

BAB IV HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Sumber data utama yang digunakan dalam penelitian ini adalah *database* Anecdotal Record SMK Walisongo Pecangaan Jepara dengan jumlah keseluruhan data ada 533 record Tahun Pelajaran 2019/2020. Data yang diterima dalam bentuk tabel pada *Microsoft Excel* sehingga mempermudah untuk dilakukan pembersihan data atau memfilter data, seperti pada gambar 4.1, adapun data selengkapnya terlampir dalam Lampiran 3. Data ini di dapatkan melalui perijinan kepada pihak SMK Walisongo Pecangaan Jepara.

NIS	NAMA	GENDER	KELAS	SIKAP PERILAKU																			
				A01	A02	A03	A04	A05	A06	A07	A08	A09	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20
1	2197 JAUHAROTUL FAJRI	P	X-KKBT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	2198 KHARISA EKA PUTRI ANGGRAINI	P	X-KKBT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	2199 NOVITA SEPTI FAULINA	P	X-KKBT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2200 PUTRI YULIA ISTIQOMAH	P	X-KKBT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	2201 QORI NANDA EKA PUTRI	P	X-KKBT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	2202 SATHI'ATUZ ZIDNI ARRIZQI	P	X-KKBT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	2203 SILVIA DESI SAFITRI	P	X-KKBT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	2204 WINDA FITRIANA SARI	P	X-KKBT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	2205 YELFI DINDA SEPTIYAN	P	X-KKBT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	2340 BUNGA NUR ISTIQOMAH	P	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	2341 DIDIK PRIYONO	L	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	2342 DINDA DAMAYANTI	P	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	2343 ELVIRA AYU RAHMAWATI	P	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	2344 GRESYANIA EKA ANANTA SALSAB	P	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	2345 IRMA SHOLEKHA	P	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	2346 KAYLA TUHASSIRONA	P	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	2347 LISTIYAFATUL FAIZAH	P	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	2348 M. IBNU ALFALAH	L	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	2349 MAZIYATUL ULAYA	P	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	2350 NIKEN PUSPITA AULIA	P	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	2351 RENI YUNITA	P	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	2352 BONIATU ARAHMANI	P	X-PS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Gambar 4.1. Data Anecdotal Record

4.2. Pre-processing Data

Data yang digunakan dalam penelitian ini adalah Data Siswa SMK Walisongo Pecangaan TP. 2019/2020 terdiri dari beberapa atribut antara lain NIS, Nama Siswa, Jenis Kelamin, Kelas dan Data Pelanggaran Siswa. Adapun atribut tersebut yaitu :

- a. NIS

Merupakan atribut yang berisi Nomor Induk Siswa di SMK Walisongo Pecangaan.

- b. Nama Siswa

Merupakan atribut yang berisi nama-nama yang tercantum di SMK Walisongo Pecangaan.

c. Jenis Kelamin

Merupakan atribut yang berisi jenis kelamin tiap-tiap siswa.

d. Kelas

Merupakan atribut yang berisi kelas-kelas siswa pada tahun pelajaran 2019/2020.

e. Pelanggaran Siswa

Merupakan atribut yang berisi pelanggaran-pelanggaran siswa pada tahun pelajaran 2019/2020.

4.2.1. Cleaning data

Pada tahapan ini dilakukan pembersihan data terhadap data yang memiliki *redundant*. Dalam *dataset* terdapat atribut NIS, Nama Siswa, Kelas dan Pelanggaran Siswa, atribut yang dihapus yaitu jenis kelamin dan Kelas. Hal itu dilakukan karena data tersebut tidak akan dipakai dalam perhitungan.

JO	NIS	NAMA	KERAJINAN												KERAPIAN															
			A16	A17	A18	A19	A20	A21	B01	B02	B03	B04	B05	B06	B07	B08	B09	B10	B11	B12	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10
242	2074	DIMAS ARIFIN ILHAM	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
243	2078	EKO LISTIAWAN	0	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
244	2080	FANYA TRISYAWATI	0	0	0	0	0	0	2	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
245	2081	FARAH DIAN PUSPITA	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
246	2086	ILHAM SUYADI	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
247	2088	KOMARIATUN AFIFA NURCAHYANI	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
248	2089	LUCKY ARDIMANSYAH	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
249	2091	LUTFIYATUL MALIKHAH	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
250	2092	M. AZKA TASLIMI CHOERI	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
251	2093	M. SALMAN FARISI FAISAL FAIS	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
252	2094	MAFTUH AHNAN	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
253	2095	MAFTUKHATU RIZQIYAH	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
254	2096	MAHMUD SUFA'AT	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
255	2097	MARIA ANA ELINA	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
256	2098	MARIA AUFA	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
257	2099	MAULA YUSUF WAFI YUDDIN	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
258	2100	MOHAMAD AZKA FAHMI	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
259	2101	MUHAMMAD ALFIN SYAIFUDIN	0	0	0	0	0	0	1	0	1	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
260	2113	MUHAMMAD UMAR FARUQ	0	0	0	0	0	1	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
261	2115	NIAM KHOLIDIN GHOFUR	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
262	2119	RAFLY YUSUF	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
263	2120	RIZKY HAKIM MAULANA	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Gambar 4.2. Hasil Cleaning Data

4.2.2. Pembobotan

Pada tahapan ini data akan dilakukan pembobotan sesuai dengan poinnya, dengan cara dikalikan jumlah kejadian dengan poin setelah itu di jumlahkan sehingga pada tahapan ini sudah diperoleh perhitungan nilai berdasarkan Anecdotal Record pada data pelanggaran.

Tabel 4.1. Data Bobot Pelanggaran

KODE	NAMA PELANGGARAN	POIN
A01	Berkuku Panjang (diberikan pewarna kuku)	2
A02	Mengganggu ketenangan Kegiatan Belajar Mengajar	2
A03	Kurang rasa setia kawan	2
A04	Mencoret-coret meja, kursi, dinding dan fasilitas sekolah lainnya	4
A05	Mengancam / mengintimidasi kawan	10
A06	Membawa / merokok di sekolah dan/atau diluar sekolah	10
A07	Bertindak tidak sopan kepada guru / karyawan	20
A08	Merusak sarana / prasarana sekolah	20
A09	Mengambil hak orang lain / mencuri / merampas	20
A10	Melompat pagar sekolah / jendela	25
A11	Bertindak tidak senonoh	30
A12	Berjudi / bermain kartu judi	30
A13	Membawa telepon seluler (HP), atau alat komunikasi lain	30
A14	Berkelahi di lingkungan sekolah	30
A15	Memalsukan tanda tangan orang lain / membuat surat ijin palsu	30
A16	Membawa / mengedarkan miras / narkoba	100
A17	Membawa senjata tajam, senjata api, VCD / buku porno	30
A18	Terlibat tawuran / perkelahian antar pelajar	30
A19	Berperilaku jorok / asusila	30
A20	Terlibat tindakan kriminal	50
A21	Berzina / hamil	100
B01	Datang terlambat <10 menit	1
B02	Datang terlambat <30 menit	2
B03	Datang Terlambag > 30 Menit	3
B04	Tidak mengikuti pelajaran tanpa ijin	4
B05	Tidak mengerjakan tugas Pekerjaan Rumah / Tugas Lain	3
B06	Tidak mengikuti kegiatan ekstra kurikuler	2
B07	Tidak masuk sekolah tanpa keterangan	4
B08	Meninggalkan kelas tanpa ijin (Bolos)	5
B09	Tidak mengikuti upacara	6
B10	Tidak mengikuti sholat dhuha	6
B11	Tidak mengikuti sholat dhuhur berjama'ah	6
B12	Tidak membawa mukena	3
C01	Tidak Memasukan Baju	1
C02	Tidak memakai kaos kaki / memakai tapi tidak sesuai ketentuan	2
C03	Tidak memakai ikat pinggang sekolah	2
C04	Atribut Tidak Lengkap	4
C05	Tidak memakai sepatu sesuai ketentuan	4
C06	Tidak memakai seragam sesuai ketentuan	5
C07	Berambut gondrong (bagi siswa putra)	5

C08	Memakai perhiasan berlebihan (bagi siswa putri)	5
C09	Bersolek berlebihan (bagi siswa putri)	10
C10	Rambut dicat / mewarnai rambut	20
C11	Telinga ditindik (bagi siswa putra)	30
C12	Memakai giwang / anting-anting (bagi siswa putra)	30
C13	Bertato	30
C14	Pakaian transparan (bagi siswa putri)	30

Tabel 4.2. Hasil Pembobotan

No	NIS	Nilai
1	2322	0
2	2323	9
3	2325	9
4	2326	8
5	2327	9
6	2358	5
7	2361	4
8	2328	15
9	2213	8
10	2206	8
11	2207	6
12	2208	11
13	2209	4
14	2210	10
15	2211	8
16	2212	4
17	2214	12
18	2215	8
19	2216	4
20	2217	9
....
533	2366	0

4.3. Proses *Clustering* dalam *K-Means*

Data yang telah melalui tahap *Pre-processing* selanjutnya akan dibagi menjadi dua cluster, sebelumnya kita tentukan centroid awal terlebih dahulu. Kita ambil data terkecil dan terbesar sebagai perhitungan pertama, *cluster 0* sebagai Siswa tidak bermasalah dan *cluster 1* sebagai siswa bermasalah. Kita

akan menggunakan persamaan *Euclidean Distance* untuk memperoleh jarak minimum data terhadap centroid.

Tabel 4.3. Nilai Centroid Awal

Cluster	x
C ₀	0
C ₁	18

Literasi Ke-1

Jarak antar data ke-1 dengan *centroid* Pertama (C₀)

$$d(x_1, c_0) = \sqrt{(0 - 0)^2} = 0$$

Jarak antar data ke-1 dengan *centroid* Kedua (C₁)

$$d(x_1, c_1) = \sqrt{(0 - 18)^2} = 18$$

Jarak antar data ke-2 dengan *centroid* Pertama (C₀)

$$d(x_2, c_0) = \sqrt{(9 - 0)^2} = 9$$

Jarak antar data ke-2 dengan *centroid* Kedua (C₁)

$$d(x_2, c_1) = \sqrt{(9 - 18)^2} = 9$$

Jarak antar data ke-3 dengan *centroid* Pertama (C₀)

$$d(x_3, c_0) = \sqrt{(9 - 0)^2} = 9$$

Jarak antar data ke-3 dengan *centroid* Kedua (C₁)

$$d(x_3, c_1) = \sqrt{(9 - 18)^2} = 9$$

Jarak antar data ke-4 dengan *centroid* Kedua (C₀)

$$d(x_4, c_0) = \sqrt{(8 - 0)^2} = 8$$

Jarak antar data ke-4 dengan *centroid* Pertama (C₁)

$$d(x_4, c_1) = \sqrt{(8 - 18)^2} = 10$$

Jarak antar data ke-5 dengan *centroid* Pertama (C_0)

$$d(x_5, c_0) = \sqrt{(9 - 0)^2} = 9$$

Jarak antar data ke-5 dengan *centroid* Kedua (C_1)

$$d(x_5, c_1) = \sqrt{(9 - 18)^2} = 9$$

Jarak antar data ke-6 dengan *centroid* Pertama (C_0)

$$d(x_6, c_0) = \sqrt{(5 - 0)^2} = 5$$

Jarak antar data ke-6 dengan *centroid* Kedua (C_1)

$$d(x_6, c_1) = \sqrt{(5 - 18)^2} = 13$$

Jarak antar data ke-7 dengan *centroid* Pertama (C_0)

$$d(x_7, c_0) = \sqrt{(4 - 0)^2} = 4$$

Jarak antar data ke-7 dengan *centroid* Kedua (C_1)

$$d(x_7, c_1) = \sqrt{(4 - 18)^2} = 14$$

Jarak antar data ke-8 dengan *centroid* Pertama (C_0)

$$d(x_8, c_0) = \sqrt{(15 - 0)^2} = 15$$

Jarak antar data ke-8 dengan *centroid* Kedua (C_1)

$$d(x_8, c_1) = \sqrt{(15 - 18)^2} = 3$$

Jarak antar data ke-9 dengan *centroid* Pertama (C_0)

$$d(x_9, c_0) = \sqrt{(8 - 0)^2} = 8$$

Jarak antar data ke-9 dengan *centroid* Kedua (C_1)

$$d(x_9, c_1) = \sqrt{(8 - 18)^2} = 10$$

Jarak antar data ke-10 dengan *centroid* Pertama (C_0)

$$d(x_{10}, c_0) = \sqrt{(8 - 0)^2} = 8$$

Jarak antar data ke-10 dengan *centroid* Kedua (C_1)

$$d(x_{10}, c_1) = \sqrt{(8 - 18)^2} = 10$$

Jarak antar data ke-11 dengan *centroid* Pertama (C_0)

$$d(x_{11}, c_0) = \sqrt{(6 - 0)^2} = 6$$

Jarak antar data ke-11 dengan *centroid* Kedua (C_1)

$$d(x_{11}, c_1) = \sqrt{(6 - 18)^2} = 12$$

Jarak antar data ke-12 dengan *centroid* Pertama (C_0)

$$d(x_{12}, c_0) = \sqrt{(11 - 0)^2} = 11$$

Jarak antar data ke-12 dengan *centroid* Kedua (C_1)

$$d(x_{12}, c_1) = \sqrt{(11 - 18)^2} = 7$$

Jarak antar data ke-13 dengan *centroid* Pertama (C_0)

$$d(x_{13}, c_0) = \sqrt{(4 - 0)^2} = 4$$

Jarak antar data ke-13 dengan *centroid* Kedua (C_1)

$$d(x_{13}, c_1) = \sqrt{(4 - 18)^2} = 14$$

Jarak antar data ke-14 dengan *centroid* Pertama (C_0)

$$d(x_{14}, c_0) = \sqrt{(10 - 0)^2} = 10$$

Jarak antar data ke-14 dengan *centroid* Kedua (C_1)

$$d(x_{14}, c_1) = \sqrt{(10 - 18)^2} = 8$$

Jarak antar data ke-15 dengan *centroid* Pertama (C_0)

$$d(x_{15}, c_0) = \sqrt{(8 - 0)^2} = 8$$

Jarak antar data ke-15 dengan *centroid* Kedua (C_1)

$$d(x_{15}, c_1) = \sqrt{(8 - 18)^2} = 10$$

Jarak antar data ke-16 dengan *centroid* Pertama (C_0)

$$d(x_{16}, c_0) = \sqrt{(4 - 0)^2} = 4$$

Jarak antar data ke-16 dengan *centroid* Kedua (C_1)

$$d(x_{16}, c_1) = \sqrt{(4 - 18)^2} = 14$$

Jarak antar data ke-17 dengan *centroid* Pertama (C_0)

$$d(x_{17}, c_0) = \sqrt{(12 - 0)^2} = 12$$

Jarak antar data ke-17 dengan *centroid* Kedua (C_1)

$$d(x_{17}, c_1) = \sqrt{(12 - 18)^2} = 6$$

Jarak antar data ke-18 dengan *centroid* Pertama (C_0)

$$d(x_{18}, c_0) = \sqrt{(8 - 0)^2} = 8$$

Jarak antar data ke-18 dengan *centroid* Kedua (C_1)

$$d(x_{18}, c_1) = \sqrt{(8 - 18)^2} = 10$$

Jarak antar data ke-19 dengan *centroid* Pertama (C_0)

$$d(x_{19}, c_0) = \sqrt{(4 - 0)^2} = 4$$

Jarak antar data ke-19 dengan *centroid* Kedua (C_1)

$$d(x_{19}, c_1) = \sqrt{(4 - 18)^2} = 14$$

Jarak antar data ke-20 dengan *centroid* Pertama (C_0)

$$d(x_{20}, c_0) = \sqrt{(9 - 0)^2} = 9$$

Jarak antar data ke-20 dengan *centroid* Kedua (C_1)

$$d(x_{20}, c_1) = \sqrt{(9 - 18)^2} = 9$$

Dari perhitungan data diatas maka diperoleh data 516 siswa dinyatakan masuk C_0 dan 17 masuk C_1 , sesuai hasil nilai yang terkecil.

Tabel 4.4. Pengelompokan data pada Literasi ke-1

NO	ID	C₀	C₁	Cluster
1	2322	0	18	C₀
2	2323	9	9	C₀
3	2325	9	9	C₀
4	2326	8	10	C₀
5	2327	9	9	C₀
6	2358	5	13	C₀
7	2361	4	14	C₀
8	2328	15	3	C₁
9	2213	8	10	C₀
10	2206	8	10	C₀
11	2207	6	12	C₀
12	2208	11	7	C₁
13	2209	4	14	C₀
14	2210	10	8	C₁
15	2211	8	10	C₀
16	2212	4	14	C₀
17	2214	12	6	C₁
18	2215	8	10	C₀
19	2216	4	14	C₀
20	2217	9	9	C₀
....
533	2366	0	18	C₀

Setelah mendapatkan label cluster untuk masing-masing data maka dicari nilai rata-ratanya dengan menjumlahkan seluruh anggota masing-masing cluster dan dibagi jumlah anggotanya.

Tabel 4.5. Centroid baru pada literasi ke-1

Cluster	x
C ₀	1,6
C ₁	12,71

Literasi Ke-2

Jarak antar data ke-1 dengan *centroid* Pertama (C₀)

$$d(x_1, c_0) = \sqrt{(0 - 1,6)^2} = 1,60$$

Jarak antar data ke-1 dengan *centroid* Kedua (C₁)

$$d(x_1, c_1) = \sqrt{(0 - 12,71)^2} = 12,71$$

Jarak antar data ke-2 dengan *centroid* Pertama (C₀)

$$d(x_2, c_0) = \sqrt{(9 - 1,6)^2} = 7,40$$

Jarak antar data ke-2 dengan *centroid* Kedua (C₁)

$$d(x_2, c_1) = \sqrt{(9 - 12,71)^2} = 3,71$$

Jarak antar data ke-3 dengan *centroid* Pertama (C₀)

$$d(x_3, c_0) = \sqrt{(9 - 1,6)^2} = 7,40$$

Jarak antar data ke-3 dengan *centroid* Kedua (C₁)

$$d(x_3, c_1) = \sqrt{(9 - 12,71)^2} = 3,71$$

Jarak antar data ke-4 dengan *centroid* Pertama (C₀)

$$d(x_4, c_0) = \sqrt{(8 - 1,6)^2} = 6,40$$

Jarak antar data ke-4 dengan *centroid* Kedua (C₁)

$$d(x_4, c_1) = \sqrt{(8 - 12,71)^2} = 4,71$$

Jarak antar data ke-5 dengan *centroid* Pertama (C₀)

$$d(x_5, c_0) = \sqrt{(9 - 1,6)^2} = 7,40$$

Jarak antar data ke-5 dengan *centroid* Kedua (C_1)

$$d(x_5, c_1) = \sqrt{(9 - 12,71)^2} = 3,71$$

Jarak antar data ke-6 dengan *centroid* Pertama (C_0)

$$d(x_6, c_0) = \sqrt{(5 - 1,6)^2} = 3,40$$

Jarak antar data ke-6 dengan *centroid* Kedua (C_1)

$$d(x_6, c_1) = \sqrt{(5 - 12,71)^2} = 7,71$$

Jarak antar data ke-7 dengan *centroid* Pertama (C_0)

$$d(x_7, c_0) = \sqrt{(4 - 1,6)^2} = 2,40$$

Jarak antar data ke-7 dengan *centroid* Kedua (C_1)

$$d(x_7, c_1) = \sqrt{(4 - 12,71)^2} = 8,71$$

Jarak antar data ke-8 dengan *centroid* Pertama (C_0)

$$d(x_8, c_0) = \sqrt{(15 - 1,6)^2} = 15$$

Jarak antar data ke-8 dengan *centroid* Kedua (C_1)

$$d(x_8, c_1) = \sqrt{(15 - 12,71)^2} = 2,29$$

Jarak antar data ke-9 dengan *centroid* Pertama (C_0)

$$d(x_9, c_0) = \sqrt{(8 - 1,6)^2} = 6,40$$

Jarak antar data ke-9 dengan *centroid* Kedua (C_1)

$$d(x_9, c_1) = \sqrt{(8 - 12,71)^2} = 4,71$$

Jarak antar data ke-10 dengan *centroid* Pertama (C_0)

$$d(x_{10}, c_0) = \sqrt{(8 - 1,6)^2} = 6,40$$

Jarak antar data ke-10 dengan *centroid* Kedua (C_1)

$$d(x_{10}, c_1) = \sqrt{(8 - 12,71)^2} = 4,71$$

Jarak antar data ke-11 dengan *centroid* Pertama (C_0)

$$d(x_{11}, c_0) = \sqrt{(6 - 1,6)^2} = 4,40$$

Jarak antar data ke-11 dengan *centroid* Kedua (C_1)

$$d(x_{11}, c_1) = \sqrt{(6 - 12,71)^2} = 6,71$$

Jarak antar data ke-12 dengan *centroid* Pertama (C_0)

$$d(x_{12}, c_0) = \sqrt{(11 - 1,6)^2} = 9,40$$

Jarak antar data ke-12 dengan *centroid* Kedua (C_1)

$$d(x_{12}, c_1) = \sqrt{(11 - 12,71)^2} = 1,71$$

Jarak antar data ke-13 dengan *centroid* Pertama (C_0)

$$d(x_{13}, c_0) = \sqrt{(4 - 1,6)^2} = 2,40$$

Jarak antar data ke-13 dengan *centroid* Kedua (C_1)

$$d(x_{13}, c_1) = \sqrt{(4 - 12,71)^2} = 8,71$$

Jarak antar data ke-14 dengan *centroid* Pertama (C_0)

$$d(x_{14}, c_0) = \sqrt{(10 - 1,6)^2} = 8,40$$

Jarak antar data ke-14 dengan *centroid* Kedua (C_1)

$$d(x_{14}, c_1) = \sqrt{(10 - 12,71)^2} = 2,71$$

Jarak antar data ke-15 dengan *centroid* Pertama (C_0)

$$d(x_{15}, c_0) = \sqrt{(8 - 1,6)^2} = 6,40$$

Jarak antar data ke-15 dengan *centroid* Kedua (C_1)

$$d(x_{15}, c_1) = \sqrt{(8 - 12,71)^2} = 4,71$$

Jarak antar data ke-16 dengan *centroid* Pertama (C_0)

$$d(x_{16}, c_0) = \sqrt{(4 - 1,6)^2} = 2,40$$

Jarak antar data ke-16 dengan *centroid* Kedua (C_1)

$$d(x_{16}, c_1) = \sqrt{(4 - 12,71)^2} = 8,71$$

Jarak antar data ke-17 dengan *centroid* Pertama (C_0)

$$d(x_{17}, c_0) = \sqrt{(12 - 1,6)^2} = 10,40$$

Jarak antar data ke-17 dengan *centroid* Kedua (C_1)

$$d(x_{17}, c_1) = \sqrt{(12 - 12,71)^2} = 0,71$$

Jarak antar data ke-18 dengan *centroid* Pertama (C_0)

$$d(x_{18}, c_0) = \sqrt{(8 - 1,6)^2} = 6,40$$

Jarak antar data ke-18 dengan *centroid* Kedua (C_1)

$$d(x_{18}, c_1) = \sqrt{(8 - 12,71)^2} = 4,71$$

Jarak antar data ke-19 dengan *centroid* Pertama (C_0)

$$d(x_{19}, c_0) = \sqrt{(4 - 1,6)^2} = 2,40$$

Jarak antar data ke-19 dengan *centroid* Kedua (C_1)

$$d(x_{19}, c_1) = \sqrt{(4 - 12,71)^2} = 8,71$$

Jarak antar data ke-20 dengan *centroid* Pertama (C_0)

$$d(x_{20}, c_0) = \sqrt{(9 - 1,6)^2} = 7,40$$

Jarak antar data ke-20 dengan *centroid* Kedua (C_1)

$$d(x_{20}, c_1) = \sqrt{(9 - 12,71)^2} = 3,71$$

Dari perhitungan data diatas maka diperoleh data 471 siswa dinyatakan masuk C_0 dan 62 masuk C_1 , sesuai hasil nilai yang terkecil.

Tabel 4.6. Pengelompokan data pada Literasi ke-2

NO	ID	C0	C1	Cluster
1	2322	1,60	12,71	C_0
2	2323	7,40	3,71	C_1

3	2325	7,40	3,71	C ₁
4	2326	6,40	4,71	C ₁
5	2327	7,40	3,71	C ₁
6	2358	3,40	7,71	C ₀
7	2361	2,40	8,71	C ₀
8	2328	13,40	2,29	C ₁
9	2213	6,40	4,71	C ₁
10	2206	6,40	4,71	C ₁
11	2207	4,40	6,71	C ₀
12	2208	9,40	1,71	C ₁
13	2209	2,40	8,71	C ₀
14	2210	8,40	2,71	C ₁
15	2211	6,40	4,71	C ₁
16	2212	2,40	8,71	C ₀
17	2214	10,40	0,71	C ₁
18	2215	6,40	4,71	C ₁
19	2216	2,40	8,71	C ₀
20	2217	7,40	3,71	C ₁
....
533	2366	1,60	112,71	C ₀

Setelah mendapatkan label cluster untuk masing-masing data maka dicari nilai rata-ratanya dengan menjumlahkan seluruh anggota masing-masing cluster dan dibagi jumlah anggotanya.

Tabel 4.7. Centroid baru pada literasi ke-2

Cluster	x
C ₀	1,8
C ₁	7,88

Literasi Ke-3

Jarak antar data ke-1 dengan *centroid* Pertama (C₀)

$$d(x_1, c_0) = \sqrt{(0 - 1,8)^2} = 1,80$$

Jarak antar data ke-1 dengan *centroid* Kedua (C_1)

$$d(x_1, c_1) = \sqrt{(0 - 7,88)^2} = 7,88$$

Jarak antar data ke-2 dengan *centroid* Pertama (C_0)

$$d(x_2, c_0) = \sqrt{(9 - 1,8)^2} = 7,20$$

Jarak antar data ke-2 dengan *centroid* Kedua (C_1)

$$d(x_2, c_1) = \sqrt{(9 - 7,88)^2} = 1,12$$

Jarak antar data ke-3 dengan *centroid* Pertama (C_0)

$$d(x_3, c_0) = \sqrt{(9 - 1,8)^2} = 7,20$$

Jarak antar data ke-3 dengan *centroid* Kedua (C_1)

$$d(x_3, c_1) = \sqrt{(9 - 7,88)^2} = 1,12$$

Jarak antar data ke-4 dengan *centroid* Pertama (C_0)

$$d(x_4, c_0) = \sqrt{(8 - 1,8)^2} = 6,20$$

Jarak antar data ke-4 dengan *centroid* Kedua (C_1)

$$d(x_4, c_1) = \sqrt{(8 - 7,88)^2} = 0,12$$

Jarak antar data ke-5 dengan *centroid* Pertama (C_0)

$$d(x_5, c_0) = \sqrt{(9 - 1,8)^2} = 7,20$$

Jarak antar data ke-5 dengan *centroid* Kedua (C_1)

$$d(x_5, c_1) = \sqrt{(9 - 7,88)^2} = 1,12$$

Jarak antar data ke-6 dengan *centroid* Pertama (C_0)

$$d(x_6, c_0) = \sqrt{(5 - 1,8)^2} = 3,20$$

Jarak antar data ke-6 dengan *centroid* Kedua (C_1)

$$d(x_6, c_1) = \sqrt{(5 - 7,88)^2} = 2,88$$

Jarak antar data ke-7 dengan *centroid* Pertama (C_0)

$$d(x_7, c_0) = \sqrt{(4 - 1,8)^2} = 2,20$$

Jarak antar data ke-7 dengan *centroid* Kedua (C_1)

$$d(x_7, c_1) = \sqrt{(4 - 7,88)^2} = 3,88$$

Jarak antar data ke-8 dengan *centroid* Pertama (C_0)

$$d(x_8, c_0) = \sqrt{(15 - 1,8)^2} = 13,20$$

Jarak antar data ke-8 dengan *centroid* Kedua (C_1)

$$d(x_8, c_1) = \sqrt{(15 - 7,88)^2} = 7,12$$

Jarak antar data ke-9 dengan *centroid* Pertama (C_0)

$$d(x_9, c_0) = \sqrt{(8 - 1,8)^2} = 6,20$$

Jarak antar data ke-9 dengan *centroid* Kedua (C_1)

$$d(x_9, c_1) = \sqrt{(8 - 7,88)^2} = 0,12$$

Jarak antar data ke-10 dengan *centroid* Pertama (C_0)

$$d(x_{10}, c_0) = \sqrt{(8 - 1,8)^2} = 6,20$$

Jarak antar data ke-10 dengan *centroid* Kedua (C_1)

$$d(x_{10}, c_1) = \sqrt{(8 - 7,88)^2} = 0,12$$

Jarak antar data ke-11 dengan *centroid* Pertama (C_0)

$$d(x_{11}, c_0) = \sqrt{(6 - 1,8)^2} = 4,20$$

Jarak antar data ke-11 dengan *centroid* Kedua (C_1)

$$d(x_{11}, c_1) = \sqrt{(6 - 7,88)^2} = 1,88$$

Jarak antar data ke-12 dengan *centroid* Pertama (C_0)

$$d(x_{12}, c_0) = \sqrt{(11 - 1,8)^2} = 9,20$$

Jarak antar data ke-12 dengan *centroid* Kedua (C_1)

$$d(x_{12}, c_1) = \sqrt{(11 - 7,88)^2} = 3,12$$

Jarak antar data ke-13 dengan *centroid* Pertama (C_0)

$$d(x_{13}, c_0) = \sqrt{(4 - 1,8)^2} = 2,20$$

Jarak antar data ke-13 dengan *centroid* Kedua (C_1)

$$d(x_{13}, c_1) = \sqrt{(4 - 7,88)^2} = 3,88$$

Jarak antar data ke-14 dengan *centroid* Pertama (C_0)

$$d(x_{14}, c_0) = \sqrt{(10 - 1,8)^2} = 8,20$$

Jarak antar data ke-14 dengan *centroid* Kedua (C_1)

$$d(x_{14}, c_1) = \sqrt{(10 - 7,88)^2} = 2,12$$

Jarak antar data ke-15 dengan *centroid* Pertama (C_0)

$$d(x_{15}, c_0) = \sqrt{(8 - 1,8)^2} = 6,20$$

Jarak antar data ke-15 dengan *centroid* Kedua (C_1)

$$d(x_{15}, c_1) = \sqrt{(8 - 7,88)^2} = 0,12$$

Jarak antar data ke-16 dengan *centroid* Pertama (C_0)

$$d(x_{16}, c_0) = \sqrt{(4 - 1,8)^2} = 2,2$$

Jarak antar data ke-16 dengan *centroid* Kedua (C_1)

$$d(x_{16}, c_1) = \sqrt{(4 - 7,88)^2} = 3,88$$

Jarak antar data ke-17 dengan *centroid* Pertama (C_0)

$$d(x_{17}, c_0) = \sqrt{(12 - 1,8)^2} = 10,20$$

Jarak antar data ke-17 dengan *centroid* Kedua (C_1)

$$d(x_{17}, c_1) = \sqrt{(12 - 7,88)^2} = 4,12$$

Jarak antar data ke-18 dengan *centroid* Pertama (C_0)

$$d(x_{18}, c_0) = \sqrt{(8 - 1,8)^2} = 6,20$$

Jarak antar data ke-18 dengan *centroid* Kedua (C_1)

$$d(x_{18}, c_1) = \sqrt{(8 - 7,88)^2} = 0,12$$

Jarak antar data ke-19 dengan *centroid* Pertama (C_0)

$$d(x_{19}, c_0) = \sqrt{(4 - 1,8)^2} = 2,20$$

Jarak antar data ke-19 dengan *centroid* Kedua (C_1)

$$d(x_{19}, c_1) = \sqrt{(4 - 7,88)^2} = 3,88$$

Jarak antar data ke-20 dengan *centroid* Pertama (C_0)

$$d(x_{20}, c_0) = \sqrt{(9 - 1,8)^2} = 7,20$$

Jarak antar data ke-20 dengan *centroid* Kedua (C_1)

$$d(x_{20}, c_1) = \sqrt{(9 - 7,88)^2} = 1,12$$

Dari perhitungan data diatas maka diperoleh data 446 siswa dinyatakan masuk C_0 dan 87 masuk C_1 , sesuai hasil nilai yang terkecil.

Tabel 4.8. Pengelompokan data pada Literasi ke-3

NO	ID	C_0	C_1	Cluster
1	2322	1,80	7,88	C_0
2	2323	7,20	1,12	C_1
3	2325	7,20	1,12	C_1
4	2326	6,20	0,12	C_1
5	2327	7,20	1,12	C_1
6	2358	3,20	2,88	C_1
7	2361	2,20	3,88	C_0
8	2328	13,20	7,12	C_1
9	2213	6,20	0,12	C_1
10	2206	6,20	0,12	C_1
11	2207	4,20	1,88	C_1

12	2208	9,20	3,12	C ₁
13	2209	2,20	3,88	C ₀
14	2210	8,20	2,12	C ₁
15	2211	6,20	0,12	C ₁
16	2212	2,20	3,88	C ₀
17	2214	10,20	4,12	C ₁
18	2215	6,20	0,12	C ₁
19	2216	2,20	3,88	C ₀
20	2217	7,20	1,12	C ₁
....
533	2366	1,80	7,88	C ₀

Setelah mendapatkan label cluster untuk masing-masing data maka dicari nilai rata-ratanya dengan menjumlahkan seluruh anggota masing-masing cluster dan dibagi jumlah anggotanya.

Tabel 4.9. Centroid baru pada literasi ke-3

Cluster	x
C ₀	1,8
C ₁	6,54

Literasi Ke-4

Jarak antar data ke-1 dengan *centroid* Pertama (C₀)

$$d(x_1, c_0) = \sqrt{(0 - 1,8)^2} = 1,8$$

Jarak antar data ke-1 dengan *centroid* Kedua (C₁)

$$d(x_1, c_1) = \sqrt{(0 - 6,54)^2} = 6,54$$

Jarak antar data ke-2 dengan *centroid* Pertama (C₀)

$$d(x_2, c_0) = \sqrt{(9 - 1,8)^2} = 7,20$$

Jarak antar data ke-2 dengan *centroid* Kedua (C₁)

$$d(x_2, c_1) = \sqrt{(9 - 6,54)^2} = 12,46$$

Jarak antar data ke-3 dengan *centroid* Pertama (C_0)

$$d(x_3, c_0) = \sqrt{(9 - 1,8)^2} = 7,20$$

Jarak antar data ke-3 dengan *centroid* Kedua (C_1)

$$d(x_3, c_1) = \sqrt{(9 - 6,54)^2} = 2,46$$

Jarak antar data ke-4 dengan *centroid* Pertama (C_0)

$$d(x_4, c_0) = \sqrt{(8 - 1,8)^2} = 6,20$$

Jarak antar data ke-4 dengan *centroid* Kedua (C_1)

$$d(x_4, c_1) = \sqrt{(8 - 6,54)^2} = 1,46$$

Jarak antar data ke-5 dengan *centroid* Pertama (C_0)

$$d(x_5, c_0) = \sqrt{(9 - 1,8)^2} = 7,20$$

Jarak antar data ke-5 dengan *centroid* Kedua (C_1)

$$d(x_5, c_1) = \sqrt{(9 - 6,54)^2} = 2,46$$

Jarak antar data ke-6 dengan *centroid* Pertama (C_0)

$$d(x_6, c_0) = \sqrt{(5 - 1,8)^2} = 3,20$$

Jarak antar data ke-6 dengan *centroid* Kedua (C_1)

$$d(x_6, c_1) = \sqrt{(5 - 6,54)^2} = 1,54$$

Jarak antar data ke-7 dengan *centroid* Pertama (C_0)

$$d(x_7, c_0) = \sqrt{(4 - 1,8)^2} = 2,20$$

Jarak antar data ke-7 dengan *centroid* Kedua (C_1)

$$d(x_7, c_1) = \sqrt{(4 - 6,54)^2} = 2,54$$

Jarak antar data ke-8 dengan *centroid* Pertama (C_0)

$$d(x_8, c_0) = \sqrt{(15 - 1,8)^2} = 13,20$$

Jarak antar data ke-8 dengan *centroid* Kedua (C_1)

$$d(x_8, c_1) = \sqrt{(15 - 6,54)^2} = 8,46$$

Jarak antar data ke-9 dengan *centroid* Pertama (C_0)

$$d(x_9, c_0) = \sqrt{(8 - 1,8)^2} = 6,20$$

Jarak antar data ke-9 dengan *centroid* Kedua (C_1)

$$d(x_9, c_1) = \sqrt{(8 - 6,54)^2} = 1,46$$

Jarak antar data ke-10 dengan *centroid* Pertama (C_0)

$$d(x_{10}, c_0) = \sqrt{(8 - 1,8)^2} = 6,20$$

Jarak antar data ke-10 dengan *centroid* Kedua (C_1)

$$d(x_{10}, c_1) = \sqrt{(8 - 6,54)^2} = 1,46$$

Jarak antar data ke-11 dengan *centroid* Pertama (C_0)

$$d(x_{11}, c_0) = \sqrt{(6 - 1,8)^2} = 4,20$$

Jarak antar data ke-11 dengan *centroid* Kedua (C_1)

$$d(x_{11}, c_1) = \sqrt{(6 - 6,54)^2} = 0,54$$

Jarak antar data ke-12 dengan *centroid* Pertama (C_0)

$$d(x_{12}, c_0) = \sqrt{(11 - 1,8)^2} = 9,20$$

Jarak antar data ke-12 dengan *centroid* Kedua (C_1)

$$d(x_{12}, c_1) = \sqrt{(11 - 6,54)^2} = 4,46$$

Jarak antar data ke-13 dengan *centroid* Pertama (C_0)

$$d(x_{13}, c_0) = \sqrt{(4 - 1,8)^2} = 2,20$$

Jarak antar data ke-13 dengan *centroid* Kedua (C_1)

$$d(x_{13}, c_1) = \sqrt{(4 - 6,54)^2} = 2,54$$

Jarak antar data ke-14 dengan *centroid* Pertama (C_0)

$$d(x_{14}, c_0) = \sqrt{(10 - 1,8)^2} = 8,20$$

Jarak antar data ke-14 dengan *centroid* Kedua (C_1)

$$d(x_{14}, c_1) = \sqrt{(10 - 6,54)^2} = 3,46$$

Jarak antar data ke-15 dengan *centroid* Pertama (C_0)

$$d(x_{15}, c_0) = \sqrt{(8 - 1,8)^2} = 6,20$$

Jarak antar data ke-15 dengan *centroid* Kedua (C_1)

$$d(x_{15}, c_1) = \sqrt{(8 - 6,54)^2} = 1,46$$

Jarak antar data ke-16 dengan *centroid* Pertama (C_0)

$$d(x_{16}, c_0) = \sqrt{(4 - 1,8)^2} = 2,2$$

Jarak antar data ke-16 dengan *centroid* Kedua (C_1)

$$d(x_{16}, c_1) = \sqrt{(4 - 6,54)^2} = 2,54$$

Jarak antar data ke-17 dengan *centroid* Pertama (C_0)

$$d(x_{17}, c_0) = \sqrt{(12 - 1,8)^2} = 10,20$$

Jarak antar data ke-17 dengan *centroid* Kedua (C_1)

$$d(x_{17}, c_1) = \sqrt{(12 - 6,54)^2} = 5,46$$

Jarak antar data ke-18 dengan *centroid* Pertama (C_0)

$$d(x_{18}, c_0) = \sqrt{(8 - 1,8)^2} = 6,20$$

Jarak antar data ke-18 dengan *centroid* Kedua (C_1)

$$d(x_{18}, c_1) = \sqrt{(8 - 6,54)^2} = 1,46$$

Jarak antar data ke-19 dengan *centroid* Pertama (C_0)

$$d(x_{19}, c_0) = \sqrt{(4 - 1,8)^2} = 2,20$$

Jarak antar data ke-19 dengan *centroid* Kedua (C_1)

$$d(x_{19}, c_1) = \sqrt{(4 - 6,54)^2} = 2,54$$

Jarak antar data ke-20 dengan *centroid* Pertama (C_0)

$$d(x_{20}, c_0) = \sqrt{(9 - 1,8)^2} = 7,20$$

Jarak antar data ke-20 dengan *centroid* Kedua (C_1)

$$d(x_{20}, c_1) = \sqrt{(9 - 6,54)^2} = 2,46$$

Dari perhitungan data diatas maka diperoleh data 446 siswa dinyatakan masuk C_0 dan 87 masuk C_1 , sesuai hasil nilai yang terkecil.

Tabel 4.10. Pengelompokan data pada Literasi ke-4

NO	ID	C_0	C_1	Cluster
1	2322	1,80	6,54	C_0
2	2323	7,20	2,46	C_1
3	2325	7,20	2,46	C_1
4	2326	6,20	1,46	C_1
5	2327	7,20	2,46	C_1
6	2358	3,20	1,54	C_1
7	2361	2,20	2,54	C_0
8	2328	13,20	8,46	C_1
9	2213	6,20	1,46	C_1
10	2206	6,20	1,46	C_1
11	2207	4,20	0,54	C_1
12	2208	9,20	4,46	C_1
13	2209	2,20	2,54	C_0
14	2210	8,20	3,46	C_1
15	2211	6,20	1,46	C_1
16	2212	2,20	2,54	C_0
17	2214	10.20	5.46	C_1
18	2215	6.20	1.46	C_1
19	2216	2.20	2.54	C_0
20	2217	7.20	2.46	C_1

....
533	2366	1,80	6,54	C ₀

Setelah mendapatkan label cluster untuk masing-masing data maka dicari nilai rata-ratanya dengan menjumlahkan seluruh anggota masing-masing cluster dan dibagi jumlah anggotanya.

Tabel 4.11 Centroid baru pada literasi ke-4

Cluster	x
C ₀	1,81
C ₁	6,53

Literasi Ke-5

Jarak antar data ke-1 dengan *centroid* Pertama (C₀)

$$d(x_1, c_0) = \sqrt{(0 - 1,81)^2} = 1,81$$

Jarak antar data ke-1 dengan *centroid* Kedua (C₁)

$$d(x_1, c_1) = \sqrt{(0 - 6,53)^2} = 6,53$$

Jarak antar data ke-2 dengan *centroid* Pertama (C₀)

$$d(x_2, c_0) = \sqrt{(9 - 1,81)^2} = 7,19$$

Jarak antar data ke-2 dengan *centroid* Kedua (C₁)

$$d(x_2, c_1) = \sqrt{(9 - 6,53)^2} = 2,47$$

Jarak antar data ke-3 dengan *centroid* Pertama (C₀)

$$d(x_3, c_0) = \sqrt{(9 - 1,81)^2} = 7,19$$

Jarak antar data ke-3 dengan *centroid* Kedua (C₁)

$$d(x_3, c_1) = \sqrt{(9 - 6,53)^2} = 2,47$$

Jarak antar data ke-4 dengan *centroid* Pertama (C_0)

$$d(x_4, c_0) = \sqrt{(8 - 1,81)^2} = 6,19$$

Jarak antar data ke-4 dengan *centroid* Kedua (C_1)

$$d(x_4, c_1) = \sqrt{(8 - 6,53)^2} = 1,47$$

Jarak antar data ke-5 dengan *centroid* Pertama (C_0)

$$d(x_5, c_0) = \sqrt{(9 - 1,81)^2} = 7,19$$

Jarak antar data ke-5 dengan *centroid* Kedua (C_1)

$$d(x_5, c_1) = \sqrt{(9 - 6,53)^2} = 2,47$$

Jarak antar data ke-6 dengan *centroid* Pertama (C_0)

$$d(x_6, c_0) = \sqrt{(5 - 1,81)^2} = 3,19$$

Jarak antar data ke-6 dengan *centroid* Kedua (C_1)

$$d(x_6, c_1) = \sqrt{(5 - 6,53)^2} = 1,53$$

Jarak antar data ke-7 dengan *centroid* Pertama (C_0)

$$d(x_7, c_0) = \sqrt{(4 - 1,81)^2} = 2,19$$

Jarak antar data ke-7 dengan *centroid* Kedua (C_1)

$$d(x_7, c_1) = \sqrt{(4 - 6,53)^2} = 2,53$$

Jarak antar data ke-8 dengan *centroid* Pertama (C_0)

$$d(x_8, c_0) = \sqrt{(15 - 1,81)^2} = 13,19$$

Jarak antar data ke-8 dengan *centroid* Kedua (C_1)

$$d(x_8, c_1) = \sqrt{(15 - 6,53)^2} = 8,47$$

Jarak antar data ke-9 dengan *centroid* Pertama (C_0)

$$d(x_9, c_0) = \sqrt{(8 - 1,81)^2} = 6,19$$

Jarak antar data ke-9 dengan *centroid* Kedua (C_1)

$$d(x_9, c_1) = \sqrt{(8 - 6,53)^2} = 1,47$$

Jarak antar data ke-10 dengan *centroid* Pertama (C_0)

$$d(x_{10}, c_0) = \sqrt{(8 - 1,81)^2} = 6,19$$

Jarak antar data ke-10 dengan *centroid* Kedua (C_1)

$$d(x_{10}, c_1) = \sqrt{(8 - 6,53)^2} = 1,47$$

Jarak antar data ke-11 dengan *centroid* Pertama (C_0)

$$d(x_{11}, c_0) = \sqrt{(6 - 1,81)^2} = 4,19$$

Jarak antar data ke-11 dengan *centroid* Kedua (C_1)

$$d(x_{11}, c_1) = \sqrt{(6 - 6,53)^2} = 0,53$$

Jarak antar data ke-12 dengan *centroid* Pertama (C_0)

$$d(x_{12}, c_0) = \sqrt{(11 - 1,81)^2} = 9,19$$

Jarak antar data ke-12 dengan *centroid* Kedua (C_1)

$$d(x_{12}, c_1) = \sqrt{(11 - 6,53)^2} = 4,47$$

Jarak antar data ke-13 dengan *centroid* Pertama (C_0)

$$d(x_{13}, c_0) = \sqrt{(4 - 1,81)^2} = 2,19$$

Jarak antar data ke-13 dengan *centroid* Kedua (C_1)

$$d(x_{13}, c_1) = \sqrt{(4 - 6,53)^2} = 2,53$$

Jarak antar data ke-14 dengan *centroid* Pertama (C_0)

$$d(x_{14}, c_0) = \sqrt{(10 - 1,81)^2} = 8,19$$

Jarak antar data ke-14 dengan *centroid* Kedua (C_1)

$$d(x_{14}, c_1) = \sqrt{(10 - 6,53)^2} = 3,47$$

Jarak antar data ke-15 dengan *centroid* Pertama (C_0)

$$d(x_{15}, c_0) = \sqrt{(8 - 1,81)^2} = 6,19$$

Jarak antar data ke-15 dengan *centroid* Kedua (C_1)

$$d(x_{15}, c_1) = \sqrt{(8 - 6,53)^2} = 1,47$$

Jarak antar data ke-16 dengan *centroid* Pertama (C_0)

$$d(x_{16}, c_0) = \sqrt{(4 - 1,81)^2} = 2,19$$

Jarak antar data ke-16 dengan *centroid* Kedua (C_1)

$$d(x_{16}, c_1) = \sqrt{(4 - 6,53)^2} = 2,53$$

Jarak antar data ke-17 dengan *centroid* Pertama (C_0)

$$d(x_{17}, c_0) = \sqrt{(12 - 1,81)^2} = 10,19$$

Jarak antar data ke-17 dengan *centroid* Kedua (C_1)

$$d(x_{17}, c_1) = \sqrt{(12 - 6,53)^2} = 5,47$$

Jarak antar data ke-18 dengan *centroid* Pertama (C_0)

$$d(x_{18}, c_0) = \sqrt{(8 - 1,81)^2} = 6,19$$

Jarak antar data ke-18 dengan *centroid* Kedua (C_1)

$$d(x_{18}, c_1) = \sqrt{(8 - 6,53)^2} = 1,47$$

Jarak antar data ke-19 dengan *centroid* Pertama (C_0)

$$d(x_{19}, c_0) = \sqrt{(4 - 1,81)^2} = 2,19$$

Jarak antar data ke-19 dengan *centroid* Kedua (C_1)

$$d(x_{19}, c_1) = \sqrt{(4 - 6,53)^2} = 2,53$$

Jarak antar data ke-20 dengan *centroid* Pertama (C_0)

$$d(x_{20}, c_0) = \sqrt{(9 - 1,81)^2} = 7,19$$

Jarak antar data ke-20 dengan *centroid* Kedua (C_1)

$$d(x_{20}, c_1) = \sqrt{(9 - 6,53)^2} = 2,47$$

Dari perhitungan data diatas maka diperoleh data 446 siswa dinyatakan masuk C_0 dan 87 masuk C_1 , sesuai hasil nilai yang terkecil.

Tabel 4.12. Pengelompokan data pada Literasi ke-5

NO	ID	C_0	C_1	Cluster
1	2322	1,81	6,53	C_0
2	2323	7,19	2,47	C_1
3	2325	7,19	2,47	C_1
4	2326	6,19	1,47	C_1
5	2327	7,19	2,47	C_1
6	2358	3,19	1,53	C_1
7	2361	2,19	2,53	C_0
8	2328	13,19	8,47	C_1
9	2213	6,19	1,47	C_1
10	2206	6,19	1,47	C_1
11	2207	4,19	0,53	C_1
12	2208	9,19	4,47	C_1
13	2209	2,19	2,53	C_0
14	2210	8,19	3,47	C_1
15	2211	6,19	1,47	C_1
....
533	2366	1,81	6,53	C_0

Setelah mendapatkan label cluster untuk masing-masing data maka dicari nilai rata-ratanya dengan menjumlahkan seluruh anggota masing-masing cluster dan dibagi jumlah anggotanya.

Tabel 4.13. Centroid baru pada literasi ke-5

Cluster	x
C_0	1,81
C_1	6,53

Karena *centroid* tidak mengalami perubahan (sama dengan *centroid* sebelumnya) maka proses *clustering* selesai, dan hasil yang dipakai pada literasi terakhir.

4.4. Evaluasi dan Validasi

Hasil pada proses *Clustering* dalam *K-Means* sebelumnya akan dilakukan validasi dengan *Davies-Bouldin Index*. Data hasil *clustering* disajikan pada Tabel 4.12, sedangkan *centroid* setiap *cluster* disajikan pada tabel 4.13 parameter jarak yang digunakan adalah *Euclidian Distance*.

Tabel 4.14 menyajikan proses perhitungan SSW. Pada tabel tersebut dikelompokkan berdasarkan *cluster*-nya, kemudian dihitung jarak setiap *centroid* masing-masing dan dihitung rata-ratanya untuk menjadi SSW. Contoh perhitungan jarak data dalam *cluster* 1 ke *centroid*-nya sebagai berikut :

Nilai SSW untuk *cluster* 1 didapat sebagai berikut :

$$\begin{aligned}
 SSW_1 &= \frac{1}{m} (d(x_1, c_0) + d(x_2, c_0) + d(x_3, c_0) + \dots + d(x_{446}, c_0)) \\
 &= \frac{1}{446} (d(4,1,81) + d(4,1,81) + d(4,1,81) + \dots + d(0,1,81)) \\
 &= \frac{1}{446} (\sqrt{(4-1,81)^2} + \sqrt{(4-1,81)^2} + d\sqrt{(4-1,81)^2} + \dots + \sqrt{(0-1,81)^2}) \\
 &= \frac{1}{446} (2,19 + 2,19 + 2,19 + \dots + 1,81) = 1,81
 \end{aligned}$$

Nilai SSW untuk *cluster* 2 dihitung dengan cara yang sama seperti di atas

$$\begin{aligned}
 SSW_2 &= \frac{1}{m} (d(x_1, c_1) + d(x_2, c_1) + d(x_3, c_1) + \dots + d(x_{47}, c_1)) \\
 &= \frac{1}{47} (d(8,1,81) + d(8,1,81) + d(8,1,81) + \dots + d(0,1,81)) \\
 &= \frac{1}{47} (\sqrt{(8-1,81)^2} + \sqrt{(8-1,81)^2} + d\sqrt{(8-1,81)^2} + \dots + \sqrt{(9-1,81)^2}) \\
 &= \frac{1}{47} (1,47 + 1,47 + 1,47 + \dots + 2,47) = 2,44
 \end{aligned}$$

Tabel 4.14. Perhitungan SSW

NIS	x	cluster	Centroid	Jarak	SSW
2197	4	C ₀	1,81	2,19	1,81
2198	4			2,19	
2199	4			2,19	
2201	4			2,19	
2202	4			2,19	
2205	4			2,19	
2340	4			2,19	
2343	4			2,19	
2348	4			2,19	
2351	4			2,19	
...	
2366	0			1,81	
2200	8	C ₀	6,53	1,47	2,44
2203	8			1,47	
2204	8			1,47	
2341	8			1,47	
2342	8			1,47	
2344	8			1,47	
2345	12			5,47	
2346	8			1,47	
2347	5			1,53	
2349	5			1,53	
...	
2132	9			2,47	

SSB didapatkan dengan menghitung jarak antar centroid, hasilnya disajikan seperti pada Tabel 4.15. Contoh perhitungan SSB pasangan diantara dua *cluster* tersebut sebagai berikut :

$$\begin{aligned}
 SSB_{1,2} &= d(c_1c_2) = \sqrt{d(x_{c1} - x_{c2})^2} \\
 &= \sqrt{(1,81 - 6,53)^2} = 4,72
 \end{aligned}$$

Tabel 4.15. Perhitungan SSB

		Data ke-i	
		SSB	
Data ke-i	1	0	4,72
	2	4,72	0

Berikut nilai R yang didapatkan dari persamaan (2.4) :

$$R_{1,2} = \frac{SSW_1 + SSW_2}{SSB_{1,2}} = \frac{1,81 + 2,44}{4,72} = 0.90$$

Hasil DBI yang didapatkan dari persamaan (2.5). Hasilnya ditampilkan seperti dibawah ini :

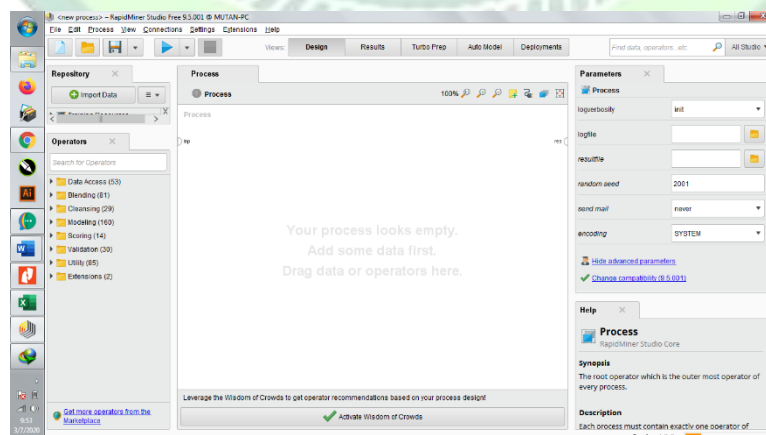
Tabel 4.16. Perhitungan R dan DBI

		Data ke-i		R Max	DBI
		R			
Data ke-i	1	0	0.90	0.90	0.90
	2	0.90	0	0.90	

Dari Tabel 4.16, terlihat bahwa nilai DBI yang didapat adalah 0.90.

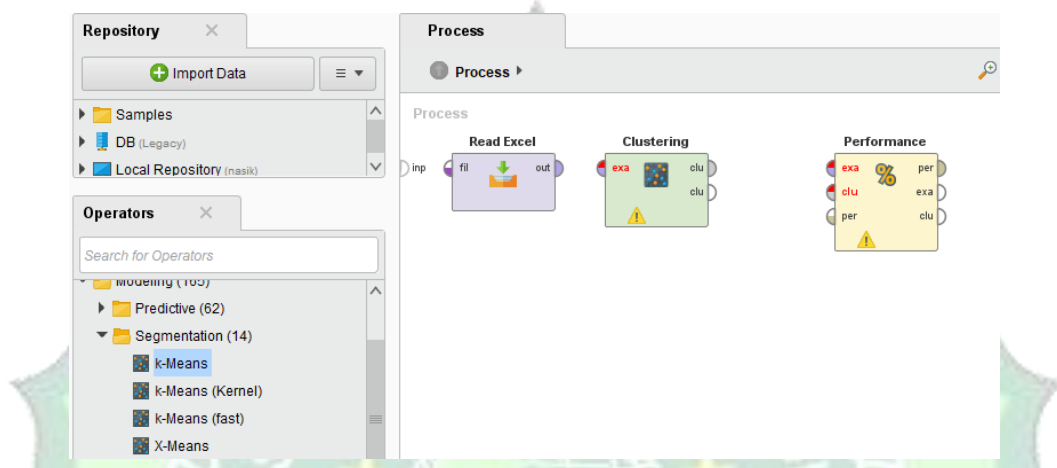
4.5. Pengujian Software *RapidMiner*

Data akan dianalisis berdasarkan nilai poinnya. Data siswa dibuat dalam *data set* baru dengan format *.xlsx (excel)* untuk bisa dianalisis *software* yang dipakai yakni *RapidMiner Studio* seperti pada gambar 4.3 sebagai berikut.



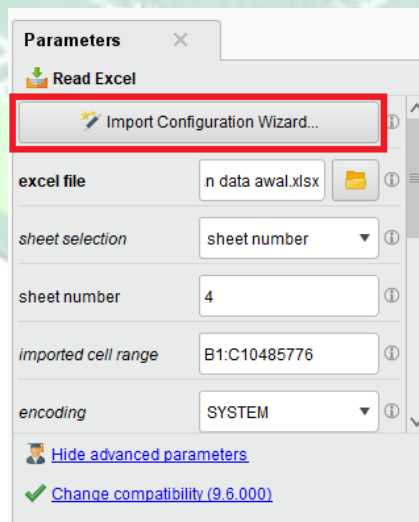
Gambar 4.3. RapidMiner Studio explorer

Setelah menentukan operator yang akan digunakan lakukan drag and drop operators seperti gambar 4.4 dibawah, Read Excel digunakan untuk membaca file excelyang telah kita buat, Clustering K-Means metode yang akan digunakan dan Performance digunakan untuk menghitung performanya. Kemudian masukan data yang mau diolah dengan menggunakan operator Read Excel dengan fitur Import Configuration Wizard seperti pada gambar 4.5

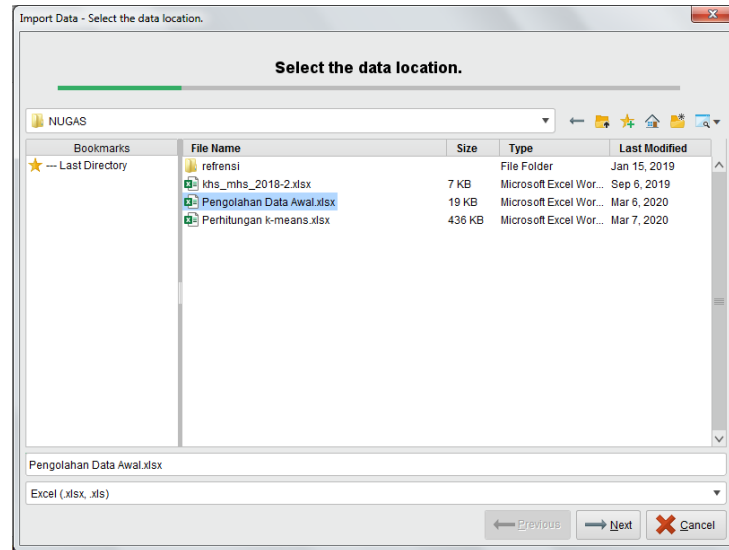


Gambar 4.4. Operator read excel, clustering dan Performance

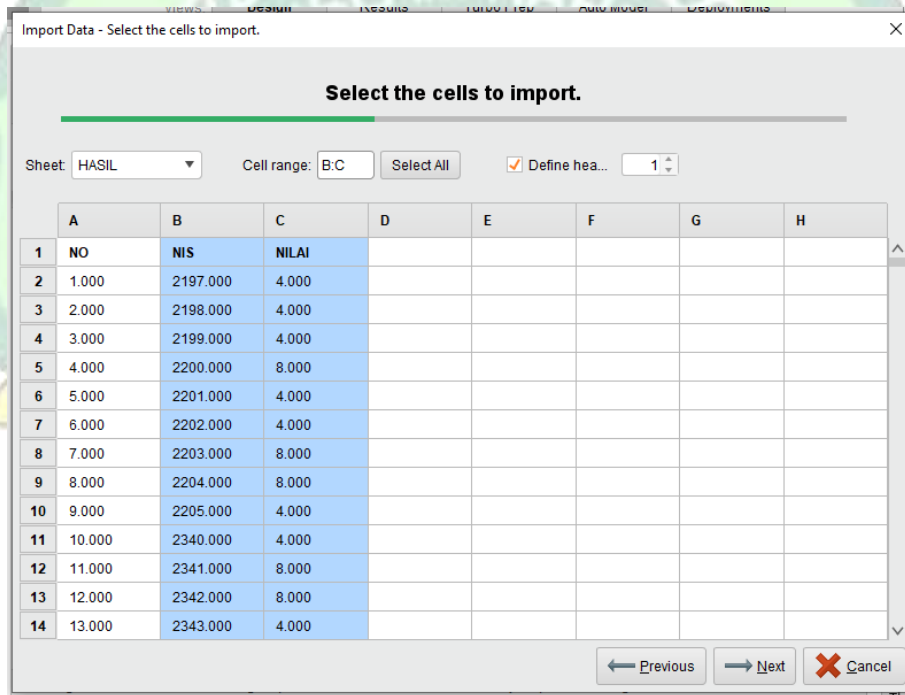
Proses open file ini dilakukan oleh operator *read excel* dan melalui parameter *Import Configuration Wizard* untuk menemukan file yang sudah dibuat dan pilih *sheet* tempat data tersebut, seperti pada gambar 4.5 sampai 4.7.



Gambar 4.5. Import Configuration Wizard

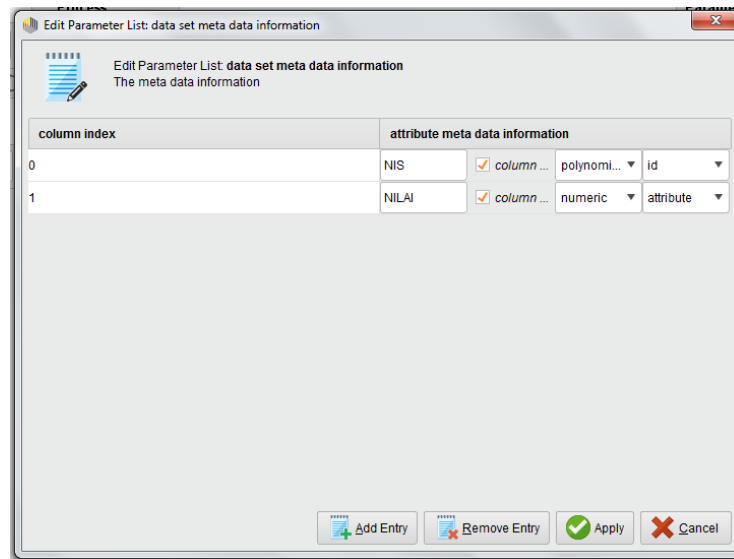


Gambar 4.6. Pilih file yang akan digunakan



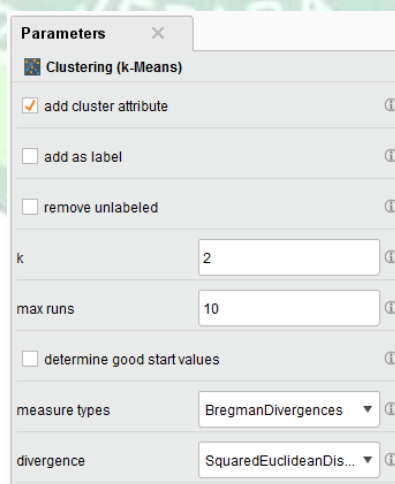
Gambar 4.7. Pilih Sheet dan Cells yang akan digunakan

agar tidak terjadi kesalahan *run* data yang sudah di *import* tadi, maka di lakukan *Edit parameter list* untuk memastikan datanya sudah benar sesuai dengan *type* datanya, jika *type* data tidak sesuai maka operator tidak akan melakukan proses *run*, terlihat pada gambar 4.8.

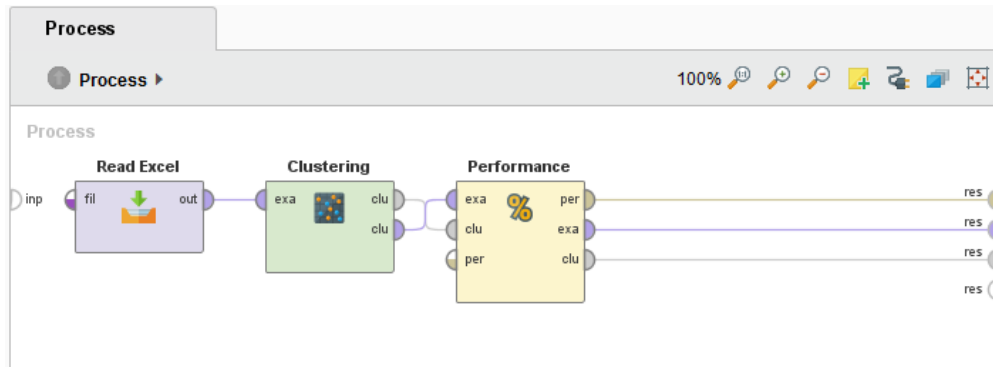


Gambar 4.8. Edit parameter list

Pada dasarnya *Rapidminer studio* dapat membaca *type data* setiap atribut seperti pada gambar 4.8, namun untuk mengurangi kesalahan pada proses run ada baiknya dipastikan kembali. Setelah *open file excel* langkah selanjutnya adalah menghubungkan *operator read excel* dengan operator *clustering k-means* dan mengatur melalui parameternya, sekaligus menentukan jumlah clusternya. Untuk menemukan jumlah yaitu 2 *cluster* seperti pada gambar 4.9. Setelah operator *read excel* dan operator *clustering k-means* diatur langkah berikutnya adalah menghubungkan kedua operator tersebut, dan hubungkan juga pada proses *output* untuk mendapatkan hasil *run* operatornya, seperti gambar 4.10.

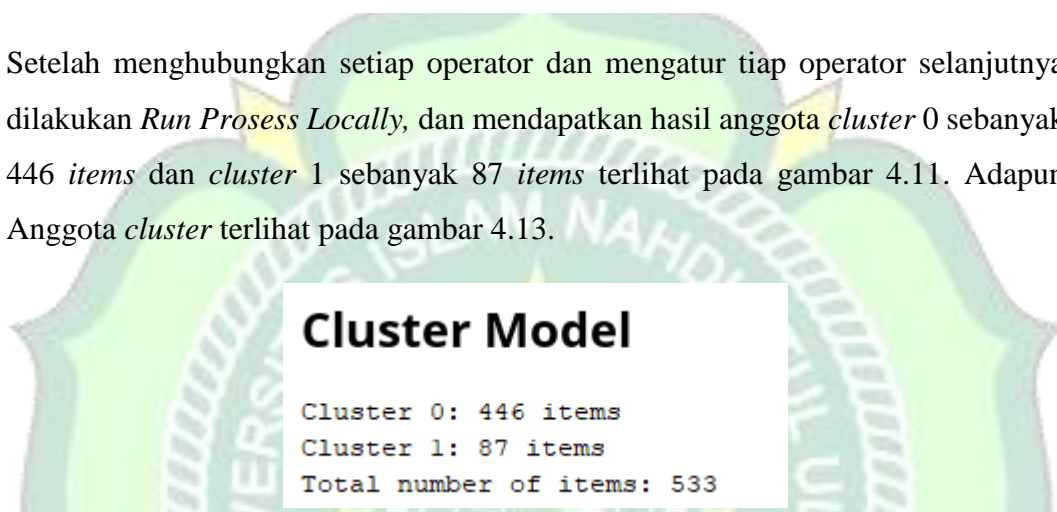


Gambar 4.9. Menentukan cluster

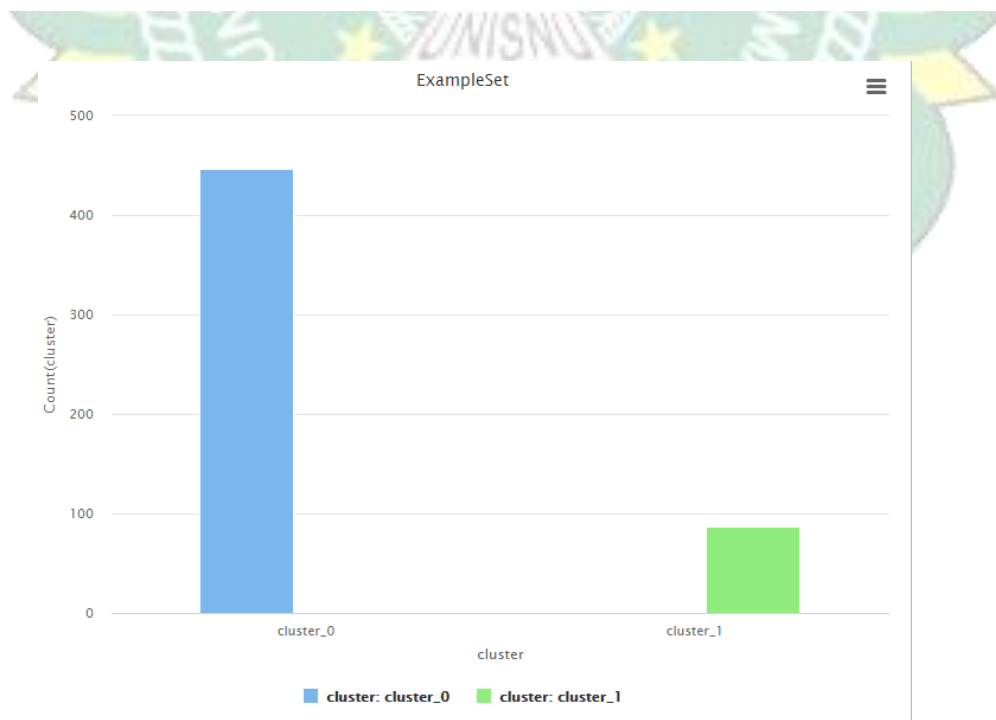


Gambar 4.10. Menghubungkan operator

Setelah menghubungkan setiap operator dan mengatur tiap operator selanjutnya dilakukan *Run Proses Locally*, dan mendapatkan hasil anggota *cluster 0* sebanyak 446 *items* dan *cluster 1* sebanyak 87 *items* terlihat pada gambar 4.11. Adapun Anggota *cluster* terlihat pada gambar 4.13.



Gambar 4.11. Hasil Cluster



Gambar 4.12. Grafik Cluster

Row No.	NIS	cluster	NILAI
1	2197	cluster_0	4
2	2198	cluster_0	4
3	2199	cluster_0	4
4	2200	cluster_1	8
5	2201	cluster_0	4
6	2202	cluster_0	4
7	2203	cluster_1	8
8	2204	cluster_1	8
9	2205	cluster_0	4
10	2340	cluster_0	4
11	2341	cluster_1	8
12	2342	cluster_1	8
13	2343	cluster_0	4
14	2344	cluster_1	8
15	2345	cluster_1	12

Gambar 4.13. Hasil Pengklompokan Anggota Cluster

Hasil Evaluasi *Cluster* dengan menggunakan *Davies bouldin index* dari perhitungan yang telah dilakukan terlihat pada gambar 4.14.

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: 3.015
Avg. within centroid distance_cluster_0: 2.165
Avg. within centroid distance_cluster_1: 7.372
Davies Bouldin: 0.397
```

Gambar 4.14. Hasil Performace