LAPORAN PRAKTIKUM
DATA WAREHOUSING DAN DATA MINING

PERTEMUAN 9
"CLUSTERING : K-MEANS"

Oleh:

NAMA : Daffa Putra Alwansyah

NIM : L200190031

KELAS : B

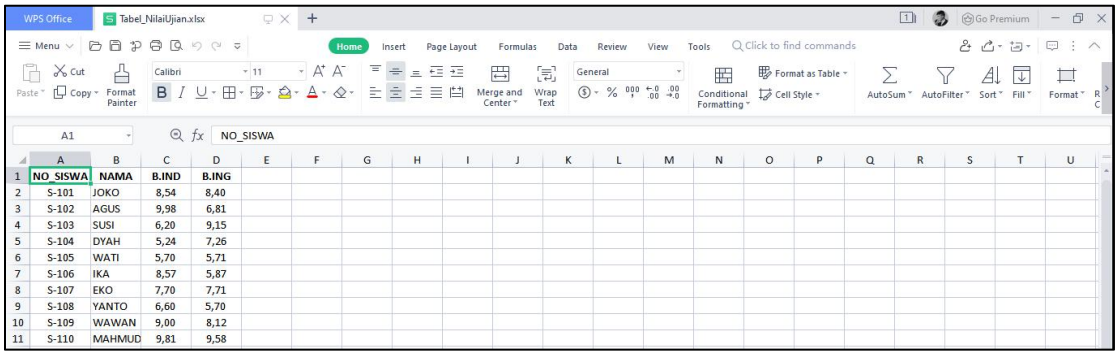
PRODI : INFORMATIKA

Fakultas Komunikasi dan Informatika

Universitas Muhammadiyah Surakarta

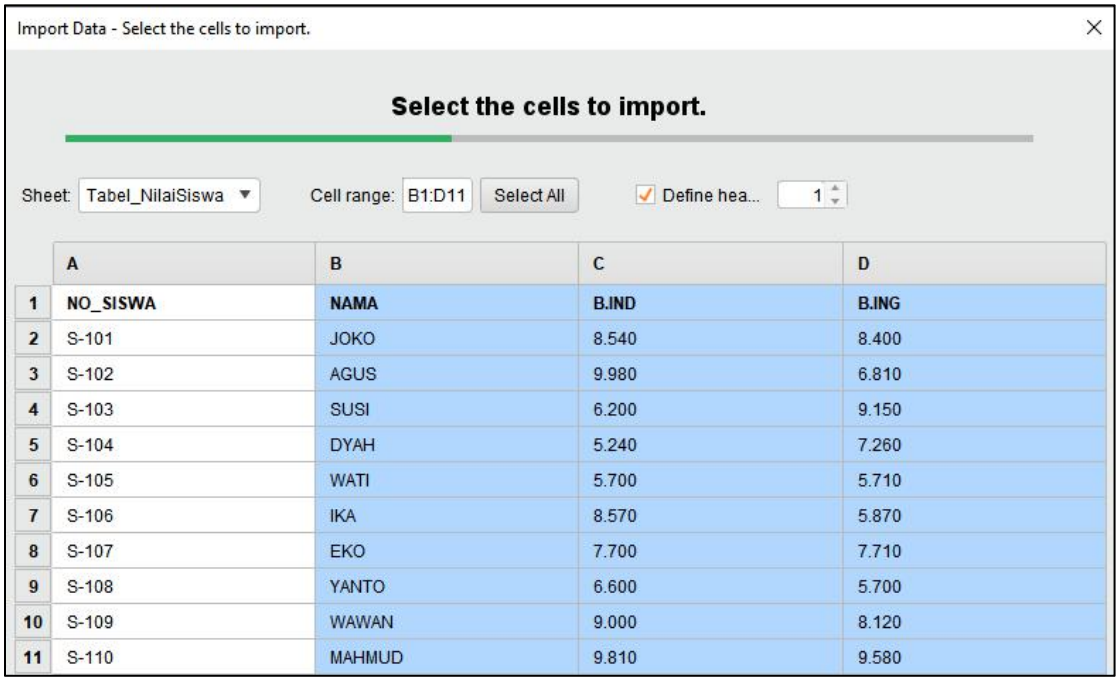
10.4 Langkah-Langkah Praktikum

1. Membuat Tabel_NilaiUjian



	A	B	C	D
1	NO_SISWA	NAMA	B.IND	B.ING
2	S-101	JOKO	8,54	8,40
3	S-102	AGUS	9,98	6,81
4	S-103	SUSI	6,20	9,15
5	S-104	DYAH	5,24	7,26
6	S-105	WATI	5,70	5,71
7	S-106	IKA	8,57	5,87
8	S-107	EKO	7,70	7,71
9	S-108	YANTO	6,60	5,70
10	S-109	WAWAN	9,00	8,12
11	S-110	MAHMUD	9,81	9,58

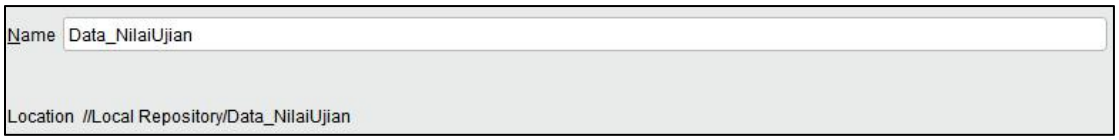
2. Seleksi tiga saja, yaitu Nama, B.IND, dan B.ING



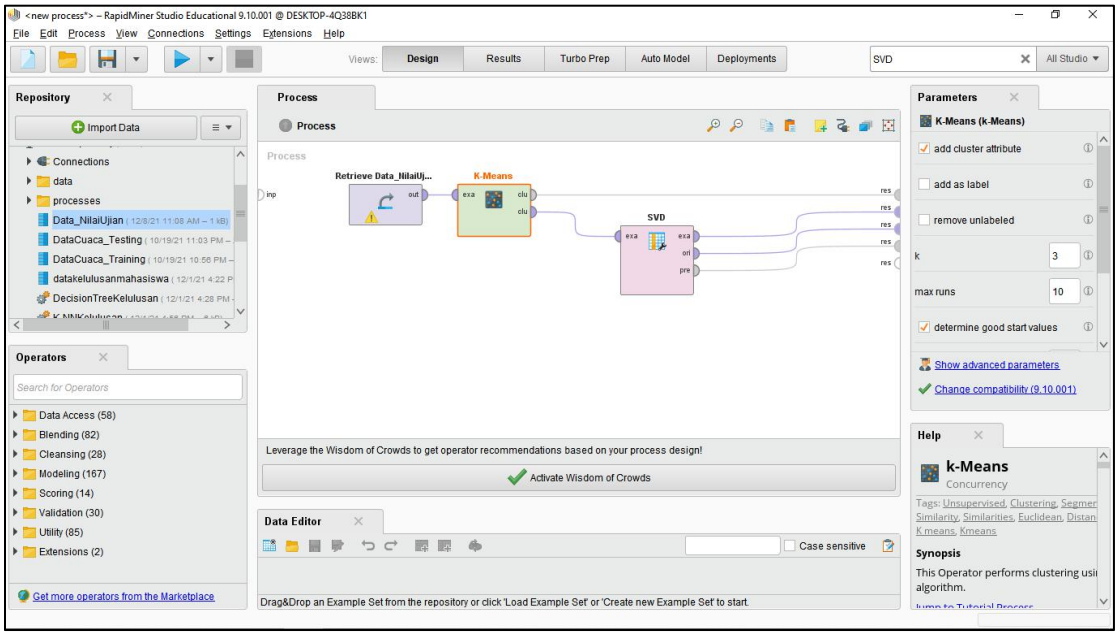
	A	B	C	D
1	NO_SISWA	NAMA	B.IND	B.ING
2	S-101	JOKO	8.540	8.400
3	S-102	AGUS	9.980	6.810
4	S-103	SUSI	6.200	9.150
5	S-104	DYAH	5.240	7.260
6	S-105	WATI	5.700	5.710
7	S-106	IKA	8.570	5.870
8	S-107	EKO	7.700	7.710
9	S-108	YANTO	6.600	5.700
10	S-109	WAWAN	9.000	8.120
11	S-110	MAHMUD	9.810	9.580

3. Ubah role Nama menjadi id.

4. Ubah menjadi Data_NilaiUjian



5. Masukan Data_NilaiUjian, K-Means (ubah K = 3), dan SVD, setelah itu sambungkan ke port-port sesuai modul, lalu klik tombol run/f11.



- Berikut adalah hasil proses Clustering dengan algoritma K-Means:

  a) SVD (Singular Value Decomposition)

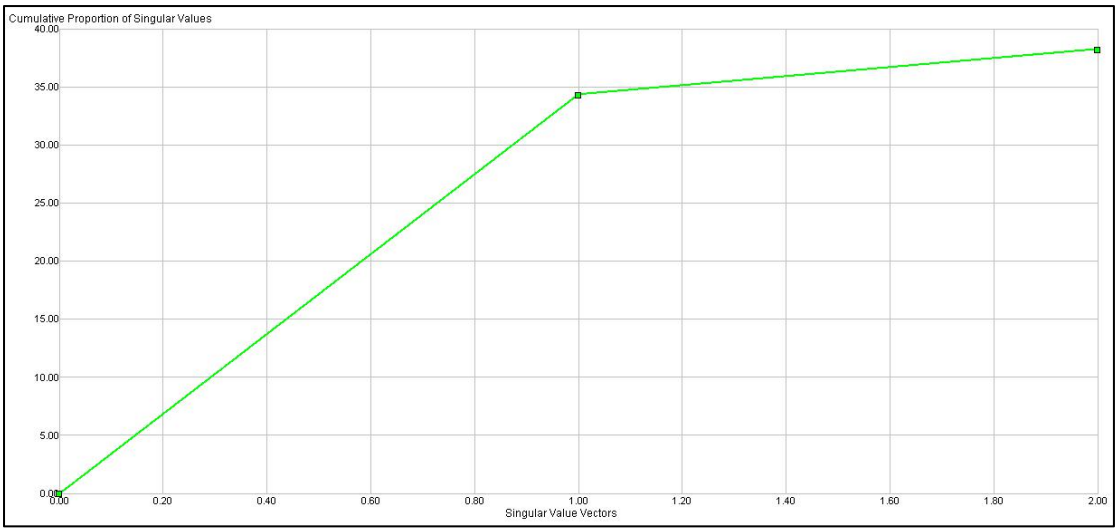
  i. Nilai Eigenvalue

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

  ii. Nilai vsd Vectors

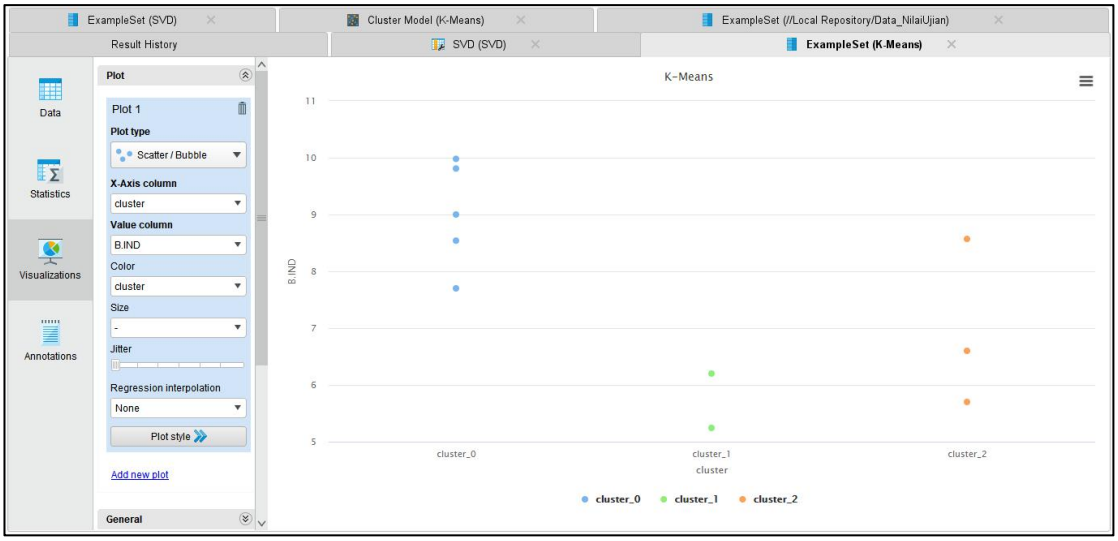
Attribute	SVD Vector 1
B.IND	0.723
B.ING	0.690

iii. Nilai Cumulative Variance

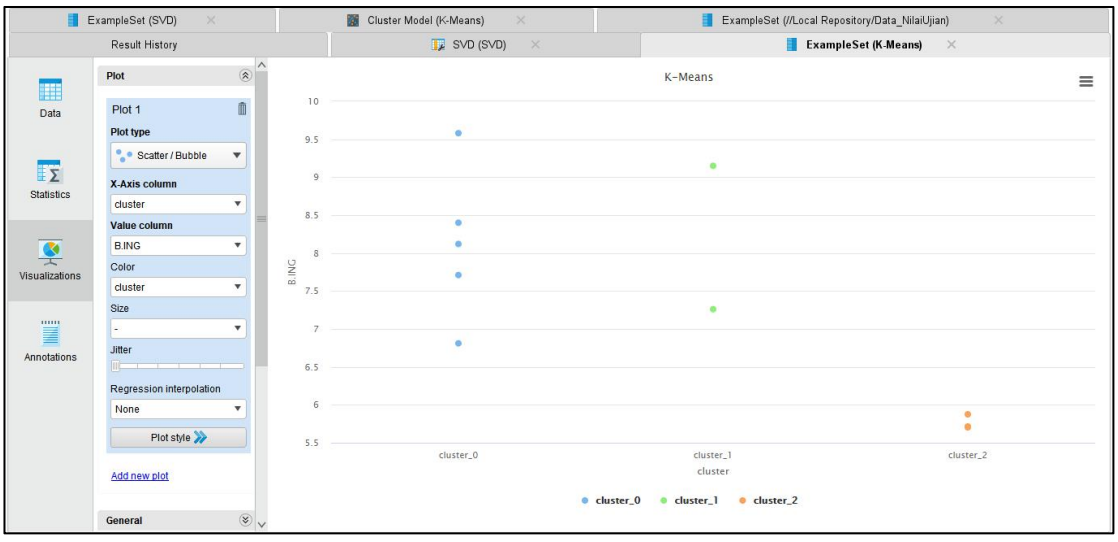


b) ExampleSet (k-Means)

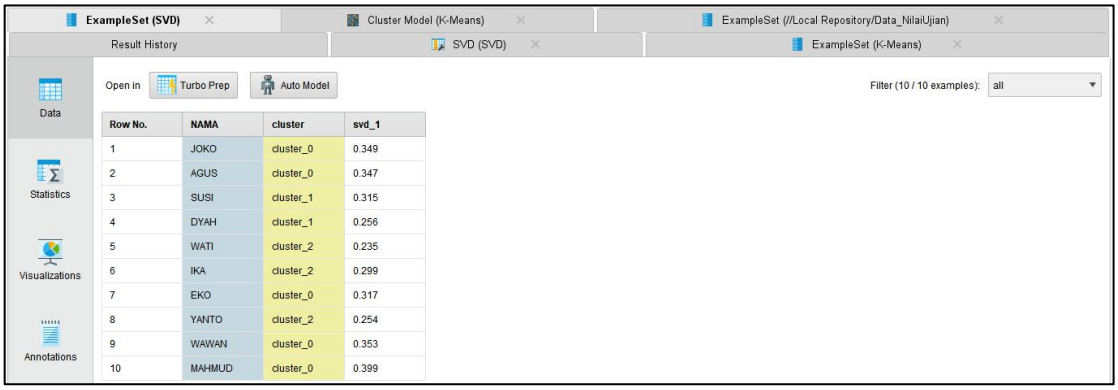
i. Kelompok siswa bidang B. Indonesia.



ii. Kelompok siswa bidang B. Inggris



c) ExampleSet(SVD)

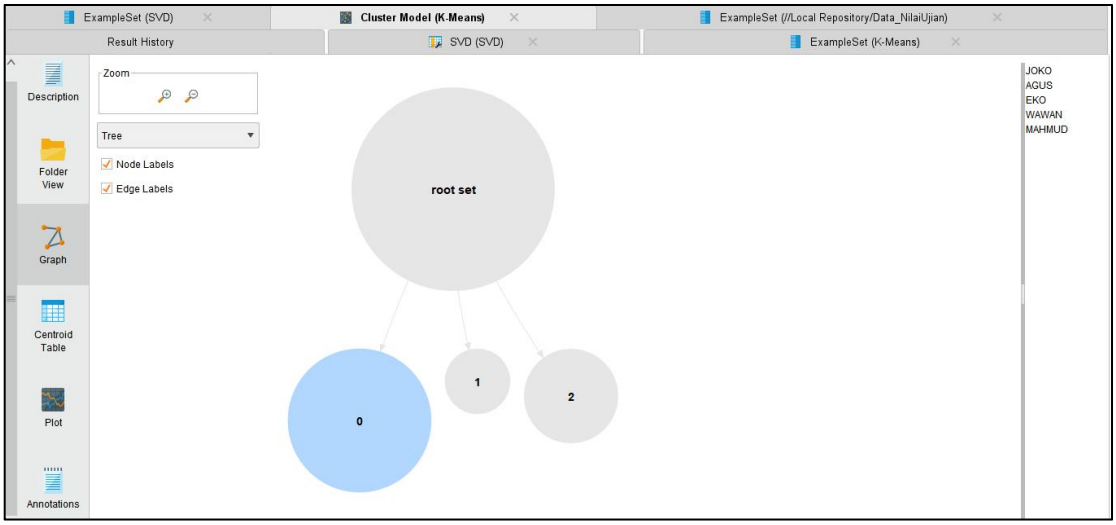


Row No.	NAMA	cluster	svd_1
1	JOKO	cluster_0	0.349
2	AGUS	cluster_0	0.347
3	SUSI	cluster_1	0.315
4	DYAH	cluster_1	0.256
5	WATI	cluster_2	0.235
6	IKA	cluster_2	0.299
7	EKO	cluster_0	0.317
8	YANTO	cluster_2	0.254
9	WAWAN	cluster_0	0.353
10	MAHMUD	cluster_0	0.399

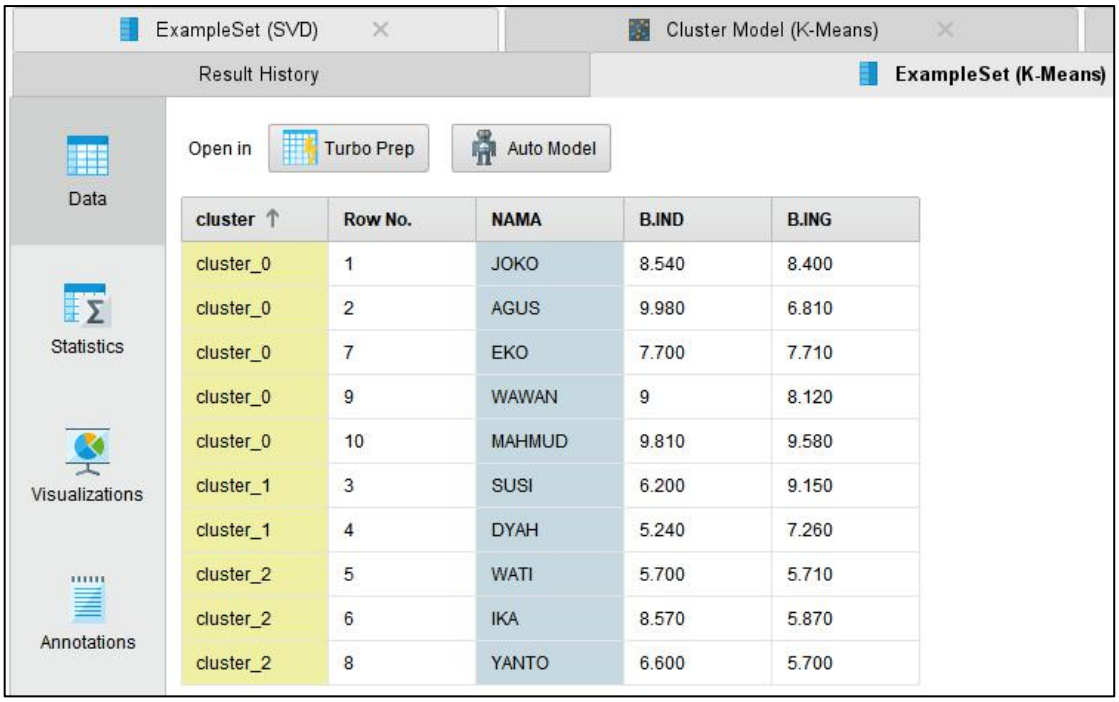
d) Cluster Model (Clustering)

i. Description

ii. Graph



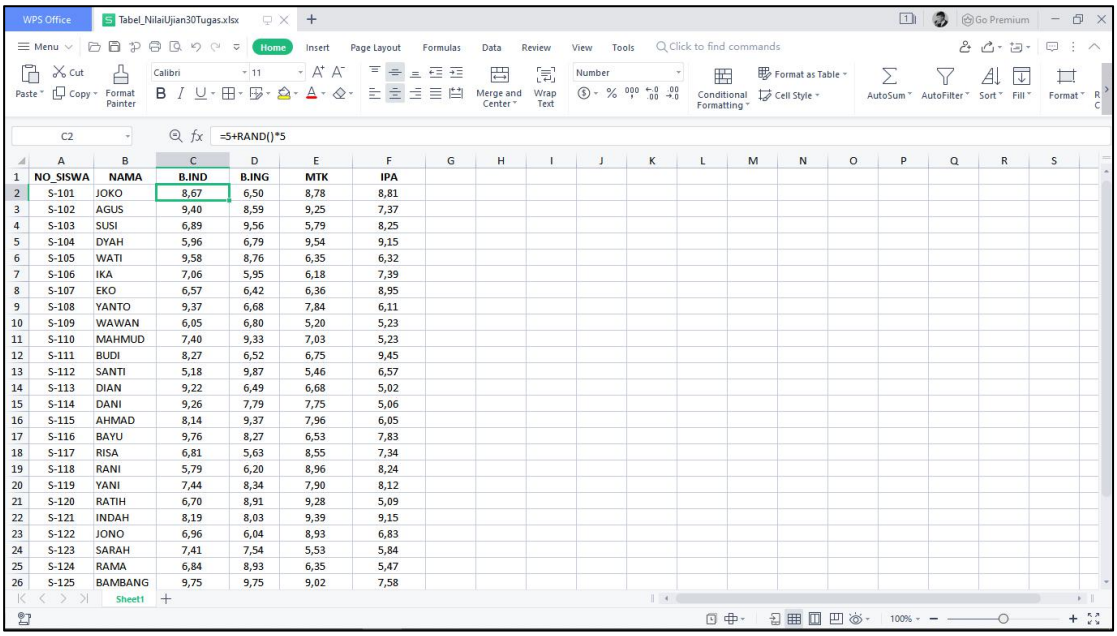
e) Hasil Algoritma K-Means



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

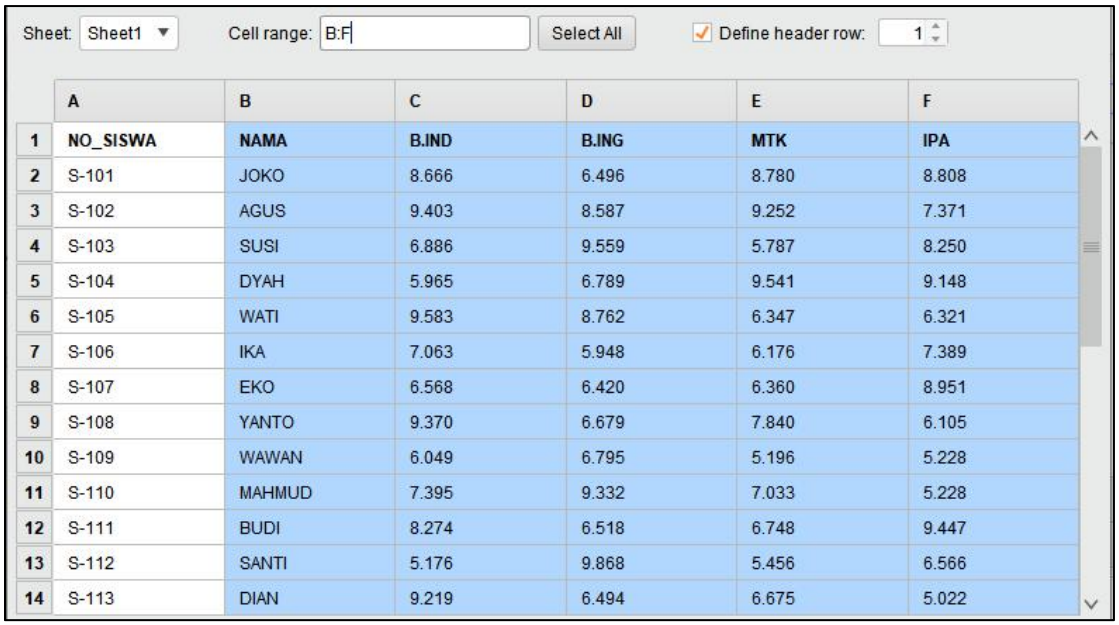
TUGAS

1. Membuat tabel Excel 30 Siswa dengan 4 mata pelajaran, pada setiap pelajaran dikasih nilai dengan rumus **=5+RAND()*5**.



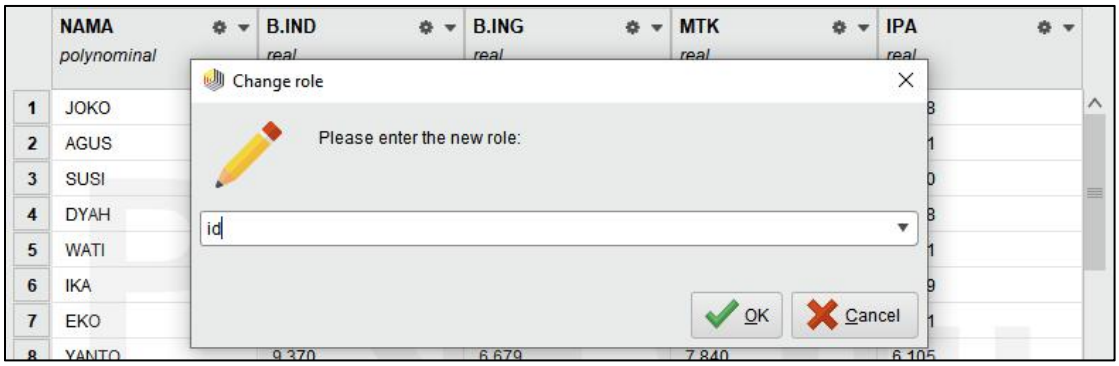
NO_SISWA	NAMA	B.IND	B.ING	MTK	IPA
S-101	JOKO	8,67	6,50	8,78	8,81
S-102	AGUS	9,40	8,59	9,25	7,37
S-103	SUSI	6,89	9,56	5,79	8,25
S-104	DYAH	5,96	6,79	9,54	9,15
S-105	WATI	9,58	8,76	6,35	6,32
S-106	IKA	7,06	5,95	6,18	7,39
S-107	EKO	6,57	6,42	6,36	8,95
S-108	YANTO	9,37	6,68	7,84	6,11
S-109	WAWAN	6,05	6,80	5,20	5,23
S-110	MAHMUD	7,40	9,33	7,03	5,23
S-111	BUDI	8,27	6,52	6,75	9,45
S-112	SANTI	5,18	9,87	5,46	6,57
S-113	DIAN	9,22	6,49	6,68	5,02
S-114	DANI	9,26	7,79	7,75	5,06
S-115	AHMAD	8,14	9,37	7,96	6,05
S-116	BAYU	9,76	8,27	6,53	7,83
S-117	RISA	6,81	5,63	8,55	7,34
S-118	RANI	5,79	6,20	8,96	8,24
S-119	YANI	7,44	8,34	7,90	8,12
S-120	RATIH	6,70	8,91	9,28	5,09
S-121	INDAH	8,19	8,03	9,39	9,15
S-122	JONO	6,96	6,04	8,93	6,83
S-123	SARAH	7,41	7,54	5,53	5,84
S-124	RAMA	6,84	8,93	6,35	5,47
S-125	BAMBANG	9,75	9,75	9,02	7,58

2. Seleksi hanya tiga bagian saja kecuali NO_SISWA.

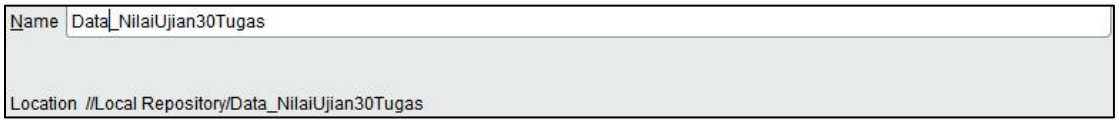

Sheet: Sheet1 Cell range: B:F Select All Define header row: 1

	A	B	C	D	E	F
1	NO_SISWA	NAMA	B.IND	B.ING	MTK	IPA
2	S-101	JOKO	8.666	6.496	8.780	8.808
3	S-102	AGUS	9.403	8.587	9.252	7.371
4	S-103	SUSI	6.886	9.559	5.787	8.250
5	S-104	DYAH	5.965	6.789	9.541	9.148
6	S-105	WATI	9.583	8.762	6.347	6.321
7	S-106	IKA	7.063	5.948	6.176	7.389
8	S-107	EKO	6.568	6.420	6.360	8.951
9	S-108	YANTO	9.370	6.679	7.840	6.105
10	S-109	WAWAN	6.049	6.795	5.196	5.228
11	S-110	MAHMUD	7.395	9.332	7.033	5.228
12	S-111	BUDI	8.274	6.518	6.748	9.447
13	S-112	SANTI	5.176	9.868	5.456	6.566
14	S-113	DIAN	9.219	6.494	6.675	5.022

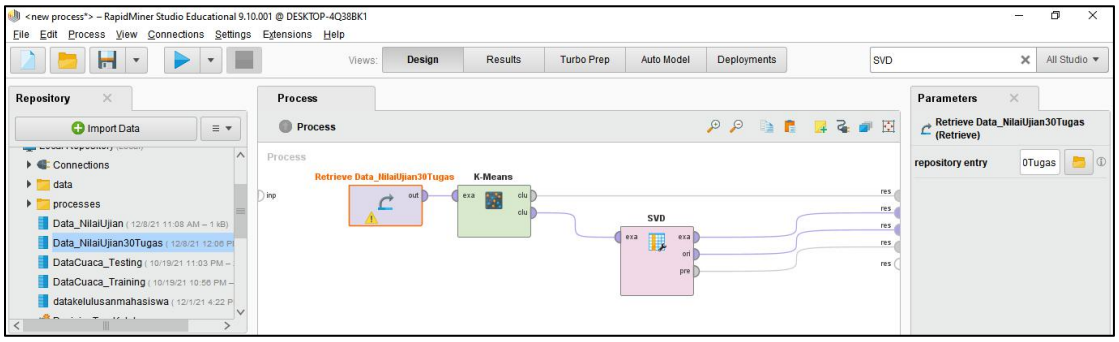
3. Ubah role nama mejadi id.



4. Simpan dengan Data_NilaiUjian30Tugas



5. Masukan operator Data_NilaiUjian30Tugas, K-Means, dan SVD, lalu sambungkan tiap port.



- Berikut adalah hasil proses Clustering dengan algoritma K-Means:

a) SVD (Singular Value Decomposition)

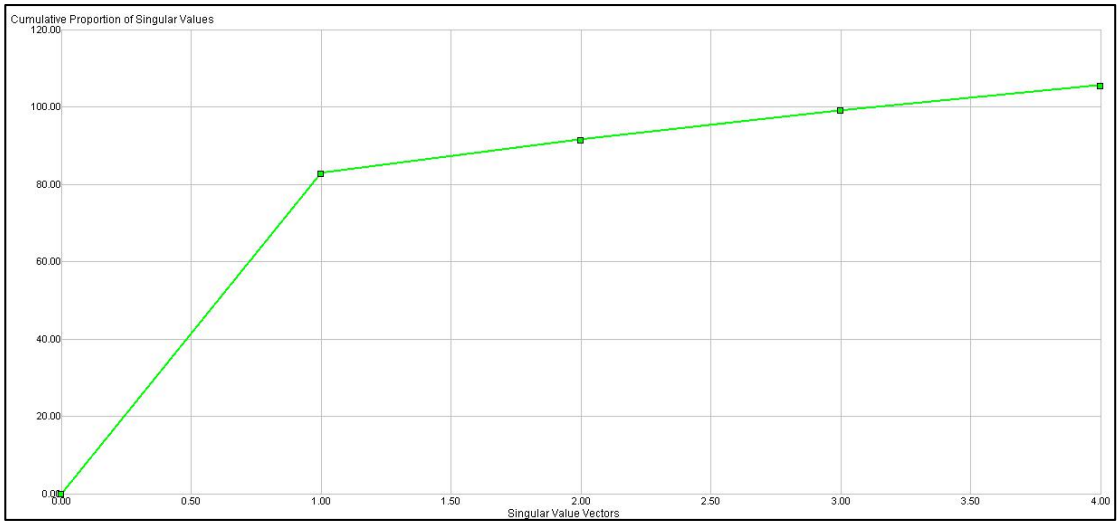
i. Nilai Eigenvalue

Component	Singular Value	Proportion of Singular Values	Cumulative Singular Values	Cumulative Proportion of Singular Val...
SVD 1	82.902	0.785	82.902	0.785
SVD 2	8.690	0.082	91.591	0.867
SVD 3	7.633	0.072	99.224	0.940
SVD 4	6.387	0.060	105.612	1.000

ii. Nilai vsd Vectors

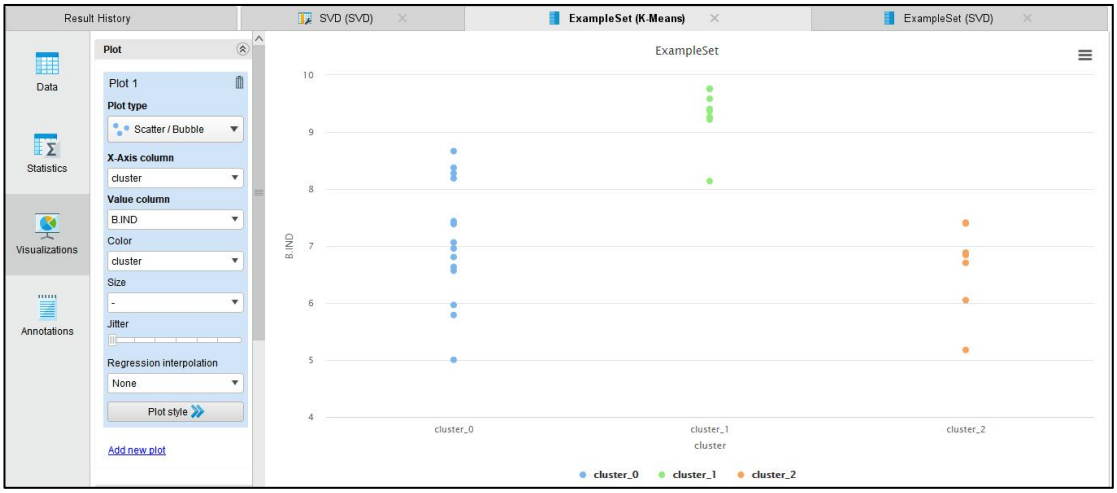
Attribute	SVD Vector 1	SVD Vector 2	SVD Vector 3
B.IND	0.508	-0.442	0.543
B.ING	0.509	-0.472	-0.698
MTK	0.500	0.226	0.396
IPA	0.483	0.729	-0.245

iii. Nilai Cumulative Variance



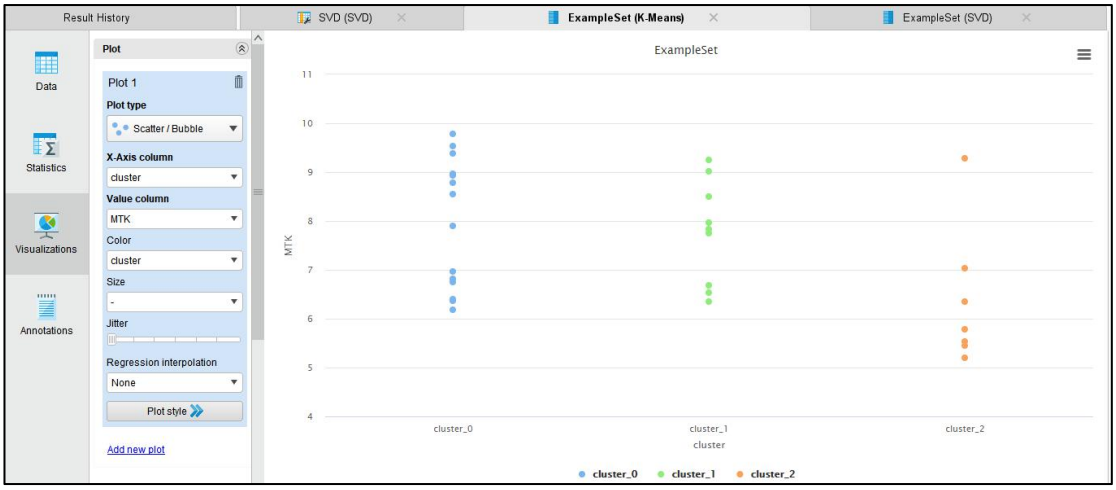
b) ExampleSet (k-Means)

i. Kelompok siswa bidang B. Indonesia.

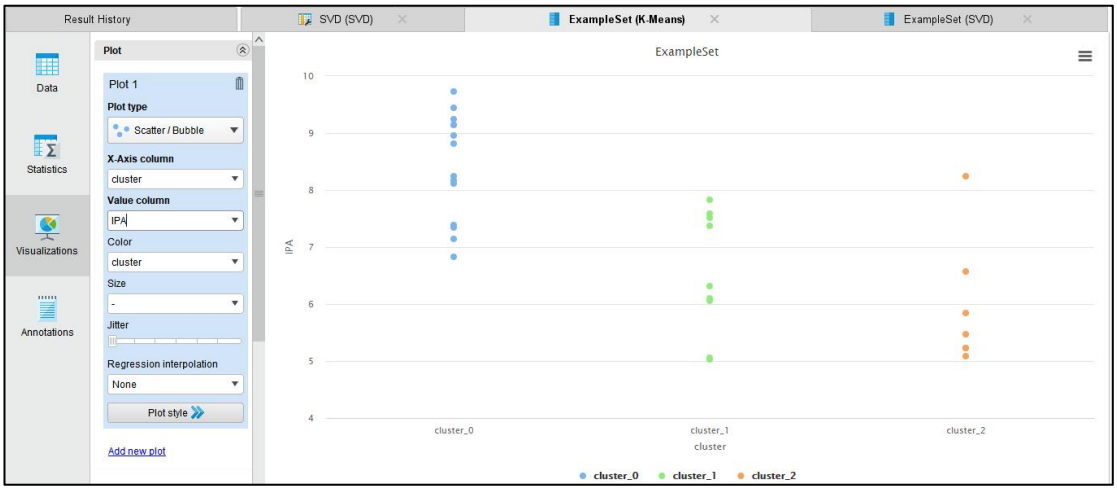


ii. Kelompok siswa bidang B. Inggris

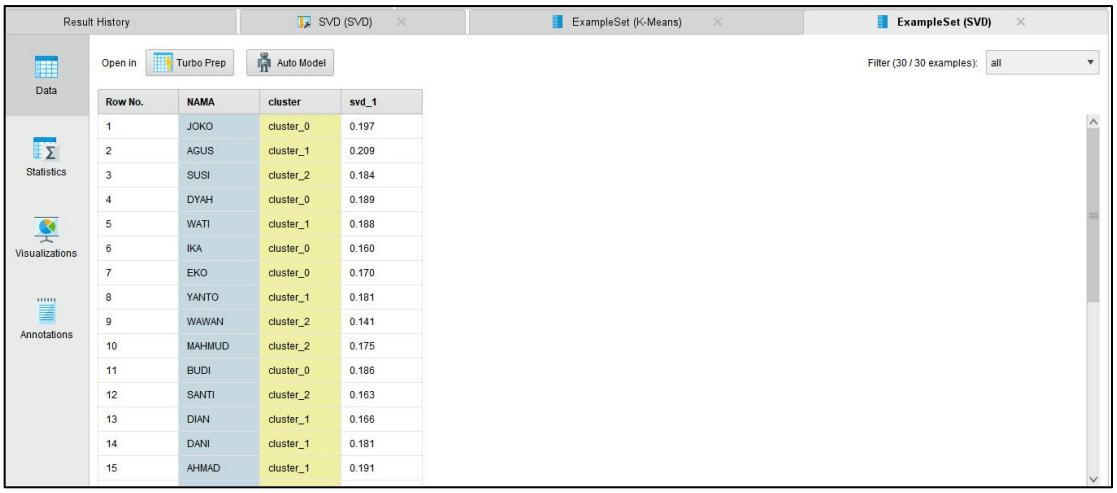
iii. Kelompok siswa bidang MTK



iv. Kelompok siswa bidang IPA



c) ExampleSet(SVD)

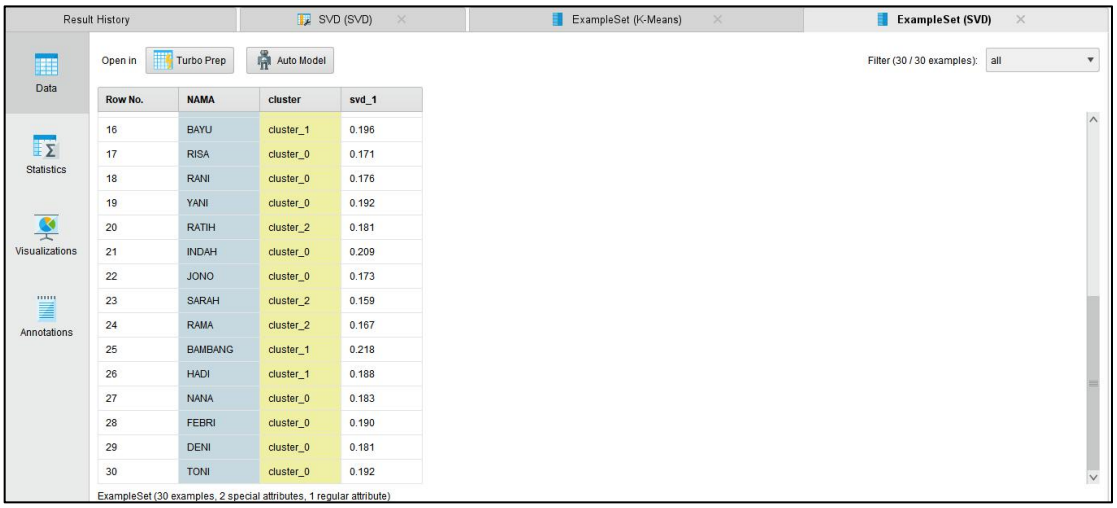


Result History | SVD (SVD) | ExampleSet (K-Means) | ExampleSet (SVD)

Data | Statistics | Visualizations | Annotations

Open in: Turbo Prep | Auto Model — Filter (30 / 30 examples): all

Row No.	NAMA	cluster	svd_1
1	JOKO	cluster_0	0.197
2	AGUS	cluster_1	0.209
3	SUSI	cluster_2	0.184
4	DYAH	cluster_0	0.189
5	WATI	cluster_1	0.188
6	IKA	cluster_0	0.160
7	EKO	cluster_0	0.170
8	YANTO	cluster_1	0.181
9	WAWAN	cluster_2	0.141
10	MAHMUD	cluster_2	0.175
11	BUDI	cluster_0	0.186
12	SANTI	cluster_2	0.163
13	DIAN	cluster_1	0.166
14	DANI	cluster_1	0.181
15	AHMAD	cluster_1	0.191



Result History | SVD (SVD) | ExampleSet (K-Means) | ExampleSet (SVD)

Data | Statistics | Visualizations | Annotations

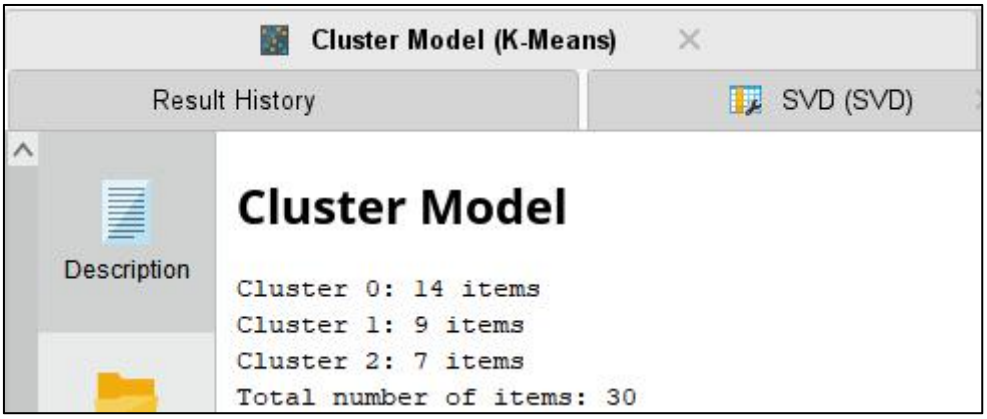
Open in: Turbo Prep | Auto Model — Filter (30 / 30 examples): all

Row No.	NAMA	cluster	svd_1
16	BAYU	cluster_1	0.196
17	RISA	cluster_0	0.171
18	RANI	cluster_0	0.176
19	YANI	cluster_0	0.192
20	RATIH	cluster_2	0.181
21	INDAH	cluster_0	0.209
22	JONO	cluster_0	0.173
23	SARAH	cluster_2	0.159
24	RAMA	cluster_2	0.167
25	BAMBANG	cluster_1	0.218
26	HADI	cluster_1	0.188
27	NANA	cluster_0	0.183
28	FEBRI	cluster_0	0.190
29	DENI	cluster_0	0.181
30	TONI	cluster_0	0.192

ExampleSet (30 examples, 2 special attributes, 1 regular attribute)

d) Cluster Model (Clustering)

i. Description



Cluster Model (K-Means)

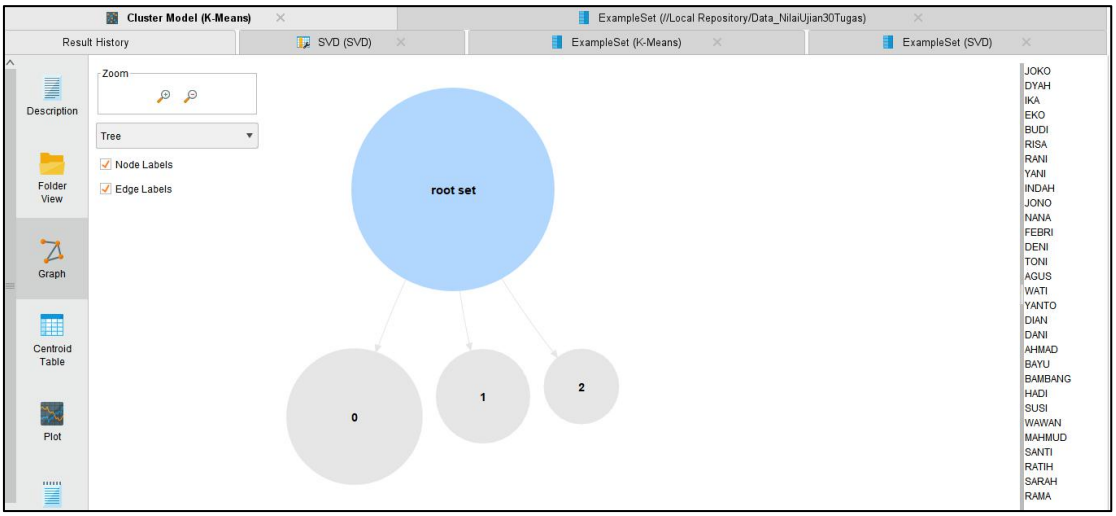
Result History | SVD (SVD)

Description

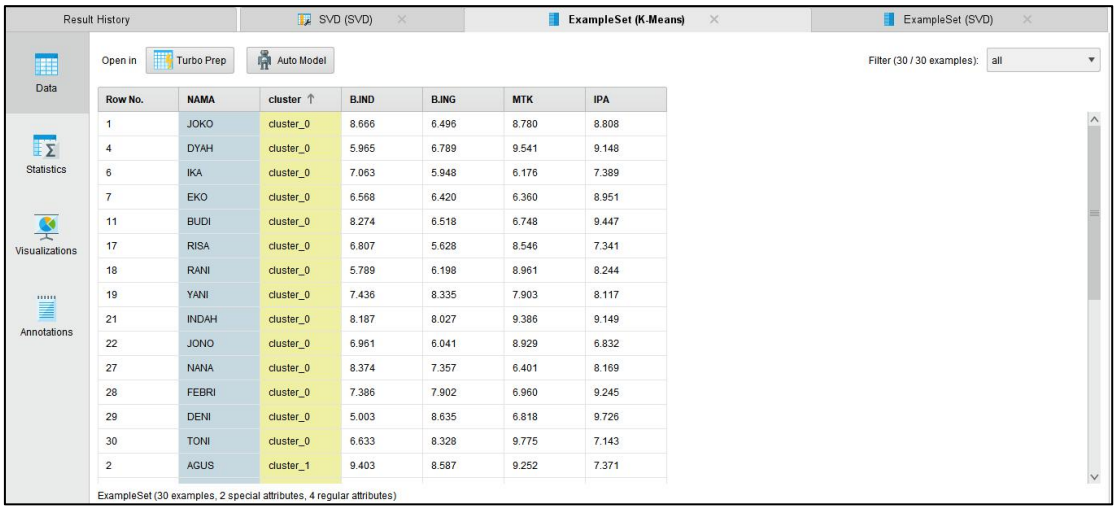
Cluster Model

Cluster 0: 14 items
Cluster 1: 9 items
Cluster 2: 7 items
Total number of items: 30

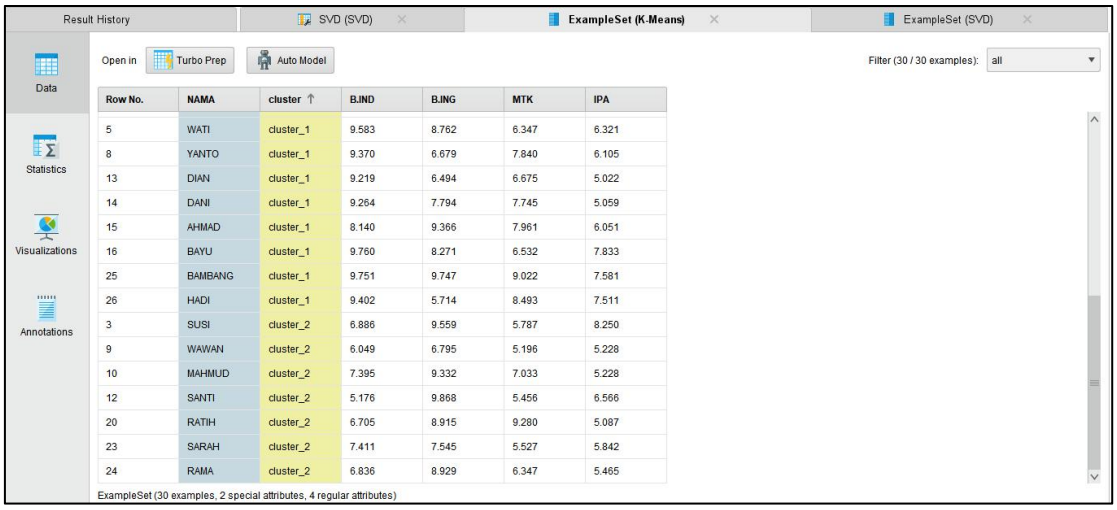
ii. Graph



e) Hasil Algoritma K-Means



Row No.	NAMA	cluster	B.IND	B.ING	MTK	IPA
1	JOKO	cluster_0	8.666	6.496	8.780	8.808
4	DYAH	cluster_0	5.965	6.789	9.541	9.148
6	IKA	cluster_0	7.063	5.948	6.176	7.389
7	EKO	cluster_0	6.568	6.420	6.360	8.951
11	BUDI	cluster_0	8.274	6.518	6.748	9.447
17	RISA	cluster_0	6.807	5.628	8.546	7.341
18	RANI	cluster_0	5.789	6.198	8.961	8.244
19	YANI	cluster_0	7.436	8.335	7.903	8.117
21	INDAH	cluster_0	8.187	8.027	9.386	9.149
22	JONO	cluster_0	6.961	6.041	8.929	6.832
27	NANA	cluster_0	8.374	7.357	6.401	8.169
28	FEBRI	cluster_0	7.386	7.902	6.960	9.245
29	DENI	cluster_0	5.003	8.635	6.818	9.726
30	TONI	cluster_0	6.633	8.328	9.775	7.143
2	AGUS	cluster_1	9.403	8.587	9.252	7.371



Row No.	NAMA	cluster	B.IND	B.ING	MTK	IPA
5	WATI	cluster_1	9.583	8.762	6.347	6.321
8	YANTO	cluster_1	9.370	6.679	7.840	6.105
13	DIAN	cluster_1	9.219	6.494	6.675	5.022
14	DANI	cluster_1	9.264	7.794	7.745	5.059
15	AHMAD	cluster_1	8.140	9.366	7.961	6.051
16	BAYU	cluster_1	9.760	8.271	6.532	7.833
25	BAMBANG	cluster_1	9.751	9.747	9.022	7.581
26	HADI	cluster_1	9.402	5.714	8.493	7.511
3	SUSI	cluster_2	6.886	9.559	5.787	8.250
9	WAWAN	cluster_2	6.049	6.795	5.196	5.228
10	MAHMUD	cluster_2	7.395	9.332	7.033	5.228
12	SANTI	cluster_2	5.176	9.868	5.456	6.566
20	RATIH	cluster_2	6.705	8.915	9.280	5.087
23	SARAH	cluster_2	7.411	7.545	5.527	5.842
24	RAMA	cluster_2	6.836	8.929	6.347	5.465