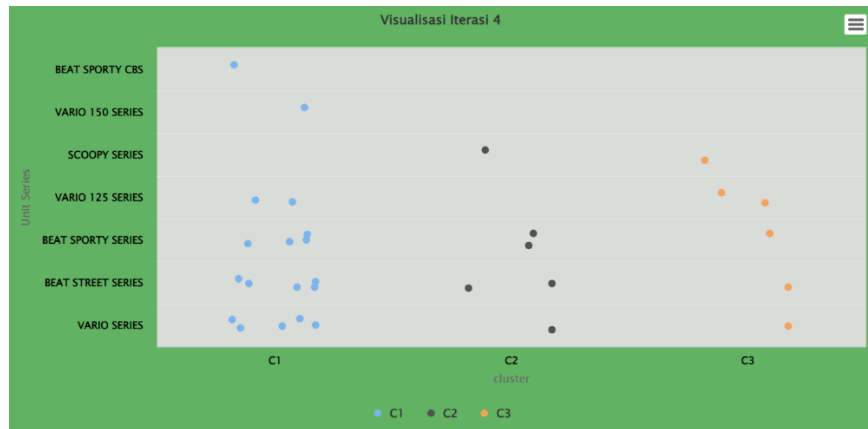
Gambar 4. 47 kelompok data C2 di Iterasi ke-4

Pada gambar 4.47 dapat dilihat data yang termasuk pada kelompok C2 di iterasi ke-4 terdapat 6 data.

ID	JP	JK	Pend	Kec	Pek
6	2	2	2	4	3
13	2	2	2	5	2
16	2	1	2	6	4
17	1	2	2	7	4
18	2	1	2	8	1
25	2	1	2	9	1

Gambar 4. 48 kelompok data C3 di Iterasi ke-4

Pada gambar 4.48 dapat dilihat bahwa data yang termasuk dalam cluster C3 di diterasi ke-4 ialah sebanyak 6 data.



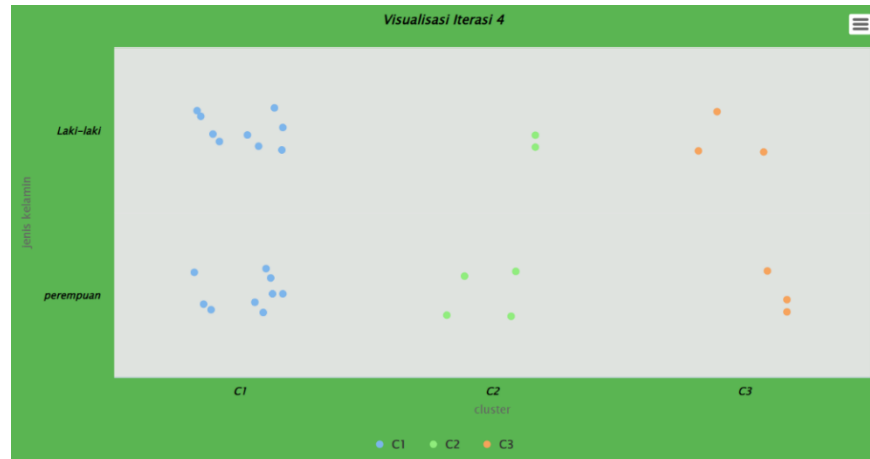
Gambar 4. 49 Visualisasi Unit Series Iterasi 4

Pada gambar 4.49 dapat dilihat bahwa data pada Atribut Unit Series tersebar cukup merata di kelompok terbanyak dengan jumlah keseluruhan C1 60%, C2 dan C3 dengan jumlah 20% data.



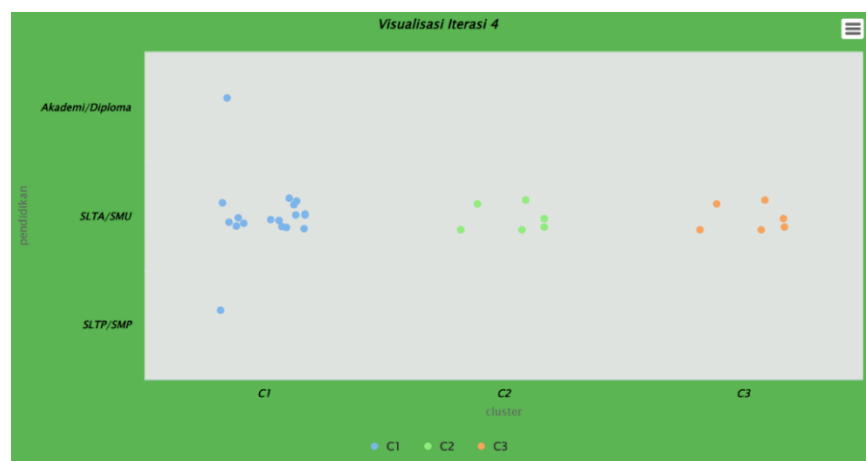
Gambar 4. 50 Visualisasi jenis Penjualan Iterasi 4

Pada gambar 4.50 dapat dilihat bahwa data pada Atribut jenis penjualan tersebar cukup merata di kelompok terbanyak dengan jumlah keseluruhan C1 60%, C2 dan C3 dengan jumlah 20% data.



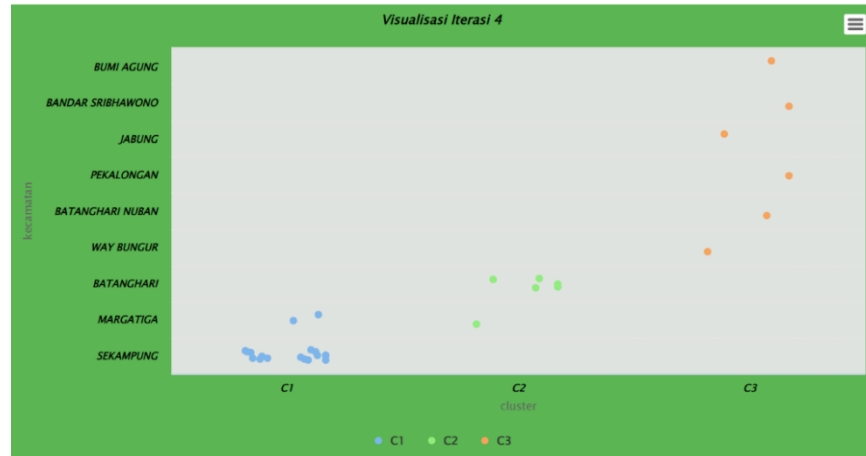
Gambar 4. 51 Visualisasi Jenis Kelamin Iterasi 4

Pada gambar 4.51 dapat dilihat bahwa data pada Atribut jenis kelamin tersebar cukup merata di kelompok terbanyak dengan jumlah keseluruhan C1 60%, C2 dan C3 dengan jumlah 20% data.



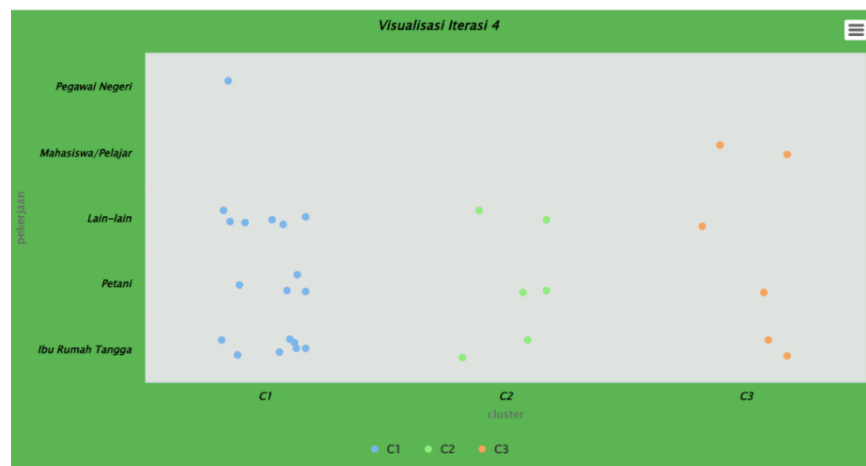
Gambar 4. 52 Visualisasi Pendidikan Iterasi 4

Pada gambar 4.52 dapat dilihat bahwa data pada Atribut Pendidikan tersebar cukup merata di kelompok terbanyak dengan jumlah keseluruhan C1 60%, C2 dan C3 dengan jumlah 20% data.



Gambar 4. 53 Visualisasi Kecamatan Iterasi 4

Pada gambar 4.53 dapat dilihat bahwa data pada Atribut Unit Series tersebar cukup merata di kelompok terbanyak dengan jumlah keseluruhan C1 60%, C2 dan C3 dengan jumlah 20% data.



Gambar 4. 54 Visualisasi Pekerjaan Iterasi 4

Pada gambar 4.54 dapat dilihat bahwa data pada Atribut pekerjaan tersebar cukup merata di kelompok terbanyak dengan jumlah keseluruhan C1 60%, C2 dan C3 dengan jumlah 20% data.

Cluster	JP	JK	Pend	Kec	Pek
C1	1,666667	1,5	2	1,111111	2,111111
C2	1,666667	1,333333	2	2,833333	2
C3	1,833333	1,5	2	6,5	2,5

Gambar 4. 55 Centroid baru di Iterasi ke-5



Pada gambar 4.55 menentukan centroid baru guna menentukan apakah masih ada perubahan atau tidak terhadap centroid sebelumnya, yaitu dengan mencari rata – rata dari masing-masing cluster terhadap masing-masing atribut.

Id	C1	C2	C3
1	1,715039	2,43242	5,868939
2	1,64898	1,5	4,841946
3	1,985285	0,5	3,574602
4	1,078694	2,217356	5,547772
5	1,223484	2,291288	5,607535
6	3,081706	1,707825	2,603417
7	0,62113	1,979057	5,547772
8	2,172357	1,258306	3,574602
9	1,268079	2,140872	5,725188
10	2,172357	1,118034	3,574602
11	1,985285	0,763763	3,574602
12	0,62113	1,892969	5,547772
13	3,936612	2,291288	1,666667
14	1,268079	2,140872	5,725188
15	1,268079	2,140872	5,725188
16	5,275438	3,774917	1,666667
17	6,240301	4,716991	1,855921
18	7,003747	5,283622	2,185813
19	1,508208	1,607275	4,630815
20	1,393326	2,217356	5,783117
21	1,393326	2,217356	5,783117
22	1,078694	2,140872	5,547772
23	1,078694	2,217356	5,547772
24	1,393326	2,217356	5,783117
25	7,989383	6,264982	2,962731
26	3,117553	3,730505	6,146363
27	1,078694	1,118034	4,558265
28	0,62113	1,979057	5,547772
29	2,34455	1,258306	3,929942
30	1,078694	2,217356	5,547772

Gambar 4. 56 jarak antara data dengan *Centroid* Iterasi 5

Pada gambar 4.56 merupakan jarak antara data dengan centroid di iterasi ke-5, pencarian jarak ini merupakan cara untuk menentukan data masuk kedalam cluster mana.

ID	JP	JK	Pend	Kec	Pek
1	1	1	1	1	1
4	2	2	2	1	3
5	1	2	2	1	3
7	2	2	2	1	2
9	2	1	2	1	1
12	2	1	2	1	2
14	2	1	2	1	1
15	2	1	2	1	1
19	1	2	2	2	3
20	1	1	2	1	1
21	1	1	2	1	1
22	2	1	2	1	3
23	2	2	2	1	3
24	1	1	2	1	1
26	2	2	3	1	5
27	2	2	2	2	2
28	2	2	2	1	2
30	2	2	2	1	3

Gambar 4. 57 kelompok data C1 di Iterasi ke-5

Pada gambar 4.57 dapat dilihat bahwa data yang termasuk pada cluster C1 di iterasi ke-5 ialah sebanyak 18 data.

Data yang termasuk dalam C2 (Iterasi ke-5)

ID	JP	JK	Pend	Kec	Pek
2	1	1	2	2	1
3	2	1	2	3	2
6	2	2	2	4	3
8	2	2	2	3	3
10	2	1	2	3	3
11	2	2	2	3	2
29	1	1	2	3	1

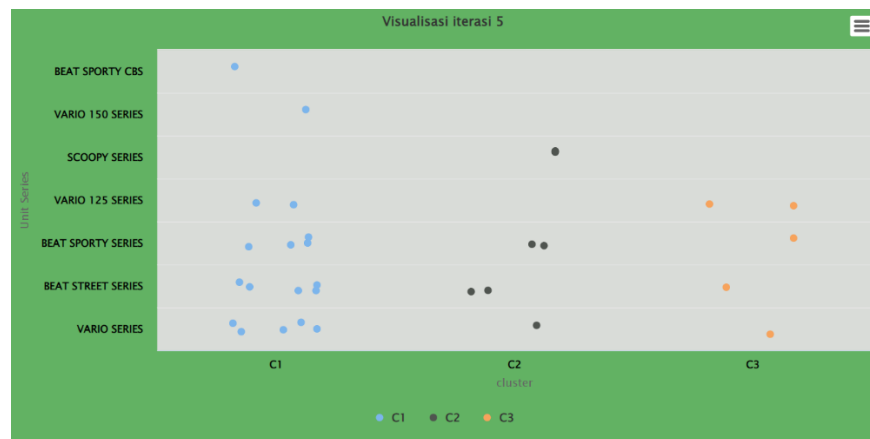
Gambar 4. 58 kelompok data C2 di Iterasi ke-5

Pada gambar 4.58 dapat dilihat bahwa data yang termasuk pada cluster C2 di iterasi ke-4 ialah sebanyak 7 data.

ID	JP	JK	Pend	Kec	Pek
13	2	2	2	5	2
16	2	1	2	6	4
17	1	2	2	7	4
18	2	1	2	8	1
25	2	1	2	9	1

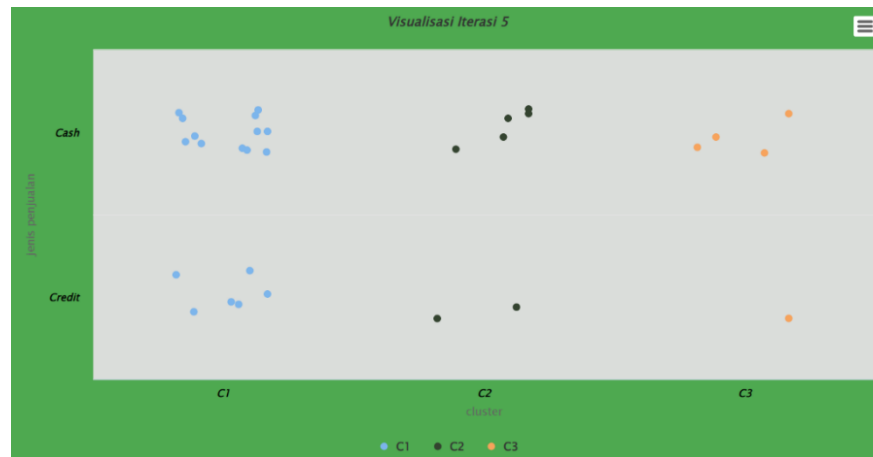
Gambar 4. 59 kelompok data C3 di Iterasi ke-5

Pada gambar 4.59 dapat dilihat bahwa data yang termasuk pada cluster C3 di iterasi ke-5 ialah sebanyak 5 data.



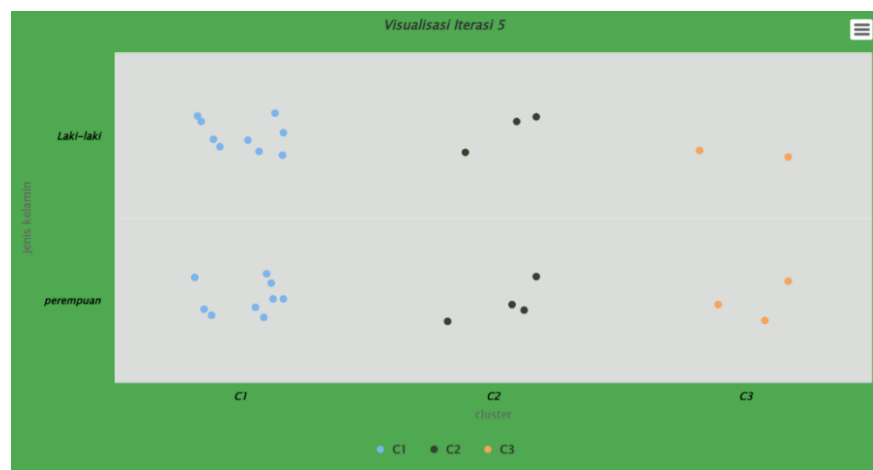
Gambar 4. 60 Visualisasi Unit Series Iterasi 5

Pada gambar 4.60 dapat dilihat bahwa data pada Atribut Unit Series tersebar cukup merata di kelompok C1 dengan jumlah 60% data, C2 dengan jumlah 23.33% data, dan C3 dengan jumlah 16.66% data.



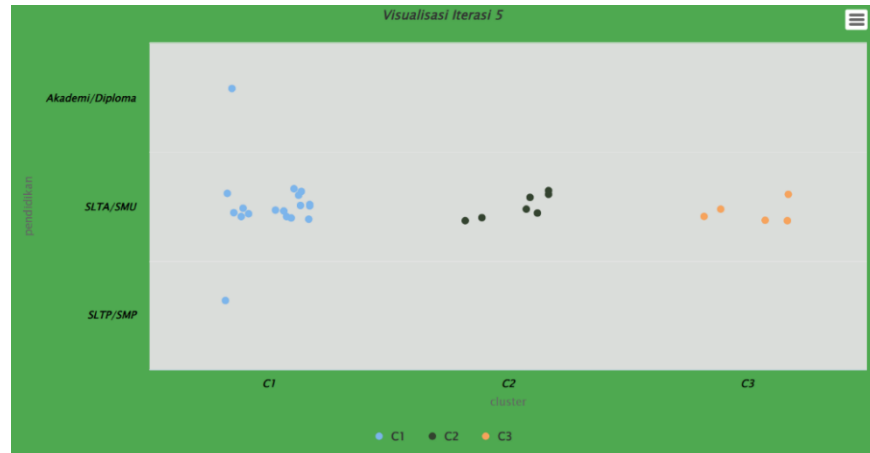
Gambar 4. 61 Visualisasi Jenis Penjualan Iterasi 5

Pada gambar 4.61 dapat dilihat bahwa data pada Atribut jenis penjualan tersebar cukup merata di kelompok C1 dengan jumlah 60% data, C2 dengan jumlah 23.33% data, dan C3 dengan jumlah 16.66% data.



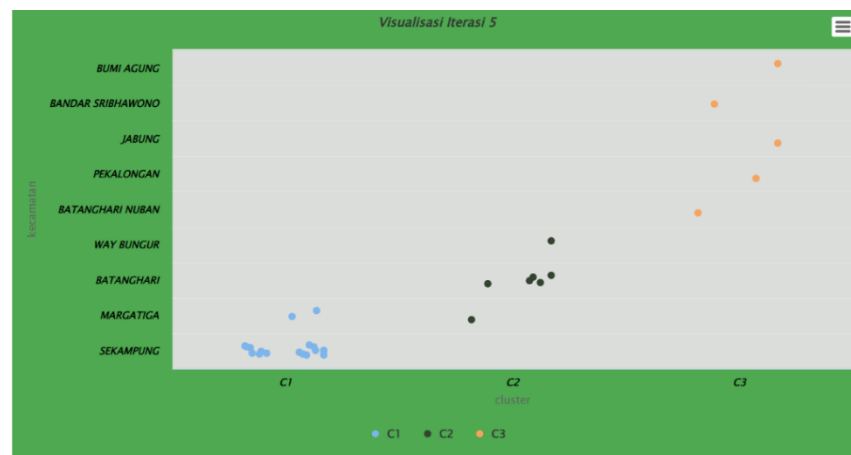
Gambar 4. 62 Visualisasi jenis Kelamin Iterasi 5

Pada gambar 4.62 dapat dilihat bahwa data pada Atribut jenis kelamin tersebar cukup merata di kelompok C1 dengan jumlah 60% data, C2 dengan jumlah 23.33% data, dan C3 dengan jumlah 16.66% data.



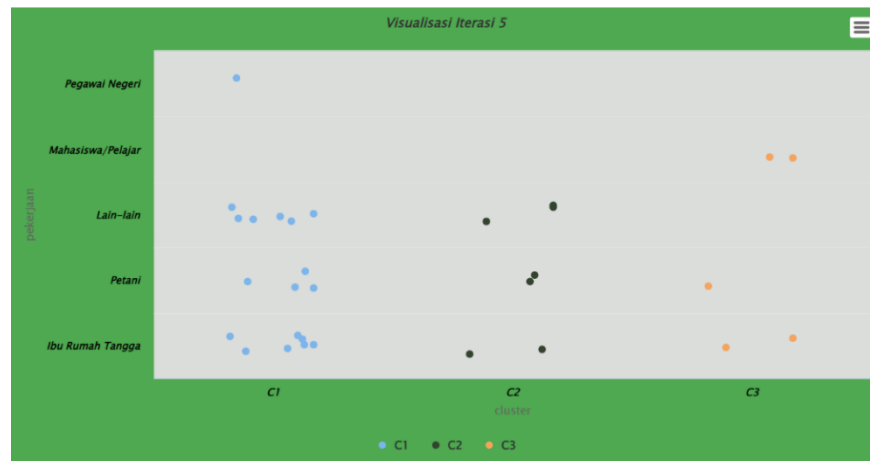
Gambar 4. 63 Visualisasi Pendidikan Iterasi 5

Pada gambar 4.63 dapat dilihat bahwa data pada Atribut pendidikan tersebar cukup merata pada C1 60%, C2 23.33%, dan C3 16.66% berdasarkan SLTA/SMU.



Gambar 4. 64 Visualisasi Kecamatan Iterasi 5

Pada gambar 4.64 dapat dilihat bahwa data pada Atribut kecamatan tersebar cukup merata di kelompok C1 60%, C2 23.33%, dan C3 masing masing wilayah tersebar 1 data dengan 16.66%.



Gambar 4. 65 Visualisasi Pekerjaan Iterasi 5

Pada gambar 4.65 dapat dilihat bahwa data pada Atribut kecamatan tersebar cukup merata di kelompok C1 60%, C2 23.33%, dan C3 dengan jumlah 16.66%.

Cluster	JP	JK	Pend	Kec	Pek
C1	1,666667	1,5	2	1,111111	2,111111
C2	1,714286	1,428571	2	3	2,142857
C3	1,8	1,4	2	7	2,4

Gambar 4. 66 Centroid baru di Iterasi ke-6

Pada gambar 4.66 menentukan centroid baru guna menentukan apakah masih ada perubahan atau tidak terhadap centroid sebelumnya, yaitu dengan mencari rata – rata dari masing-masing cluster terhadap masing-masing atribut.

Id	C1	C2	C3
1	1,715039	2,645751	6,305553
2	1,64898	1,732051	5,268776
3	1,985285	0,534522	4,04475
4	1,078694	2,267787	6,063003
5	1,223484	2,360387	6,112283
6	3,081706	1,46385	3,1241
7	0,62113	2,104417	6,046487
8	2,172357	1,069045	4,093898
9	1,268079	2,360387	6,177378
10	2,172357	1	4,069398
11	1,985285	0,654654	4,069398
12	0,62113	2,070197	6,029925
13	3,936612	2,104417	2,135416
14	1,268079	2,360387	6,177378
15	1,268079	2,360387	6,177378
16	5,275438	3,56571	1,939072
17	6,240301	4,503967	1,886796
18	7,003747	5,154748	1,777639
19	1,508208	1,603567	5,134199
20	1,393326	2,44949	6,225753
21	1,393326	2,44949	6,225753
22	1,078694	2,236068	6,046487
23	1,078694	2,267787	6,063003
24	1,393326	2,44949	6,225753
25	7,989383	6,129554	2,481935
26	3,117553	3,683942	6,645299
27	1,078694	1,195229	5,05569
28	0,62113	2,104417	6,046487
29	2,34455	1,414214	4,331282
30	1,078694	2,267787	6,063003

Gambar 4. 67 jarak antara data dengan *Centroid* di Iterasi ke-6

Pada gambar 4.67 merupakan jarak antara data dengan centroid di iterasi ke-6, pencarian jarak ini merupakan cara untuk menentukan data masuk kedalam cluster mana.

ID	JP	JK	Penc	Kec	Pek
1	1	1	1	1	1
2	1	1	2	2	1
4	2	2	2	1	3
5	1	2	2	1	3
7	2	2	2	1	2
9	2	1	2	1	1
12	2	1	2	1	2
14	2	1	2	1	1
15	2	1	2	1	1
19	1	2	2	2	3
20	1	1	2	1	1
21	1	1	2	1	1
22	2	1	2	1	3
23	2	2	2	1	3
24	1	1	2	1	1
26	2	2	3	1	5
27	2	2	2	2	2
28	2	2	2	1	2
30	2	2	2	1	3

Gambar 4. 68 kelompok data C1 di Iterasi ke-6

Pada gambar 4.68 dapat dilihat bahwa data yang termasuk pada cluster C1 di iterasi ke-6 ialah sebanyak 19 data.

ID	JP	JK	Penc	Kec	Pek
3	2	1	2	3	2
6	2	2	2	4	3
8	2	2	2	3	3
10	2	1	2	3	3
11	2	2	2	3	2
13	2	2	2	5	2
29	1	1	2	3	1

Gambar 4. 69 kelompok data C2 di Iterasi ke-6

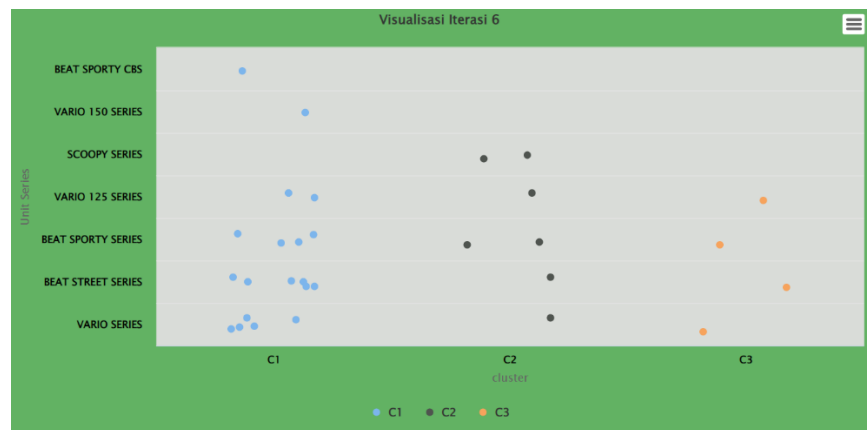
Pada gambar 4.69 dapat dilihat bahwa data yang termasuk pada cluster C2 di iterasi ke-6 ialah sebanyak 7 data.



ID	JP	JK	Penc	Kec	Pek
16	2	1	2	6	4
17	1	2	2	7	4
18	2	1	2	8	1
25	2	1	2	9	1

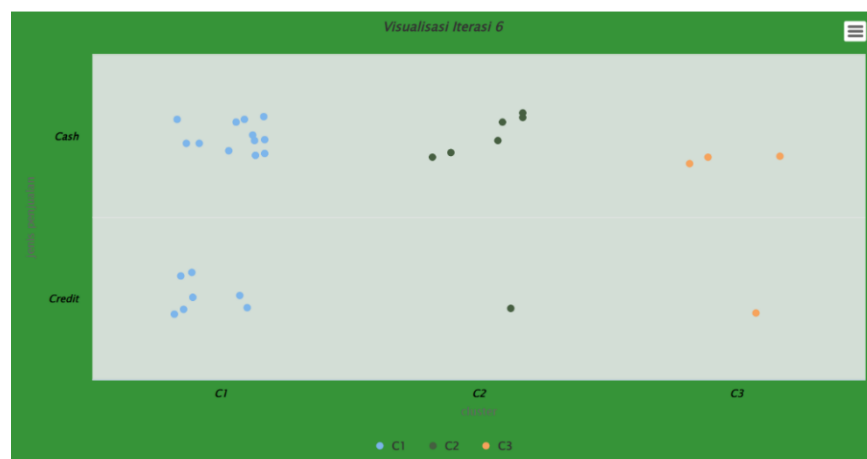
Gambar 4. 70 kelompok data C3 di Iterasi ke-6

Pada gambar 4.70 dapat dilihat bahwa data yang termasuk pada cluster C3 di iterasi ke-6 ialah sebanyak 4 data



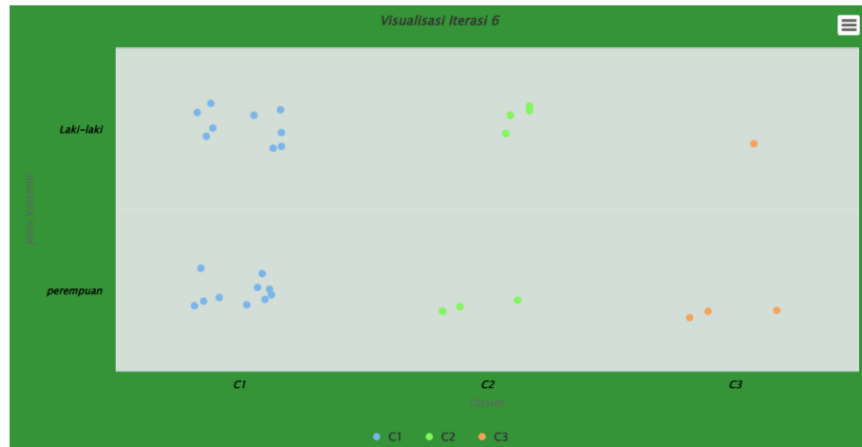
Gambar 4. 71 Visualisasi Unit Series Iterasi 6

Pada gambar 4.71 dapat dilihat bahwa data pada Atribut Unit Series tersebar cukup merata di kelompok C1 dengan jumlah 63.33% data, C2 dengan jumlah 23.3% data, dan C3 dengan jumlah 13.33% data.



Gambar 4. 72 Visualisasi Jenis Penjualan iterasi 6

Pada gambar 4.72 dapat dilihat bahwa data pada Atribut jenis penjualan tersebar cukup merata di kelompok C1 dengan jumlah 63.33% data, C2 dengan jumlah 23.3% data, dan C3 dengan jumlah 13.33% data.



Gambar 4. 73 Visualisasi Jenis Kelamin iterasi 6

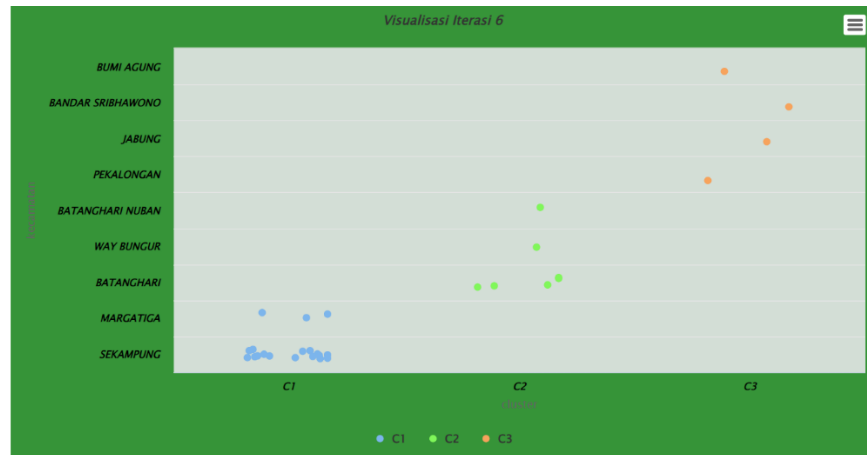
Pada gambar 4.73 dapat dilihat bahwa data pada Atribut jenis kelamin tersebar cukup merata di kelompok C1 dengan jumlah 63.33% data, C2 dengan jumlah 23.3% data, dan C3 dengan jumlah 13.33% data.



Gambar 4. 74 Visualisasi Pendidikan iterasi 6

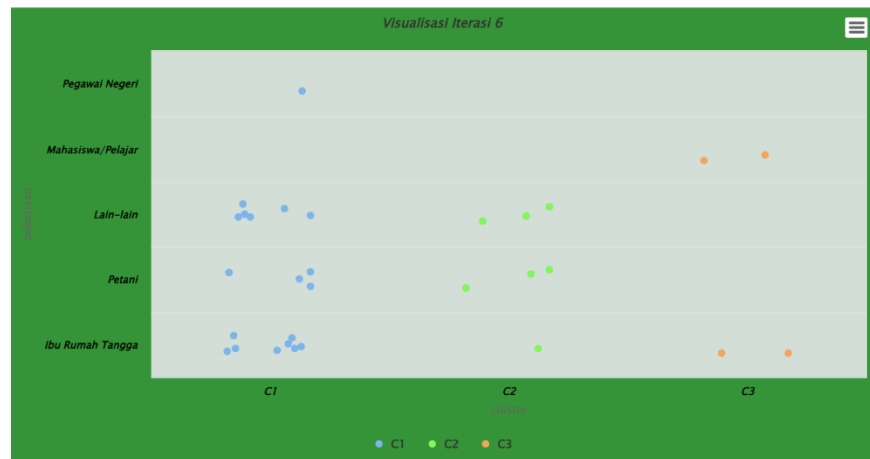
Pada gambar 4.74 dapat dilihat bahwa data pada Atribut pendidikan tersebar cukup merata di C1 dengan jumlah 63.33% data, C2 dengan

jumlah 23.3% data, dan C3 dengan jumlah 13.33% data pada Pendidikan SLTA/SMU.



Gambar 4. 75 Visualisasi Kecamatan iterasi 6

Pada gambar 4.75 dapat dilihat bahwa data pada Atribut kecamatan tersebar cukup merata di kelompok C1 dengan jumlah 63.33% data pada wilayah sekampung, C2 dengan jumlah 23.3% data, dan C3 dengan jumlah 13.33% data.



Gambar 4. 76 Visualisasi Pekerjaan iterasi 6

Pada gambar 4.76 dapat dilihat bahwa data pada Atribut pekerjaan tersebar cukup merata di kelompok C1 dengan jumlah 63.33% data, C2 dengan jumlah 23.3% data, dan C3 dengan jumlah 13.33% data.

Cluster	JP	JK	Pend	Kec	Pek
C1	1,631579	1,473684	2	1,157895	2,052632
C2	1,857143	1,571429	2	3,428571	2,285714
C3	1,75	1,25	2	7,5	2,5

Gambar 4. 77 *Centroid* baru di Iterasi ke-7

Pada gambar 4.77 menentukan *centroid* baru guna menentukan apakah masih ada perubahan atau tidak terhadap *centroid* sebelumnya, yaitu dengan mencari rata – rata dari masing-masing cluster terhadap masing-masing atribut.

Id	C1	C2	C3
1	1,660191	3,100362	6,791539
2	1,562192	2,18062	5,755432
3	1,9381	0,782461	4,541476
4	1,1555	2,571429	6,566963
5	1,264254	2,706756	6,604922
6	3,063954	1,020204	3,622844
7	0,663659	2,486699	6,566963
8	2,168779	0,947607	4,596194
9	1,221914	2,810331	6,680195
10	2,156611	1,020204	4,541476
11	1,951631	0,685119	4,596194
12	0,622745	2,51526	6,528782
13	3,895804	1,65985	2,66927
14	1,221914	2,810331	6,680195
15	1,221914	2,810331	6,680195
16	5,253412	3,146102	2,150581
17	6,212756	4,075812	1,903943
18	6,948564	4,785181	1,620185
19	1,510811	1,862629	5,623611
20	1,325229	2,934663	6,717514
21	1,325229	2,934663	6,717514
22	1,132497	2,599058	6,528782
23	1,1555	2,571429	6,566963
24	1,325229	2,934663	6,717514
25	7,93516	5,748114	2,150581
26	3,181926	3,803865	7,079901
27	1,060497	1,525297	5,578978
28	0,663659	2,486699	6,566963
29	2,26377	1,702339	4,808846
30	1,1555	2,571429	6,566963

Gambar 4. 78 jarak antara data dengan *Centroid* di Iterasi ke-7

Pada gambar 4.78 merupakan jarak antara data dengan centroid di iterasi ke-7, pencarian jarak ini merupakan cara untuk menentukan data masuk kedalam cluster mana.

ID	JP	JK	Pend	Kec	Pek
1	1	1	1	1	1
2	1	1	2	2	1
4	2	2	2	1	3
5	1	2	2	1	3
7	2	2	2	1	2
9	2	1	2	1	1
12	2	1	2	1	2
14	2	1	2	1	1
15	2	1	2	1	1
19	1	2	2	2	3
20	1	1	2	1	1
21	1	1	2	1	1
22	2	1	2	1	3
23	2	2	2	1	3
24	1	1	2	1	1
26	2	2	3	1	5
27	2	2	2	2	2
28	2	2	2	1	2
30	2	2	2	1	3

Gambar 4. 79 kelompok data C1 di Iterasi ke-7

Pada gambar 4.79 dapat dilihat bahwa data yang termasuk pada cluster C1 di iterasi ke-7 ialah sebanyak 19 data.

ID	JP	JK	Pend	Kec	Pek
3	2	1	2	3	2
6	2	2	2	4	3
8	2	2	2	3	3
10	2	1	2	3	3
11	2	2	2	3	2
13	2	2	2	5	2
29	1	1	2	3	1

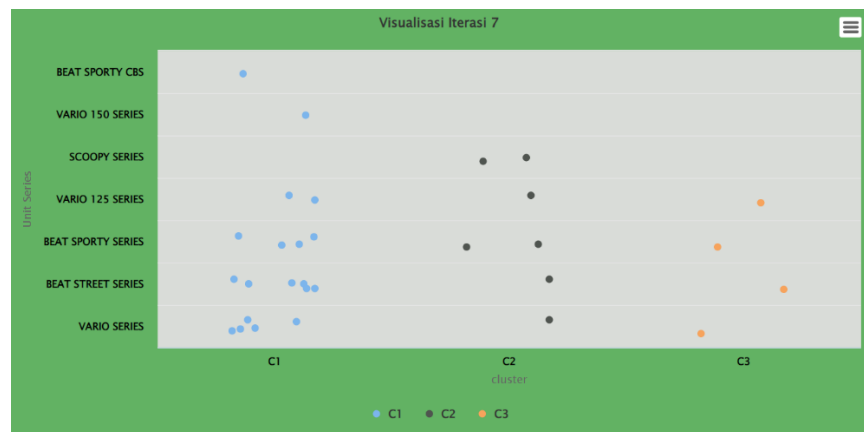
Gambar 4. 80 kelompok data C2 di Iterasi ke-7

Pada gambar 4.80 dapat dilihat bahwa data yang termasuk pada cluster C2 di iterasi ke-7 ialah sebanyak 7 data.

ID	JP	JK	Pend	Kec	Pek
16	2	1	2	6	4
17	1	2	2	7	4
18	2	1	2	8	1
25	2	1	2	9	1

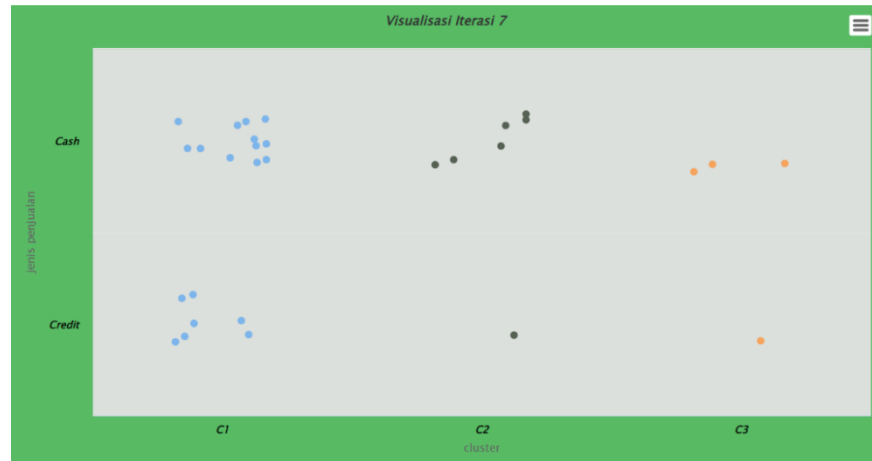
Gambar 4. 81 kelompok data C3 di Iterasi ke-7

Pada gambar 4.81 dapat dilihat bahwa data yang termasuk pada cluster C4 di iterasi ke-7 ialah sebanyak 4 data.



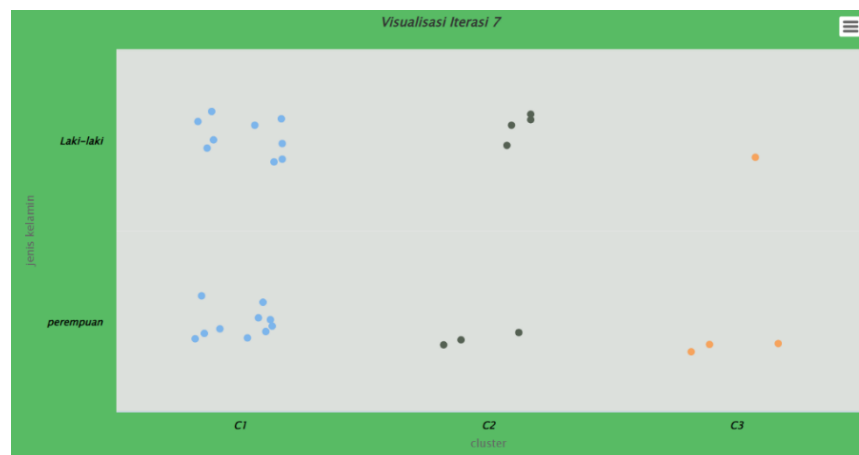
Gambar 4. 82 Visualisasi Unit Series iterasi 7

Pada gambar 4.82 dapat dilihat bahwa data pada Atribut Unit Series tersebar cukup merata di kelompok C1 dengan jumlah 63.3% data, C2 dengan jumlah 23.3% data, dan C3 dengan jumlah 13.3% data.



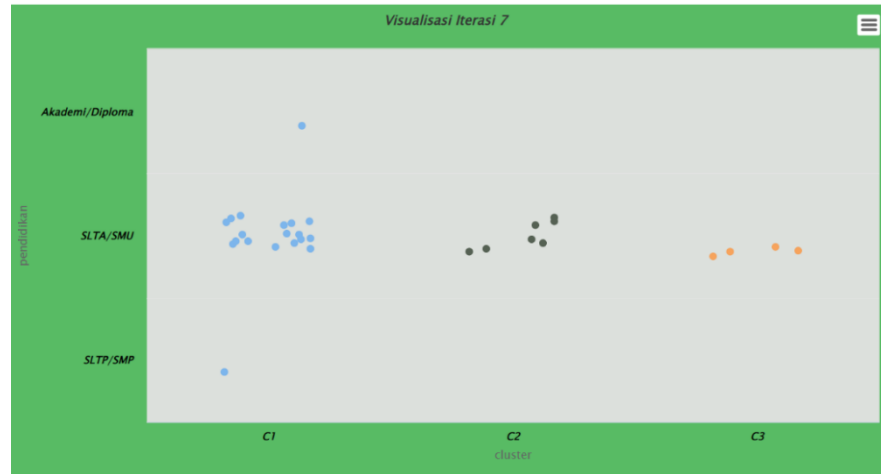
Gambar 4. 83 Visualisasi Jenis Penjualan iterasi 7

Pada gambar 4.83 dapat dilihat bahwa data pada Atribut jenis penjualan tersebar cukup merata di kelompok C1 dengan jumlah penjualan tersebar cukup merata di kelompok C1 dengan jumlah 63.33% data, C2 dengan jumlah 23.3% data, dan C3 dengan jumlah 13.33% data.



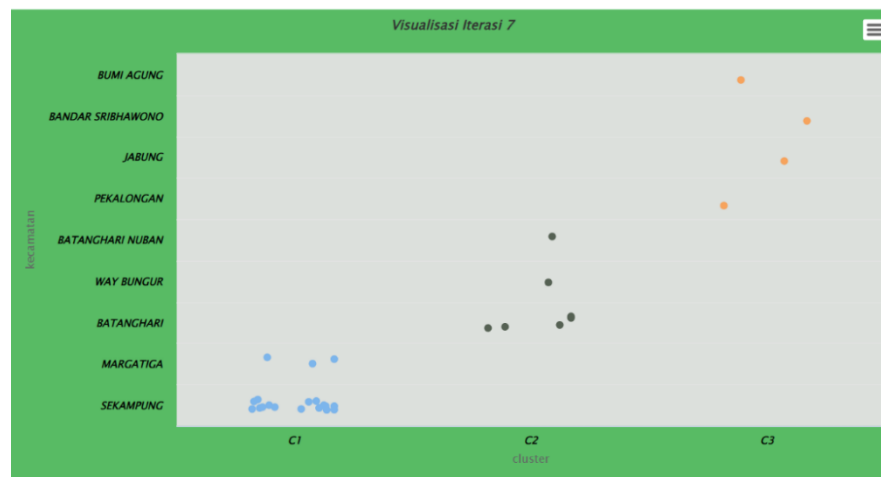
Gambar 4. 84 Visualisasi Jenis Kelamin iterasi 7

Pada gambar 4.84 dapat dilihat bahwa data pada Atribut jenis kelamin pada kelompok C1 dengan jumlah 63.33% data, C2 dengan jumlah 23.3% data, dan C3 dengan jumlah 13.33% data.



Gambar 4. 85 Visualisasi Pendidikan iterasi 7

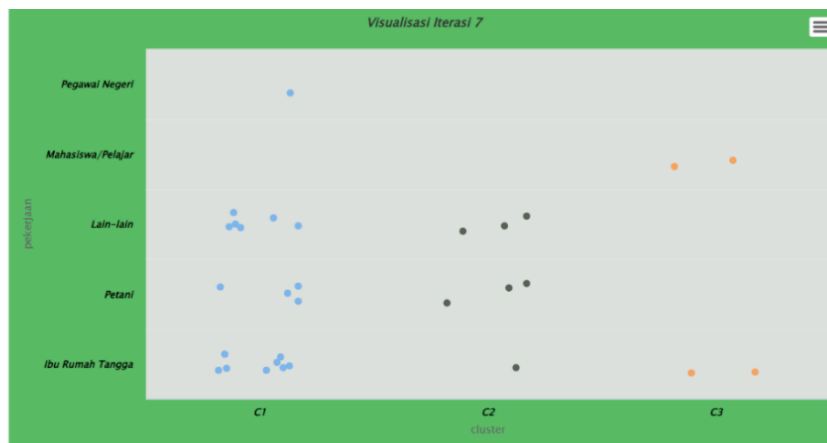
Pada gambar 4.85 dapat dilihat bahwa data pada Atribut pendidikan tersebar cukup merata di kelompok C1 dengan jumlah 63.33% data, C2 dengan jumlah 23.3% data, dan C3 dengan jumlah 13.33% data dengan Pendidikan SLTA/SMU.



Gambar 4. 86 Visualisasi Kecamatan iterasi 7

Pada gambar 4.86 dapat dilihat bahwa data pada Atribut kecamatan tersebar cukup merata di kelompok C1 pada wilayah sekampung dengan jumlah 63.33% data, C2 dengan jumlah 23.3% data, dan C3 dengan jumlah 13.33% data.



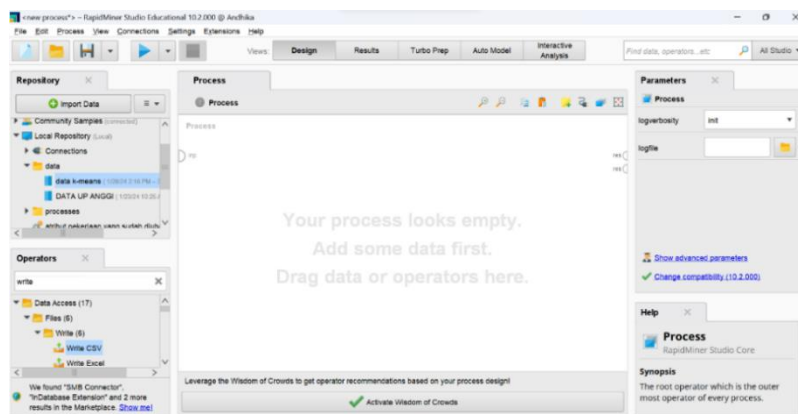


Gambar 4. 87 Visualisasi Pekerjaan iterasi 7

Pada gambar 4.87 dapat dilihat bahwa data pada Atribut pekerjaan tersebar pada kelompok C1 dengan jumlah 63.33% data, C2 dengan jumlah 23.33% data, dan C3 dengan jumlah 13.33% data.

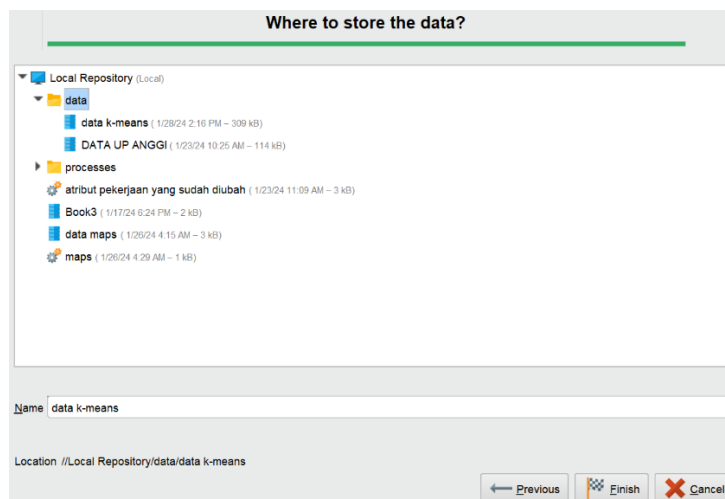
#### 4.6 Pengujian *software Rapid Miner*

Atribut yang digunakan sebagai segmentasi geografi dan segmentasi psikografi adalah Unit Series, jenis penjualan, jenis kelamin, Pendidikan, kecamatan, dan pekerjaan berdasarkan Tingkat penjualan yang tertinggi, sedang dan terendah. Data tersebut dibuat dalam data set baru dengan format .xlsx (*excel*) agar dapat dianalisis, *software* yang dipakai adalah *rapid miner studio* pada gambar 4.86 sebagai berikut.



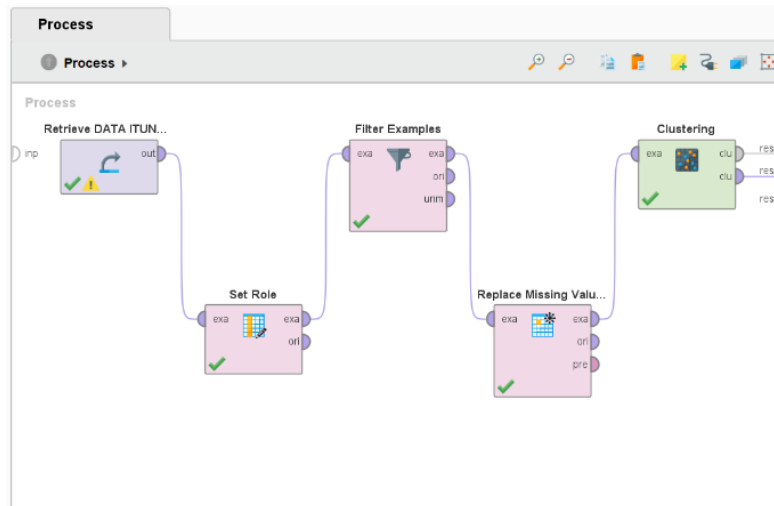
Gambar 4. 88 *Rapid Miner Studio*

Proses memasukan data excel dengan meng-klik finish agar excel tersimpan didalam file tools rapid miner dan selanjutnya Tarik data kedalam blank process dan dilanjutkan dengan menghubungkan operator yang akan digunakan.



Gambar 4. 89 Open File Excel

Setelah menentukan operator yang akan digunakan, lakukan drag and drop operators seperti pada gambar 4.89 dibawah, kemudian *open file data set operator read excel*, Setelah *open file excel* langkah selanjutnya adalah menghubungkan operator *read excel*, set role untuk menentukan label, filter examples untuk memfilter data yang tidak digunakan, replace missing value untuk mengisi data kosong pada setiap variabel, dan operator *clustering k-means* untuk menghitung data dan mengelompokkan data.



Gambar 4. 90 menghubungkan operator

Parameters dibawah adalah untuk menentukan Cluster yang telah ditentukan, setelah menentukan clusternya kemudian klik “Run”, dan muncul cluster model seperti gambar 4.90.

Parameters	
<b>Clustering (k-Means)</b>	
<input checked="" type="checkbox"/>	add cluster attribute
<input type="checkbox"/>	add as label
<input type="checkbox"/>	remove unlabeled
k	3
max runs	10
<input checked="" type="checkbox"/>	determine good start values
measure types	MixedMeasures
mixed measure	MixedEuclideanDistance
max optimization steps	100

Gambar 4. 91 Menentukan *Cluster*

## Cluster Model

Cluster 0: 10 items  
 Cluster 1: 5 items  
 Cluster 2: 15 items  
 Total number of items: 30

Gambar 4. 92 Hasil *Cluster*

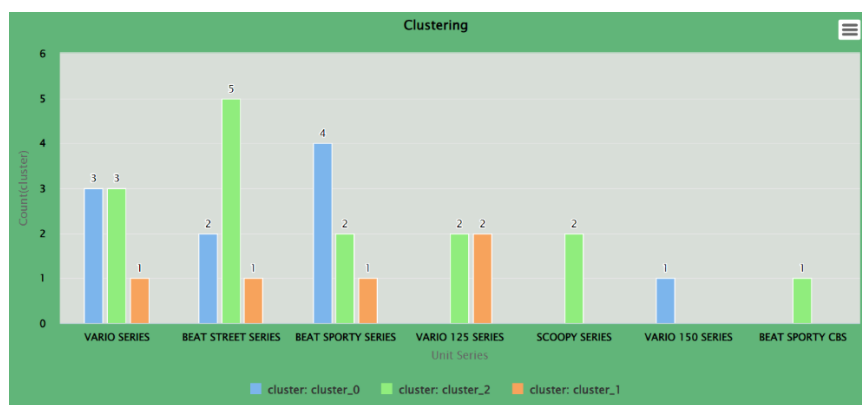
Dengan menentukan *cluster* pada data-data yang digunakan untuk mencari kelompok yang sejenis dan dapat terlihat jarak *Centroid* antar *cluster*.

Attribute	cluster_0	cluster_1	cluster_2
Kode Jenis Penjualan	1.400	1.800	1.967
Kode Jenis Kelamin	1	1.400	1.800
Kode pendidikan	1.900	2	2.067
Kode kecamatan	1.300	7	1.967
Kode pekerjaan	1.100	2.400	2.800

Gambar 4. 93 *Centroid k-means* Dalam Implementasi *Rapid Miner*

Gambar diatas adalah jarak *Centroid* Cluster1, 2 dan Cluster 3. Dengan masing-masing Atribut yang telah ditentukan.

Visualisasi data hubungan antara cluster dengan pekerjaan. Cluster tertinggi ada pada kode no 1 yaitu ibu rumah tangga dengan jumlah 11, sedang ada pada kode no 3 dengan jumlah 9 dan terendah ada pada kode no 5 yaitu pegawai negeri dengan jumlah 1 dari 30 sampel data.



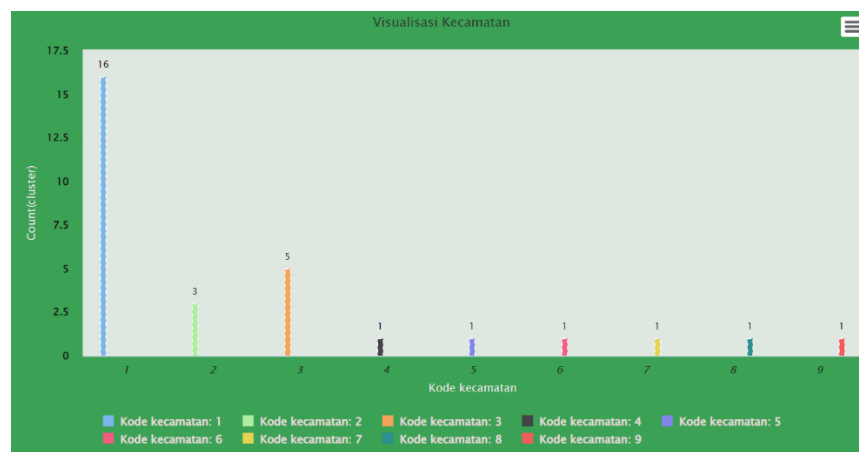
Gambar 4. 94 Visualisasi Unit Series

Visualisasi Data hubungan antara cluster dengan Unit Series penjualan tertinggi ada pada C2 Beat Street Series, C1 Beat Sporty Series, dan dari 30 sampel data lainnya menunjukkan sama rata dalam tiap Clusternya.



Gambar 4. 95 Visualisasi Pendidikan

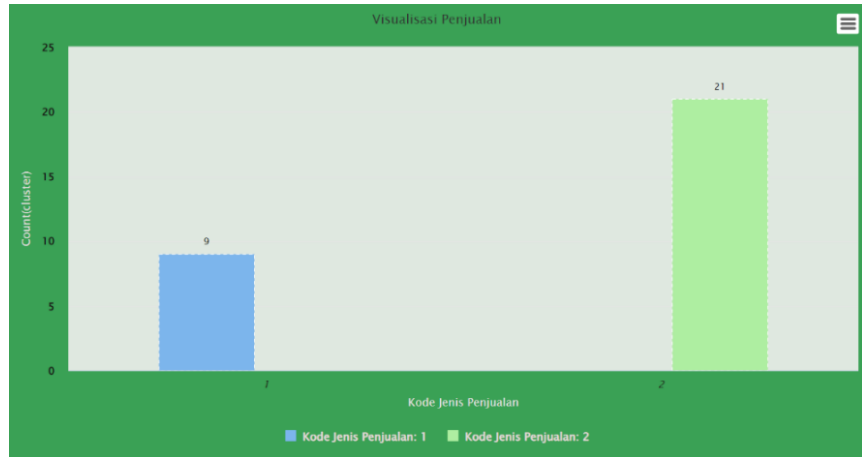
Visualisasi data hubungan antara cluster dengan Pendidikan. Tertinggi ada padaa kode no 2 SLTA/SMU dengan jumlah 28 dan teredah ada pada kode no 1 SLTP/SMP dan no 3 Akademi/ diploma dengan jumlah 1 dari 30 sampel data.



Gambar 4. 96 Visualisasi Kecamatan

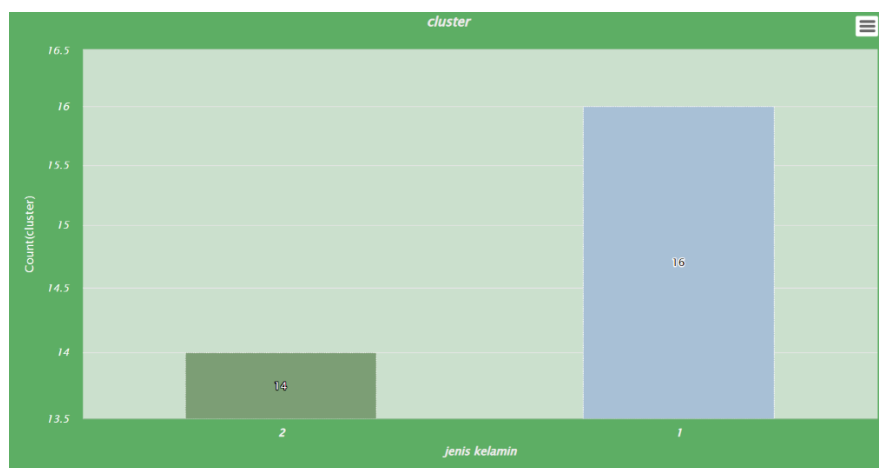
Visualisasi data hubungan antara cluster dengan kecamatan. Bagian tertinggi ada pada kode no 1 wilayah sekampung dengan jumlah 16, bagian sedang ada pada kode no 3 batanghari dengan jummlah 5 dan

kode no 2 margatiga dengan jumlah 3, dan terendah ada pada wilayah bagian way bungur, Batanghari nuban, pekalongan, jabung, bandar sribawono, dan bumi angung dengan jumlah 1.



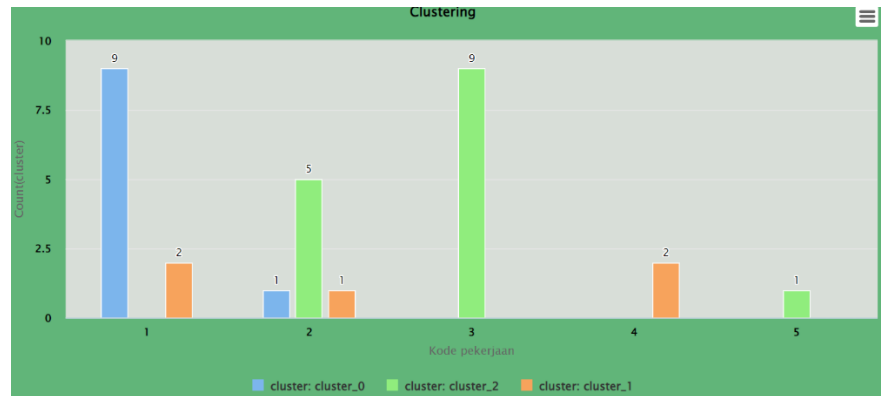
Gambar 4. 97 Visualisasi Jenis Penjualan

Visualisasi data hubungan antara cluster dengan jenis penjualan, yang tertinggi ada pada kode no 2 yaitu 21 dan terendah ada pada kode no 1 yaitu 9.



Gambar 4. 98 Visualisasi Jenis Kelamin

Visualisasi data hubungan antara cluster dengan jenis kelamin, yang tertinggi ada pada kode no 1 yaitu 16 dan terendah ada pada kode no 2 yaitu 14.



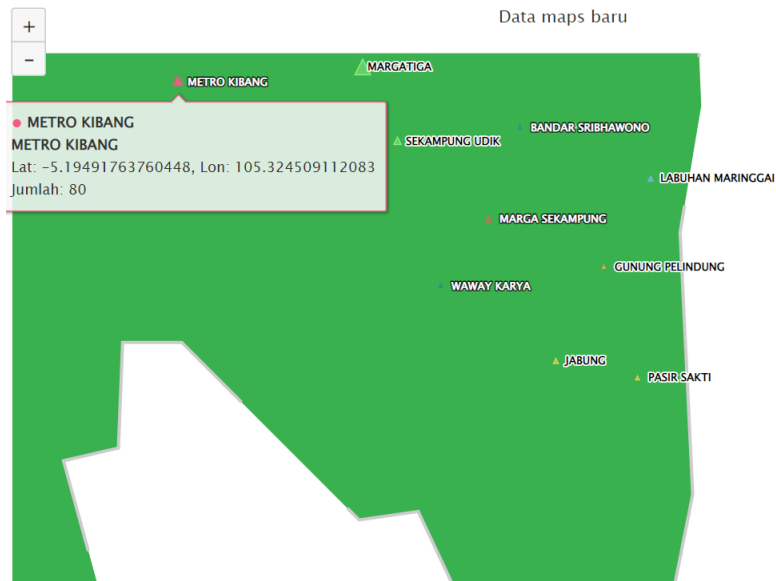
Gambar 4. 99 Visualisasi Pekerjaan

Visualisasi data hubungan antara cluster dengan pekerjaan, menunjukkan Tingkat tertinggi ada pada kode no 1 dan 3 dengan jumlah 9 data, tingkat sedan pada kode no 2 C2 dengan jumlah 5 dan terendah ada pada kode no 5 dari 30 sampel data.

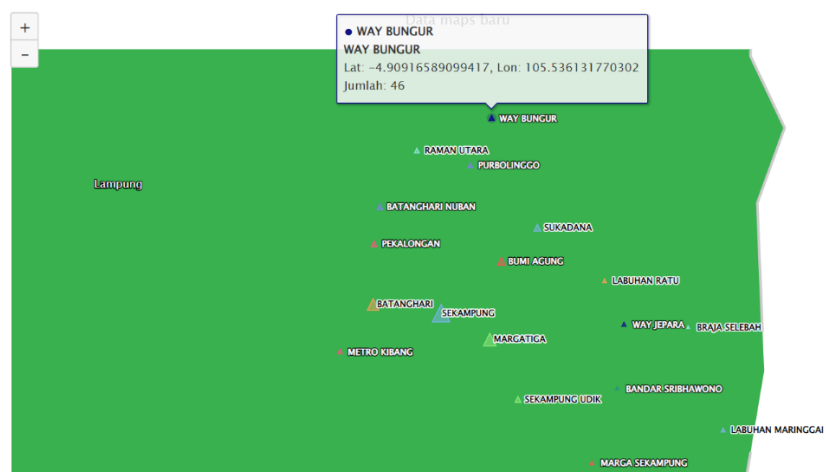
Gambar dibawah menentukan Longitude dan Latitude dengan menggunakan *Gmaps* agar dapat mengetahui titik wilayah masing masing kecamatan dari Kab. Lampung Timur. Pada gambar dibawah adalah hasil longitude dan latitude yang telah berbentuk *Maps*.

kecamatan	Latitude	longitude	Jumlah
SEKAMPUNG	-5.15436045	105.4562479	3047
MARGATIGA	-5.19209849	105.5165857	1315
BATANGHARI	-5.13771496	105.3698941	1179
WAY BUNGUR	-4.90916589	105.5361318	46
BATANGHARI NUBAN	-5.01346923	105.3864258	146
PEKALONGAN	-5.06039402	105.3766854	125
JABUNG	-5.50937939	105.6961907	14
BANDAR SRIBHAWONO	-5.26478656	105.6749708	12
BUMI AGUNG	-5.09293955	105.5374417	432
RAMAN UTARA	-4.94465679	105.4388	79
SUKADANA	-5.05302861	105.5865152	216
SEKAMPUNG UDIK	-5.27084009	105.547098	191
LABUHAN RATU	-5.12620119	105.6668155	26
WAY JEPARA	-5.18347456	105.6885861	23
PURBOLINGGO	-4.96769537	105.5059232	81
METRO KIBANG	-5.19491764	105.3245091	80
PASIR SAKTI	-5.53239247	105.779584	4
WAWAY KARYA	-5.42385812	105.5822492	3
MARGA SEKAMPUNG	-5.35825072	105.6366369	17
BRAJA SELEBAH	-5.19236748	105.7706806	13
LABUHAN MARINGGAI	-5.32665037	105.8062109	22
MATARAM BARU	-5.26709443	105.7313043	6
GUNUNG PELINDUNG	-5.41511789	105.7522635	1

Gambar 4. 100 Data Visualisasi *Maps*



Gambar 4. 101 Kecamatan Berdasarkan Visualisasi Maps



Gambar 4. 102 Kecamatan Berdasarkan Visualisasi Maps

Dari kedua visualisasi diatas dapat dilihat daerah sebaran penjualan pada PT. Tunas Dwimatra di Kab. Lampung timur, daerah penjualan tertinggi yaitu ada pada kecamatan sekampung yaitu sebanyak 43.05%, sedangkan penjualan terendah ada pada daerah Waway Karya yaitu dengan 0.04% dan Gunung Pelindung sebanyak 0.01%. Daerah lainnya, seperti Margatiga sebanyak 18.58%, Batanghari sebanyak 16.66%, Bumi Agung sebanyak 6.10%, Sukadana sebanyak 3.05%, Sekampung Udik sebanyak 2.70%, Batanghari



Nuban sebanyak 2.06%, Pekalongan sebanyak 1.77%, Purbolingo 1.14%, Metro Kibang sebanyak 1.13%, Raman Utara sebanyak 1.12%, Way Bungur sebanyak 0.55%, Labuhan ratu sebanyak 0.37%, Way jepera sebanyak 0.32%, Labuhan maringgai sebanyak 0.31%, Marga sekampung sebanyak 0.24%, Jabung sebanyak 0.20%, Braja sebelah sebanyak 0.18%, Bandar Sribhawono sebanyak 0.17%, Mataram baru sebanyak 0.08%, dan Pasir sakti sebanyak 0.06%