

Nama : Pawitro Purbangkoro

NIM : L200170045

Kelas : C

## MODUL10

### 10.1. Algoritma K-Means Menggunakan Rapid Miner

1. Buatlah tabel data nilai ujian siswa dengan nama **Tabel\_NilaiUjian.xls**.

	A	B	C	D
1	NO_SISWA	NAMA	B.IND	B.ING
2	S-101	JOKO	8,54	8,4
3	S-102	AGUS	9,98	6,81
4	S-103	SUSI	6,2	9,15
5	S-104	DYAH	5,24	7,26
6	S-105	WATI	5,7	5,71
7	S-106	IKA	8,57	5,87
8	S-107	EKO	7,7	7,71
9	S-108	YANTO	6,6	5,7
10	S-109	WAWAN	9	8,12
11	S-110	MAHMUD	9,81	9,58

2. Jalankan aplikasi RapidMiner
3. Gunakan file **Tabel\_NilaiUjian.xls** sebagai data yang digunakan dalam proses Clustering. Kita hanya menggunakan 3 kolom (nama siswa, nilai bahasa indonesia, nilai bahasa inggris).

Import Data - Select the cells to import.

**Select the cells to import.**

Sheet: Sheet1 Cell range: B:D Select All ☒ Define header row: 1

	A	B	C	D
1	NO_SISWA	NAMA	B.IND	B.ING
2	S-101	JOKO	8.540	8.400
3	S-102	AGUS	9.980	6.810
4	S-103	SUSI	6.200	9.150
5	S-104	DYAH	5.240	7.260
6	S-105	WATI	5.700	5.710
7	S-106	IKA	8.570	5.870
8	S-107	EKO	7.700	7.710
9	S-108	YANTO	6.600	5.700
10	S-109	WAWAN	9.000	8.120
11	S-110	MAHMUD	9.810	9.580

← Previous
Next →
✖ Cancel

- Ubah kolom NAMA menjadi id, dengan cara klik **Change role >> id >> OK**.

NAMA *polynomial* ⚙️ B.IND *real*

1 JOKO 8.540

Please enter the new role:

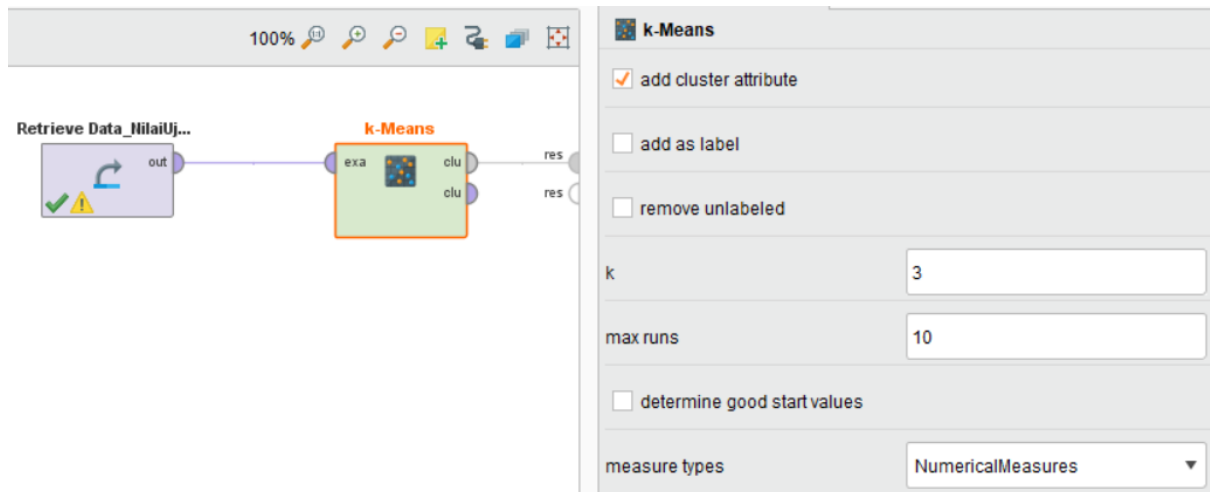
id

✓ OK ✖ Cancel

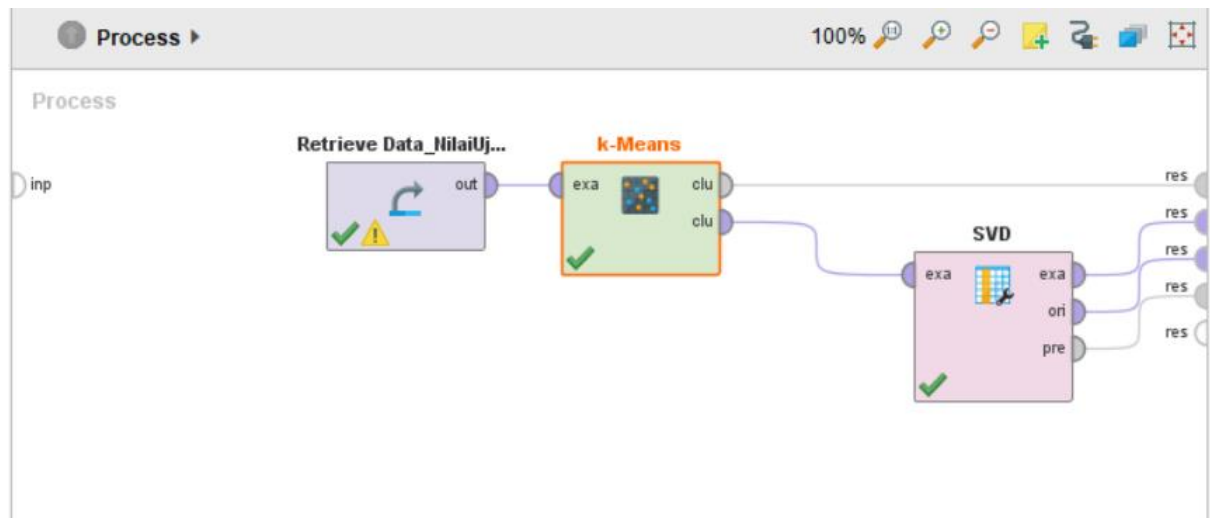
10 MAHMUD 9.810

- Beri nama **Data\_NilaiUjian** dan masukkan pada repositories. Kemudian klik **Finish**.
- Gunakan **Data\_NilaiUjian** ini dan masukkan ke dalam area process.

7. Tambahkan operator **k-Means**. Ubah nama operator ini menjadi k-Means. Hubungkan seperti gambar di bawah. Ubah parameter k=3 pada operator ini.



8. Tambahkan pula operator **SDV(Singular Value Decomposition)**.



9. Jalankan dengan menekan tombol Run (F11).
10. Berikut hasil proses Clustering dengan algoritma k-Means :

**a. SDV(Singular Value Decomposition)**

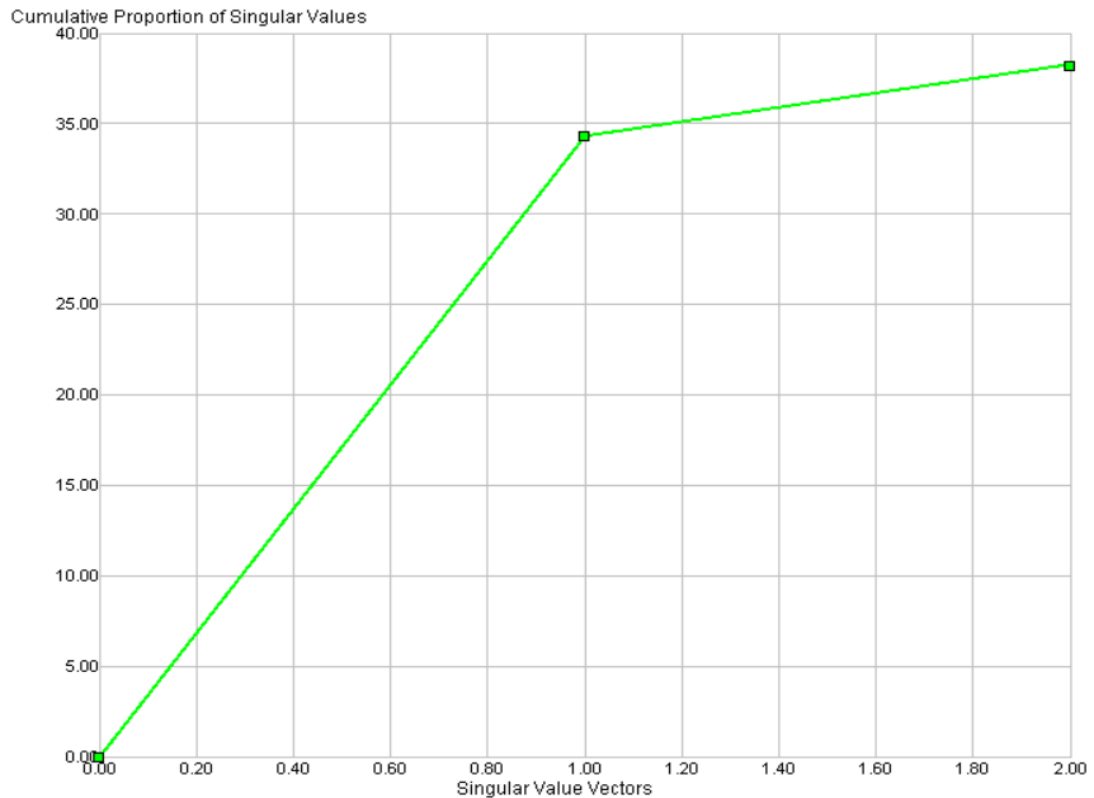
**i. Nilai Eigenvalue**

Component	Singular Value	Proportion of Singul...	Cumulative Singular ...	Cumulative Proporti...
SVD 1	34.340	0.898	34.340	0.898
SVD 2	3.906	0.102	38.246	1.000

**ii. Nilai Svd vectors**

Attribute	SVD Vector 1
B.IND	0.723
B.ING	0.690

### iii. Nilai Cumulative variance



### b. Example (k-Means)

Hasil ini kita lihat dengan mode Plot View menggunakan grafik Scatter untuk menentukan kelompok siswa (cluster) yang dicalonkan untuk maju ke dalam olimpiade mata pelajaran berdasarkan nilai tertinggi ujian.

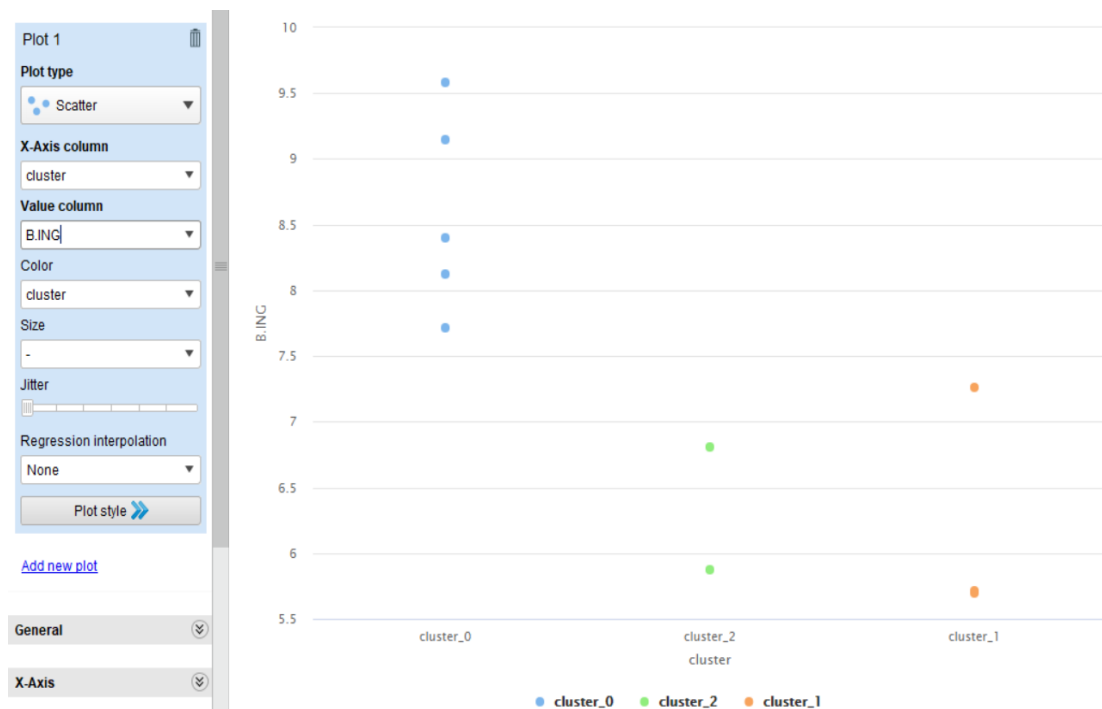
Ketentuan :

Plotter	= Scatter
x-Axis	= cluster
y-Axis	= B.IND, B.ING(diubah-ubah)
Color Column	= cluster
Jitter	= bisa diubah-ubah untuk melihat distribusi data secara lebih detail.

**i. Kelompok siswa bidang B. Indonesia**



**i. Kelompok siswa bidang B. Inggris**





**c. ExampleSet (SDV)**

Row No.	NAMA	cluster ↑	B.IND	B.ING
1	JOKO	cluster_0	8.540	8.400
3	SUSI	cluster_0	6.200	9.150
7	EKO	cluster_0	7.700	7.710
9	WAWAN	cluster_0	9	8.120
10	MAHMUD	cluster_0	9.810	9.580
4	DYAH	cluster_1	5.240	7.260
5	WATI	cluster_1	5.700	5.710
8	YANTO	cluster_1	6.600	5.700
2	AGUS	cluster_2	9.980	6.810
6	IKA	cluster_2	8.570	5.870

**d. Cluster Model (Clustering)**

**i. Description**

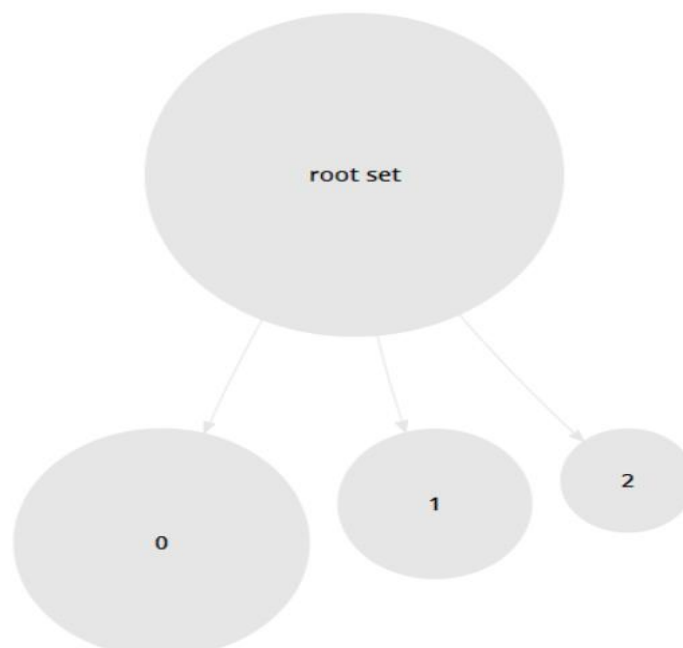
  
Description

  
Folder  
View

### Cluster Model

Cluster 0: 5 items  
Cluster 1: 3 items  
Cluster 2: 2 items  
Total number of items: 10

**ii. Graph**



## 10.2. Interpretasi Hasil Algoritma K-Means

CLUSTER	NO_SISWA	NAMA	B.IND	B.ING
0	S-101	JOKO	8,54	8,4
0	S-103	SUSI	6,2	9,15
0	S-107	EKO	7,7	7,71
0	S-109	WAWAN	9	8,12
0	S-110	MAHMUD	9,81	9,58
1	S-104	DYAH	5,24	7,26
1	S-105	WATI	5,7	5,71
1	S-108	YANTO	6,6	5,7
1	S-102	AGUS	9,98	6,81
1	S-106	IKA	8,57	5,87

## Tugas

1. Buatlah tabel berikut dengan menggunakan Microsoft Excel!

1	NO_SISWA	NAMA	B.IND	B.ING
2	S-101	JOKO	6,68	9,21
3	S-102	AGUS	9,02	6,42
4	S-103	SUSI	7,95	6,04
5	S-104	DYAH	6,65	6,89
6	S-105	WATI	7,40	7,59
7	S-106	IKA	8,46	7,63
8	S-107	EKO	9,48	9,16
9	S-108	YANTO	8,17	9,04
10	S-109	WAWAN	5,95	5,80
11	S-110	MAHMUD	6,51	8,94
12	S-111	BUDI	5,25	6,83
13	S-112	SANTI	6,50	7,17
14	S-113	DIAN	7,43	9,31
15	S-114	DANI	8,04	9,88
16	S-115	AHMAD	5,12	6,35
17	S-116	BAYU	9,97	9,71
18	S-117	RISA	7,96	8,98
19	S-118	RANI	7,46	5,29
20	S-119	YANI	6,67	9,15
21	S-120	RATIH	9,84	5,12
22	S-121	INDAH	9,37	6,69
23	S-122	JONO	9,31	9,28
24	S-123	SARAH	7,20	9,08
25	S-124	RAMA	7,99	7,05
26	S-125	BAMBANG	7,57	5,38
27	S-126	HADI	8,61	8,64
28	S-127	NANA	9,99	6,74
29	S-128	FEBRI	7,43	5,55
30	S-129	DENI	8,15	7,51
31	S-130	TONI	6,26	8,05

2. Ketentuan jumlah Cluster = 4.

Hasil :

**a. SDV (Singular Value Decomposition)**

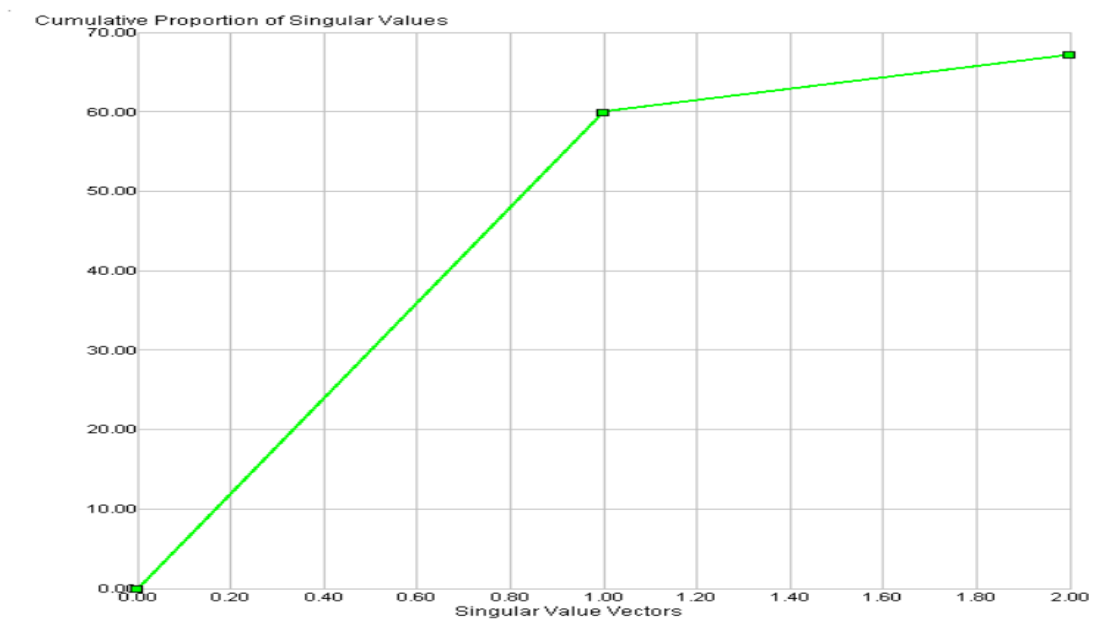
**i. Nilai Eigenvalue**

Component	Singular Value	Proportion of Si...	Cumulative Sin...	Cumulative Pro...
SVD 1	60.034	0.893	60.034	0.893
SVD 2	7.199	0.107	67.233	1.000

**ii. Nilai Sdv vectors**

Attribute	SVD Vector 1
B.IND	0.712
B.ING	0.702

**iii. Nilai Cumulative Variance**



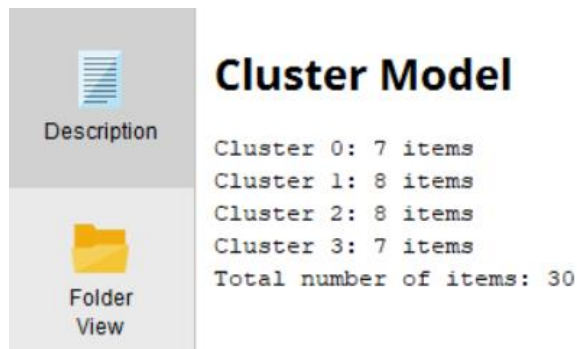


**b. ExampleSet (SDV)**

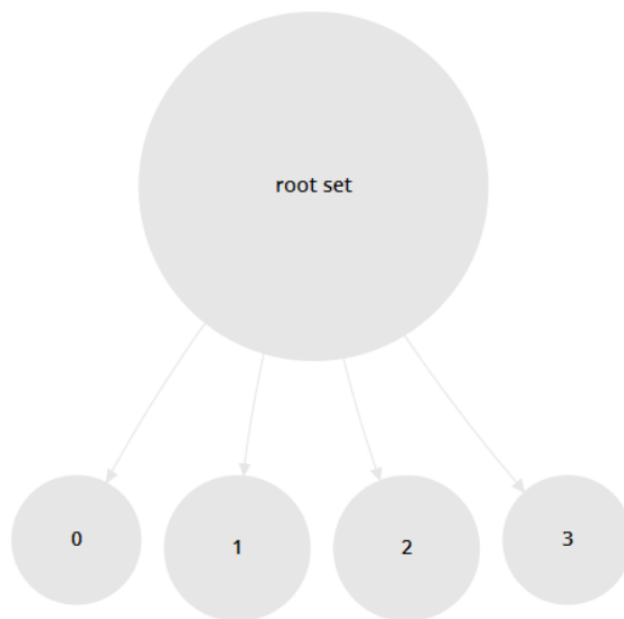
Row No.	NAMA	cluster ↑	svd_1
1	JOKO	cluster_0	0.187
5	WATI	cluster_0	0.177
10	MAHMUD	cluster_0	0.182
13	DIAN	cluster_0	0.197
19	YANI	cluster_0	0.186
23	SARAH	cluster_0	0.192
30	TONI	cluster_0	0.168
4	DYAH	cluster_1	0.159
9	WAWAN	cluster_1	0.138
11	BUDI	cluster_1	0.142
12	SANTI	cluster_1	0.161
15	AHMAD	cluster_1	0.135
18	RANI	cluster_1	0.150
25	BAMBANG	cluster_1	0.153
28	FEBRI	cluster_1	0.153
2	AGUS	cluster_2	0.182
3	SUSI	cluster_2	0.165
6	IKA	cluster_2	0.190
20	RATIH	cluster_2	0.177
21	INDAH	cluster_2	0.189
24	RAMA	cluster_2	0.177
27	NANA	cluster_2	0.197
29	DENI	cluster_2	0.185
7	EKO	cluster_3	0.220
8	YANTO	cluster_3	0.203
14	DANI	cluster_3	0.211
16	BAYU	cluster_3	0.232
17	RISA	cluster_3	0.199
22	JONO	cluster_3	0.219
26	HADI	cluster_3	0.203

**c. Description**

**i. Description**



**ii. Graph**



3. Tuliskan masing-masing nama siswa yang terdapat dalam Kelompok Cluster 0, Cluster 1, dan Cluster 3

- Cluster 0 : JOKO,WATI, MAHMUD, DIAN, YANI, SARAH, TONI
- Cluster 1 : DYAH, WAWAN, BUDI, SANTI, AHMAD, RANI, BAMBANG, FEBRI
- Cluster 2 : AGUS, SUSI, IKA, RATIH, INDAH, RAMA, NANA, DENI
- Cluster 3 : EKO, YANTO, DANI, BAYU, RISA, JONO, HADI