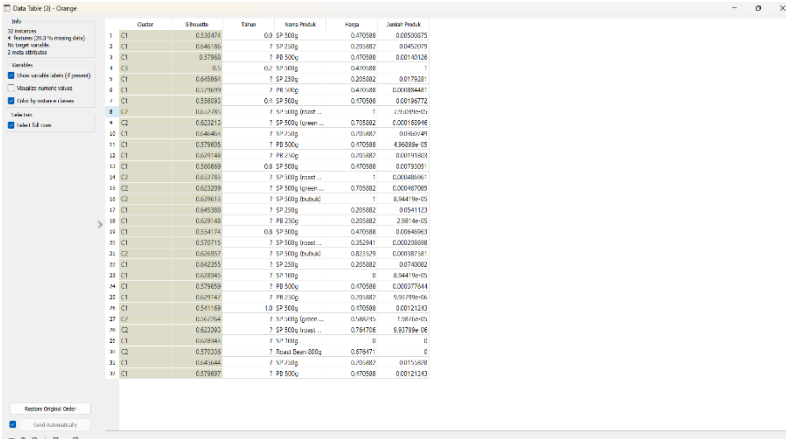


BAB V

PENGUJIAN DAN ANALISIS HASIL

Dataset penjualan kopi dikumpulkan dari KSU FA Masa Kabupaten Ngada antara tahun 2018 dan 2023. Program Orange memproses data dengan membaginya menjadi beberapa kluster dan menggunakan metode k-means.



Cluster	ID Barista	Tahun	Nama Produk	Harga	Jumlah Penjualan
1 C1	0.633472	0.3	SP 100g	0.470382	0.0003075
2 C1	0.646186	7	SP 250g	0.201882	0.0403071
3 C1	0.37966	7	PB 500g	0.470382	0.0014026
4 C1	0.6	0.2	SP 100g	0.461082	1
5 C1	0.645866	7	SP 250g	0.205802	0.0176281
6 C1	0.379699	7	PH 100g	0.461082	0.0004447
7 C1	0.656602	0.1	SP 500g	0.470382	0.0002672
8 C2	0.632172	7	SP 500g lyeven ...	0.735882	0.00018946
9 C2	0.646466	7	SP 250g	0.201882	0.0001491
10 C1	0.379655	7	PB 500g	0.470382	4.9689e-05
11 C1	0.639148	7	PH 250g	0.701882	0.00179459
12 C1	0.656669	0.8	SP 500g	0.470382	0.00730951
13 C1	0.637781	7	SP 100g lyeven ...	1	0.00048961
14 C2	0.632059	7	SP 500g lyeven ...	0.735882	0.000487095
15 C2	0.639633	7	SP 100g lyeven ...	1	0.04419e-05
16 C1	0.645388	7	SP 250g	0.205802	0.0541123
17 C1	0.639148	7	PB 250g	0.205802	2.3874e-05
18 C1	0.634134	0.8	SP 500g	0.470382	0.00040967
19 C1	0.379715	7	SP 100g lyeven ...	0.251941	0.000206688
20 C2	0.638957	7	SP 100g lyeven ...	0.823329	0.000887581
21 C1	0.642255	7	SP 250g	0.205802	0.01740662
22 C1	0.637846	7	SP 100g	0	0.04617e-05
23 C1	0.379659	7	PB 500g	0.470382	0.000275844
24 C1	0.639148	7	PH 250g	0.701882	0.00179459
25 C1	0.641159	1.8	SP 500g	0.470382	0.00121453
26 C2	0.632191	7	SP 500g lyeven ...	0.734704	5.93739e-06
27 C1	0.645466	4	SP 100g	0	0
28 C2	0.670236	7	Busui Daun 500g	0.676471	0
29 C1	0.646444	7	SP 250g	0.201882	0.01753029
30 C1	0.379697	7	PB 500g	0.470382	0.00121453

Gambar 5. 1 Mengelompokkan Data Berdasarkan *Cluster* Dengan Algoritma *K-Means* Menggunakan *Orange*

5.1 Sampel *Dataset* Latih

Tabel 5.1 menampilkan contoh dataset hasil penjualan kopi, yang meliputi parameter Nama Produk, Tahun, Harga, Jumlah Penjualan, Total Penjualan, dan Harga Total Penjualan.

Tahun	Nama Produk	Harga	Jumlah Produk	Total Harga Penjualan	Total Penjualan
2018	SP 500g	90,000	505	Rp 45,450,000.00	Rp 262,980,000.00
	SP 250g	45,000	4550	Rp 204,750,000.00	
	PB 500g	90,000	142	Rp 12,780,000.00	
2019	SP 500g	90,000	100625	Rp 9,056,250, 000.00	Rp 9,145,575,000.00
	SP 250g	45,000	1805	Rp 81,225,000.00	
	PB 500g	90,000	90	Rp 8,100,000.00	
2020	SP 500g	90,000	199	Rp 17,910,000.00	Rp 194,535,000.00
	SP 500g (roast bean)	180,000	9	Rp 1,620,000.00	
	SP 500g (green bean)	130,000	18	Rp 2,340,000.00	

	SP 250g	45,000	3631	Rp 163,395,000.00	
	PB 500g	90,000	6	Rp 540,000.00	
	PB 250g	45,000	194	Rp 8,730,000.00	
2021	SP 500g	90,000	799	Rp 71,910,000.00	Rp 334,200,000.00
	SP 500g (roast bean)	180,000	50	Rp 9,000,000.00	
	SP 500g (green bean)	130,000	48	Rp 6,240,000.00	
	SP 500g (bubuk)	180,000	10	Rp 1,800,000.00	
	SP 250g	45,000	5446	Rp 245,070,000.00	
	PB 250g	45,000	4	Rp 180,000.00	

2022	SP 500g	90,000	652	Rp 58,680,000.00	Rp 405,080,000.00
	SP 500g (roast bean)	70,000	22	Rp 1,540,000.00	
	SP 500g (bubuk)	150,000	40	Rp 6,000,000.00	
	SP 250g	45,000	7448	Rp 335,160,000.00	
	SP 100g	10,000	10	Rp 100,000.00	
	PB 500g	90,000	39	Rp 3,510,000.00	
	PB 250g	45,000	2	Rp 90,000.00	
2023	SP 500g	90,000	123	Rp 11,070,000.00	Rp 93,490,000.00
	SP 500g (green bean)	110,000	3	Rp 330,000.00	

SP 500g (roast bean)	140,000	2	Rp 280,000.00
SP 100g	10,000	1	Rp 10,000.00
Roast Bean 800g	125,000	1	Rp 125,000.00
SP 250g	45,000	1569	Rp 70,605,000.00
PB 500g	90,000	123	Rp 11,070,000.00

Tabel 5. 1 Data Penjualan Kopi Tahun 2018-2023

5.2 Sampel Dataset Uji

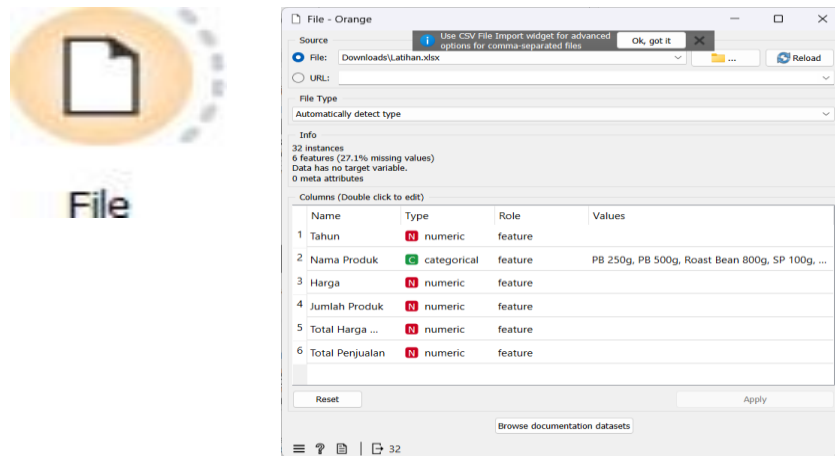
Sampel dataset pengujian digunakan untuk meramalkan cluster menggunakan algoritma k-means, seperti yang ditunjukkan pada Tabel 5.2 dan 5.3.

Tabel 5. 2 Data Uji *Cluster* Penjualan Kopi

<i>Cluster 1</i>	X	Y	<i>Cluster 2</i>	X	y	<i>Cluster 3</i>	x	y
SP 500g	450,000	102903	SP 250g	270,000	24449	SP 500g (bubuk)	330,000	50
SP 500g (roast bean)	570,000	83	PB 500g	450,000	400	SP 100g	20,000	11
SP 500g (green bean)	370,000	69	PB 250g	135,000	200	Roast Bean 800g	125,000	1

5.3 Proses *Import Data set*

Setelah mendapatkan data set, muat data set pelatihan dan data uji hasil penjualan kopi ke dalam aplikasi Orange melalui file widget. Tampilan file widget pada program Orange adalah sebagai berikut.



Gambar 5. 2 *Import Dataset* Latih dan Uji

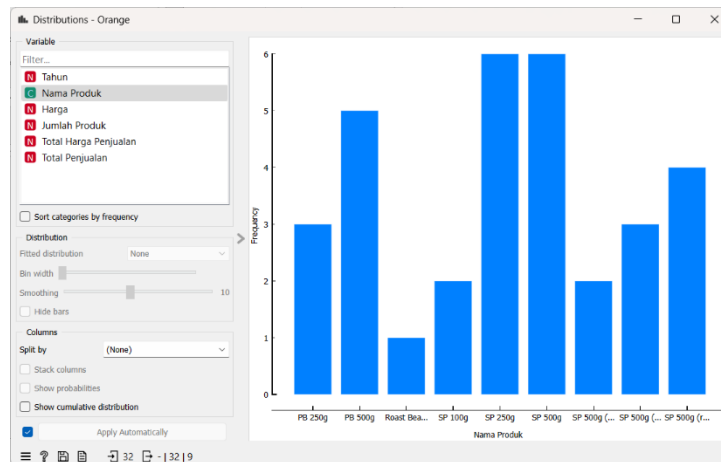
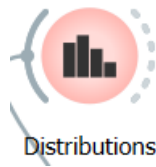
Widget file mengimpor file excel yang berisi data pelatihan dan data uji pertanian yang akan digunakan. Pada area source, Anda dapat memilih data set yang dikirimkan dalam bentuk file atau alamat URL; jika file, tentukan file excel yang akan digunakan. Data file excel yang dimasukkan akan diinformasikan, termasuk ukuran data set, serta jumlah dan jenis fitur data di excel. Atribut ditampilkan dalam bentuk kolom-kolom yang menjelaskan fungsi dan nilai atribut. Tabel tersebut berisi nama atribut, tipe data atau jenis variabel, role untuk mendefinisikan atribut, dan nilai yang ada pada atribut. Bagian tipe data pada tabel menunjukkan bahwa isi dari atribut yang ada merupakan tipe data numerik, kecuali nama produk yang merupakan tipe data kategorikal. Kemudian ada kolom role, yang memberikan atribut pada variabel yang menjadi target, fitur, dan meta data. Untuk data Tahun dimulai dari tahun 2018-2023, data Nama Produk berisi nama produk yang ada pada data

yang di-*import* seperti PB 250g, PB 500g, Roast Bean 800g, SP 100g, SP 250g, SP 500g, SP 500g (bubuk), SP 500g (green bean), SP 500g (roast bean), data harga merupakan harga penjualan setiap produk yang dijual, data jumlah produk merupakan jumlah produk yang terjual, data total harga penjumlahan merupakan data harga dikali sama jumlah produk terjual, data total penjualan merupakan penjumlahan dari total harga penjumlahan. Setelah semua data di-*import*, data tersebut akan ditampilkan dalam tabel dengan menggunakan *widget* data tabel. Seperti yang terlihat pada Gambar 5.3 di bawah ini.



Tahun	Nama Produk	Harga	Jumlah Produk	Total Harga Penjualan	Total Penjualan
2018	SP 500g	90000	505	45450000	262980000
?	SP 250g	45000	4550	204750000	?
?	PB 500g	90000	142	12780000	?
2019	SP 500g	90000	100625	9056250000	9145575000
?	SP 250g	45000	1805	81225000	?
?	PB 500g	90000	90	8100000	?
2020	SP 500g	90000	199	17910000	194535000
?	SP 500g (roast ...	180000	9	1620000	?
?	SP 500g (green ...	130000	18	2340000	?
?	SP 250g	45000	3631	163395000	?
?	PB 500g	90000	6	540000	?
?	PB 250g	45000	194	8730000	?
2021	SP 500g	90000	799	71910000	334200000
?	SP 500g (roast ...	180000	50	9000000	?
?	SP 500g (green ...	130000	48	6240000	?
?	SP 500g (bubuk)	180000	10	1800000	?
?	SP 250g	45000	5446	245070000	?
?	PB 250g	45000	4	180000	?
2022	SP 500g	90000	652	58680000	405180000
?	SP 500g (roast ...	70000	22	1540000	?
?	SP 500g (bubuk)	150000	40	6000000	?
?	SP 250g	45000	7448	335160000	?

Gambar 5. 3 Data Latih



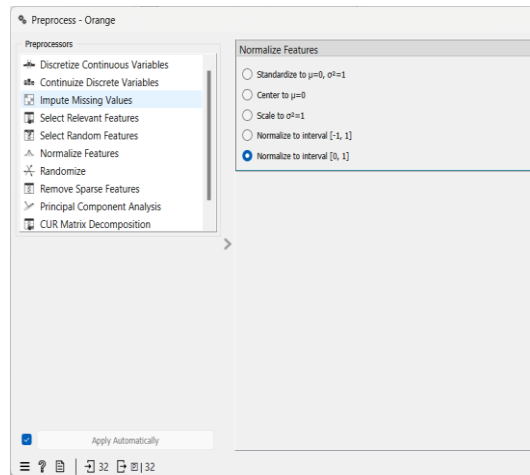
Gambar 5. 4 Distribusi Data Latih

Pada Gambar 5.3 di atas terlihat data tabel yang dihasilkan setelah kita meng-*import file* dari *file excel* yakni berupa tahun, nama produk, harga, jumlah produk, totalharga penjualan dan total penjualan. Semua data yang ditampilkan dalam data tabel ini adalah data asli dari *file excel* sebelum melalui tahap *preprocessing* data. Selain itu, kita juga bisa melihat hasil pendistribusian data menggunakan *widget distributions*. Yang dimana, *widget* ini berfungsi untuk menampilkan data hasil *import file excel*. Selain itu, pada *widget* ini juga kita bisa menampilkan data sesuai filter seperti data tahun, nama produk, harga, jumlah produk, total harga penjualan dan total penjualan.

5.4 *Pre-processing*

Persiapan data mengikuti tampilan data dalam widget tabel data. Pembersihan data terjadi pada tahap pra-pemrosesan. Pada tahap ini juga dilakukan normalisasi data dengan memilih Normalize to interval [0,1]. Cara ini dilakukan agar interval yang dihasilkan berkisar antara 0 hingga

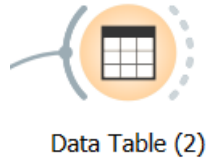
1. Seperti yang diilustrasikan pada Gambar 5.5 di bawah ini.



Gambar 5. 5 Proses *Pre-processing*

Data Hasil Penjualan Kopi

Gambar 5.6 menunjukkan hasil pra-pemrosesan data yang ditampilkan dalam widget tabel data. Hasilnya, atribut yang ditampilkan selama tahap pra-pemrosesan adalah atribut yang digunakan untuk tahap prapemrosesan data ini. Untuk memisahkan atribut-atribut tersebut, widget pilih kolom digunakan, yang juga memisahkan target.



Data Table (1) - Orange

Info
32 instances
5 features (27.1 % missing data)
No target variable.
No meta attributes.

Variables
 Show variable labels (if present)
 Visualize numeric values
 Color by instance classes
Selection
 Select full rows

Tahun	Nama Produk	Harga	Jumlah Produk	Total Harga Penjualan	Total Penjualan
1	0.0 SP 500g	0.470588	0.00500875	0.00501753	0.0187239
2	? SP 250g	0.205882	0.0452079	0.0226076	?
3	? PB 500g	0.470588	0.00140126	0.00141008	?
4	0.2 SP 500g	0.470588	1	1	1
5	? SP 250g	0.205882	0.0179281	0.00896785	?
6	? PB 500g	0.470588	0.00088481	0.000893307	?
7	0.4 SP 500g	0.470588	0.00196772	0.00197654	0.0111626
8	? SP 500g (roast ...	?	1	7.95039e-05	0.000177778
9	? SP 500g (green ...	0.705882	0.000168946	0.000257281	?
10	? SP 250g	0.205882	0.0360749	0.0180412	?
11	? PB 500g	0.470588	4.96899e-05	5.85232e-05	?
12	? PB 250g	0.205882	0.00191803	0.000962872	?
13	0.6 SP 500g	0.470588	0.00793051	0.00793928	0.0265917
14	? SP 500g (roast ...	?	1	0.000486961	0.000992686
15	? SP 500g (green ...	0.705882	0.000467085	0.000687923	?
16	? SP 500g (bubuk)	?	1	8.94419e-05	0.000197654
17	? SP 250g	0.205882	0.0541123	0.0270598	?
18	? PB 250g	0.205882	2.9814e-05	1.87716e-05	?
19	0.8 SP 500g	0.470588	0.00646963	0.00647841	0.034433
20	? SP 500g (roast ...	0.352941	0.000208698	0.000168944	?
21	? SP 500g (bubuk)	0.823529	0.000387581	0.000661422	?
22	? SP 250g	0.205882	0.0740082	0.0370076	?

Gambar 5. 6 Widget Data Tabel Hasil

Preprocessing Data Hasil Penjualan Kopi



Select Columns - Orange

Ignored (1)
Filter
Tahun

Features (5)
Filter
Nama Produk
Harga
Jumlah Produk
Total Harga Penjualan
Total Penjualan

Target
Metas

Reset Ignore new variables by default Send Automati

Gambar 5. 7 Widget Select Columns

Hasil *Preprocessing* Data Hasil

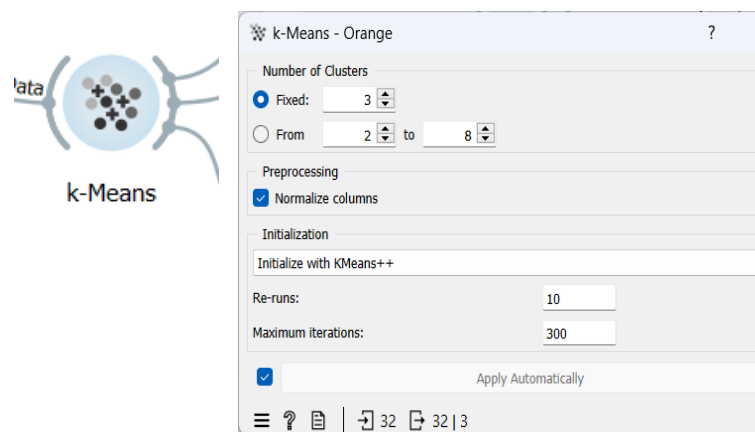
Penjualan Kopi

Untuk *wiget select columns* digunakan untuk memilih kolom mana yang ingin digunakan sebagai perhitungan algoritma *k-means*. Jika ada data

yang tidak ingin dihitung maka akan dimasukkan pada kotak *ignored* seperti data tahun yang tidak masuk dalam perhitungan algoritma *k-means*.

5.5 Algoritma *K-Means*

Penerapan algoritma k-means pada data set yang telah disiapkan dan penggunaan widget K-Means Orange untuk melakukan proses perhitungan menggunakan metode k-means untuk mendapatkan data hasil clustering. Pada tahap ini, perhitungan algoritma k-means dilakukan. Gambar 5.7 menunjukkan input jumlah cluster dari 2 hingga 8, normalisasi kolom, inisiasi dengan metode k-means, dan menjalankan kembali sebanyak 10 iterasi dengan maksimal 300 iterasi menggunakan widget k-means.



Gambar 5. 8 Widget *K-Means*

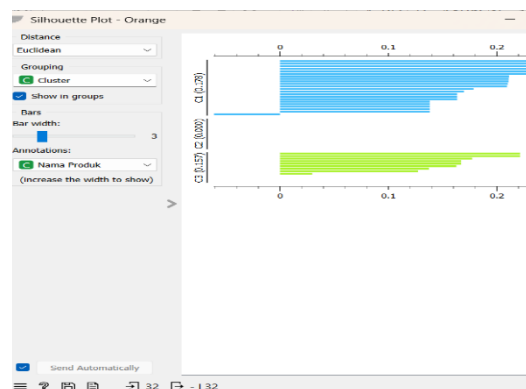
Data Table (3)

Cluster	Silhouette	Nama Produk	Harga	Jumlah Produk	Price Harga Pokok	Total Produk
1	0.538961	SP 500g	0.447082	0.0020085	0.0004751	0.0181099
1	0.531444	SP 750g	0.705687	0.5405779	0.0195206	?
1	0.538966	PK 500g	0.427058	0.07187136	0.00141308	?
4	0.5	SP 500g	0.470588	1	1	1
4	0.491123	SP 250g	0.205582	0.179381	0.00097155	?
6	0.538961	PK 500g	0.470588	0.00288481	0.00092307	?
7	0.538966	SP 500g	0.470588	0.00199772	0.00197854	0.0111628
8	0.481119	SP 250g Isolat	1	7.92339e-05	0.00177778	?
8	0.491123	SP 500g Isolat	0.705682	0.00718946	0.00257281	?
10	0.473071	PK 500g	0.470588	0.0000049	0.0180412	?
11	0.480919	PK 500g	0.447082	4.46899e-05	5.8152e-06	?
12	0.535276	PK 750g	0.705687	0.00191585	0.00000007	?
13	0.538968	SP 500g	0.470588	0.00290051	0.00798058	0.005917
14	0.531471	SP 250g Isolat	1	0.00000081	0.00000086	?
19	0.531471	SP 500g Isolat	0.705682	0.00040785	0.00000763	?
19	0.531471	SP 500g Isolat	1	0.94115e-05	0.00197854	?
17	0.531471	SP 250g	0.205582	0.0511123	0.0270598	?
18	0.531471	PK 250g	0.205582	2.8119e-05	1.87176e-05	?
18	0.538979	SP 500g	0.470588	0.00046983	0.00047841	0.004433
26	0.538971	SP 500g Isolat	0.532961	0.000328988	0.00116844	?
21	0.538948	SP 500g Isolat	0.532959	0.000391181	0.00001627	?
22	0.541991	SP 750g	0.705687	0.0040082	0.00000006	?
23	0.616215	PK 150g	0.470588	8.94479e-05	2.098e-05	?
24	0.538968	PK 500g	0.470588	0.000377644	0.000000478	?
25	0.531471	PK 500g	0.205582	0.00199e-06	0.00000e-06	?
26	0.538953	SP 500g	0.470588	0.00121343	0.00121306	0
17	0.531471	SP 500g Isolat	0.532959	1.9075e-05	3.5110e-05	?
18	0.531471	SP 250g Isolat	0.204706	5.92279e-05	2.92127e-05	?
18	0.531471	SP 150g	0	0	0	?
18	0.541628	Isolat Besar 800g	0.678411	0	1.2089e-05	?
21	0.538979	SP 250g	0.205582	0.0150028	0.00199188	?
22	0.538962	PK 500g	0.470588	0.00171245	0.00107156	?

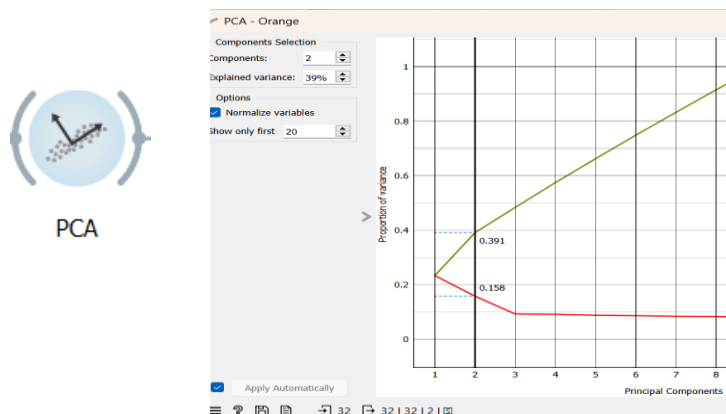
Gambar 5. 9 Widget Data Table

Penerapan algoritma k-means pada data set yang telah disiapkan dan penggunaan widget K-Means Orange untuk melakukan proses perhitungan menggunakan metode k-means untuk mendapatkan data hasil clustering. Pada tahap ini, perhitungan algoritma k-means dilakukan. Gambar 5.7 menunjukkan input jumlah cluster dari 2 hingga 8, normalisasi kolom, inisiasi dengan metode k-means, dan menjalankan kembali sebanyak 10 iterasi dengan maksimal 300 iterasi menggunakan widget k-means.

Silhouette Plot



Gambar 5. 10 Widget Silhoutte Plot



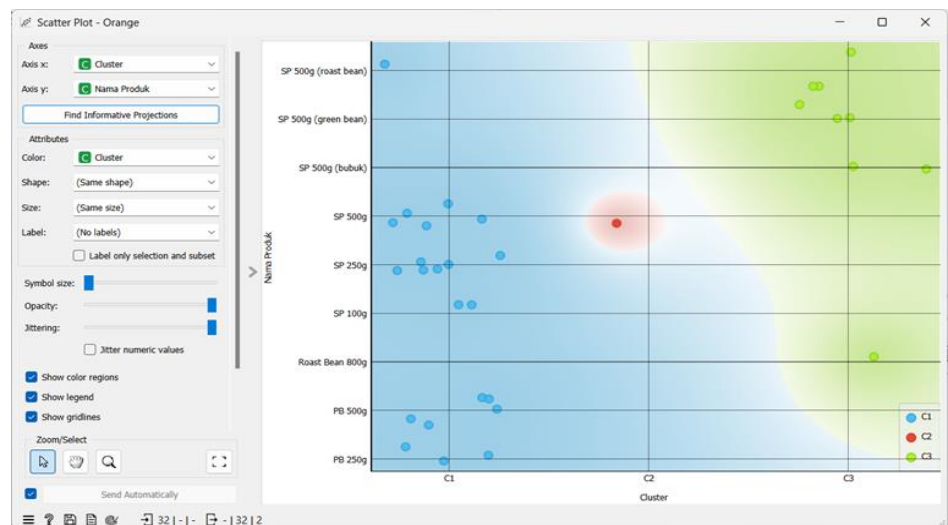
Gambar 5. 11 *Widget PCA*

Pada Gambar 5.11 di atas, menampilkan hasil visualisasi data yang telah disederhanakan menggunakan *widget PCA*. Yang dimana, pada nilai *component variance* didapatkan hasil 0.157 yang sama nilai dengan *cluster 3*, sedangkan pada nilai *cumulative variance* didapatkan hasil 0.178 yang sama nilai dengan *cluster 1*.

5.6 Analisis Hasil

Gambar 5.12 menggambarkan penerapan widget Scatter Plot untuk memvisualisasikan hasil cluster $c=3$ dari data penjualan produk KSU Fa Masa berdasarkan nama produk, yang dilakukan oleh KSU Fa Masa dengan menggunakan algoritma K-Means dan alat bantu orange. Pada *widget scatter plot* ini menjelaskan hasil *cluster* berdasarkan nama-nama produk yang di produksi oleh KSU Fa Masa yang terdiri dari 3 *cluster*, dalam setiap *cluster* ini terdapat beberapa macam produk yang mana antara *cluster* satu dengan yang lainnya itu berbeda. Dalam *cluster* ke-1 terdapat beberapa produk

diantaranya yaitu, PB 250g, PB 500g, SP 100g, SP 250g, SP 500g, SP 500g (*roast bean*). Kemudian dalam *cluster* ke-2 hanya terdapat satu jenis produk yaitu SP 500g. Kemudian dalam *cluster* ke-3 yaitu *Roast Bean* 800g, SP 500g (bubuk), SP 500g (*green bean*), SP 500g (*roast bean*).



Gambar 5. 12 Proses Pengujian Dengan Widget Scatter Plot

Berdasarkan hasil akhir penelitian mengenai data transaksi penjualan berdasarkan Gambar 5.9 berikut adalah rincian dan analisisnya rincian data transaksi penjualan berdasarkan *cluster*:

a) Cluster 1 C1

1. Jumlah transaksi : 21 transaksi
2. Transaksi terendah : Rp. 540.000
3. Transaksi tertinggi : Rp. 71.910.000

b) Cluster 2 (C2)

1. Jumlah transaksi penjualan: 9 transaksi

2. Transaksi tertinggi: Rp. 9.056.250

c) *Cluster 3 (C3)*

1. Jumlah transaksi: 1 transaksi

2. Transaksi terendah: Rp. 10.000

3. Transaksi penjualan tertinggi Rp.
335.160.000

Berdasarkan data-data tersebut dapat dilihat bahwa untuk data transaksi penjualan terendah terdapat pada *cluster 3 (3)* dengan nilai transaksi sebesar Rp. 10.000. Dan transaksi penjualan tertinggi terjadi pada *cluster 1 (1)* dengan nilai transaksi sebesar Rp. 71.910.000. *Cluster 1(C)* memiliki jumlah transaksi terbanyak dan juga mencakup transaksi penjualan tertinggi ini menunjukkan performa yang stabil dan tinggi dalam penjualan. *Cluster 2 (C2)* memiliki jumlah transaksi yang lebih sedikit dibanding *Cluster 1*. Transaksi terendah di *cluster* ini adalah Rp 9.056.250, yang menunjukkan bahwa meskipun jumlah transaksinya lebih sedikit, nilai transaksi terendahnya relatif tinggi. *Cluster 3 (3)* hanya memiliki satu transaksi, yang merupakan transaksi terendah sebesar Rp 10.000 dan transaksi tertinggi yang sangat signifikan sebesar Rp 335.160.000. Ini menunjukkan adanya anomali atau variasi yang sangat besar dalam data transaksi *cluster* ini.