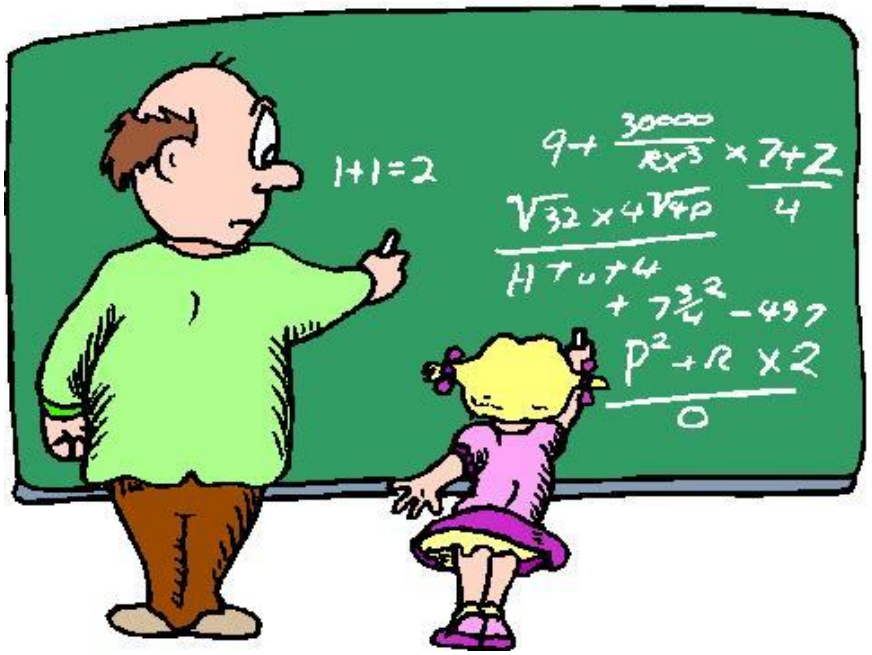
PENGELOMPOKAN BIMBINGAN BELAJAR MENGGUNAKAN METODE *CLUSTERING* DI SMA NEGERI 1 CILAKU KABUPATEN CIANJUR



SIDANG SKRIPSI
Oleh
Adry Rahadinata
10108069

Penguji 1
Andri Heryandi, S.T., M.T.

Penguji 2
Adam Mukharil Bachtiar, S.Kom.

Penguji 3
Eko Budi S., S.Kom.



SMA Negeri 1 Cilaku Kabupaten Cianjur

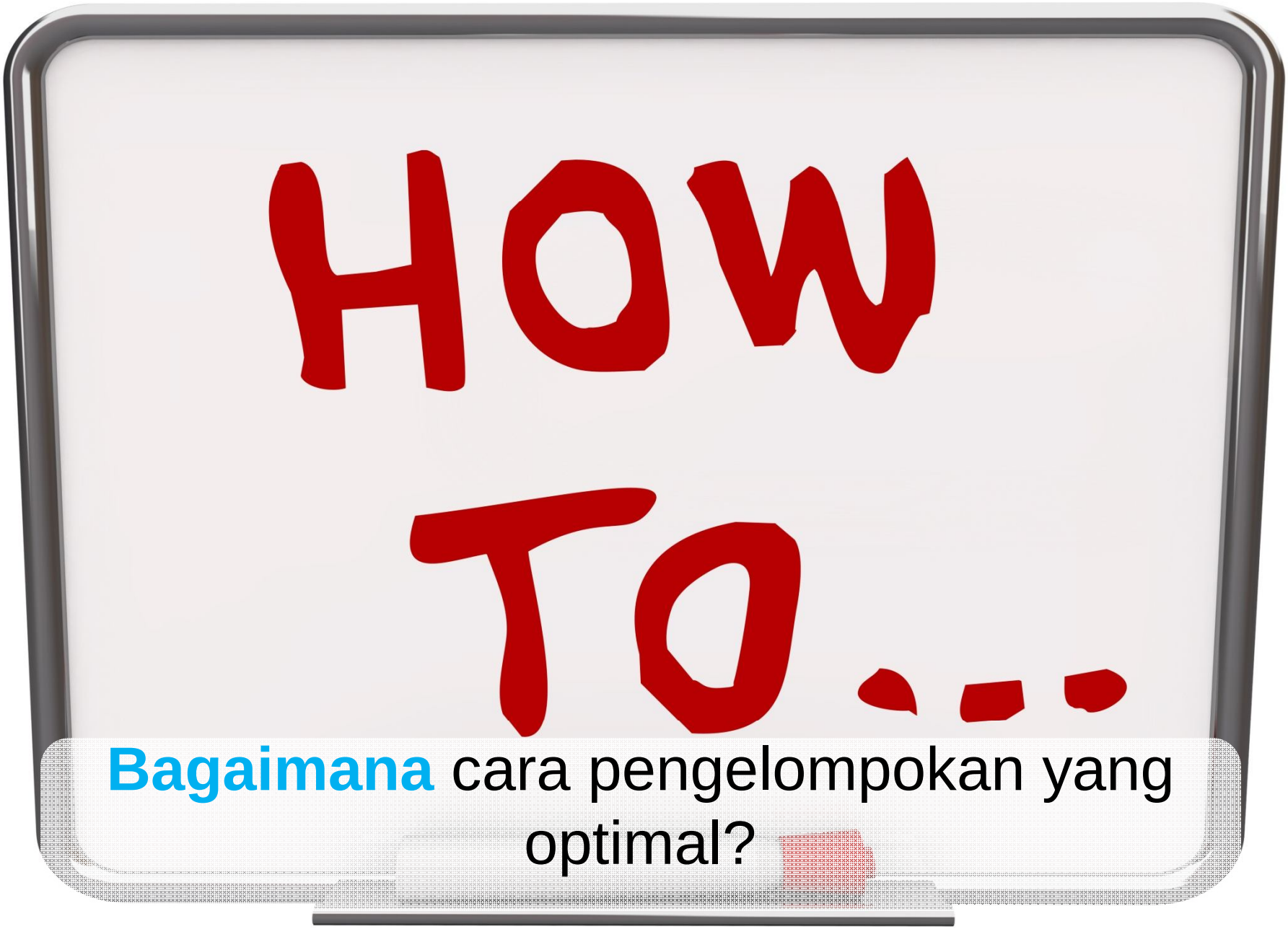
Bimbingan belajar (**pemantapan**)




Kemampuan siswa yang berbeda-beda.

Kurangnya minat siswa dalam mengikuti proses bimbingan belajar.




HOW
TO...
Bagaimana cara pengelompokan yang optimal?



Maksud

Membangun perangkat lunak



Tujuan

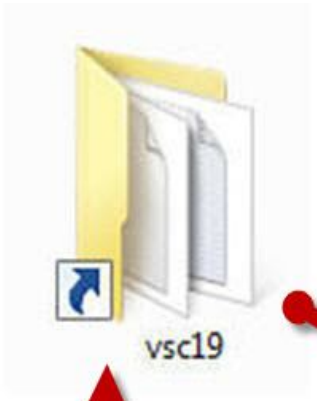
Membantu pihak sekolah

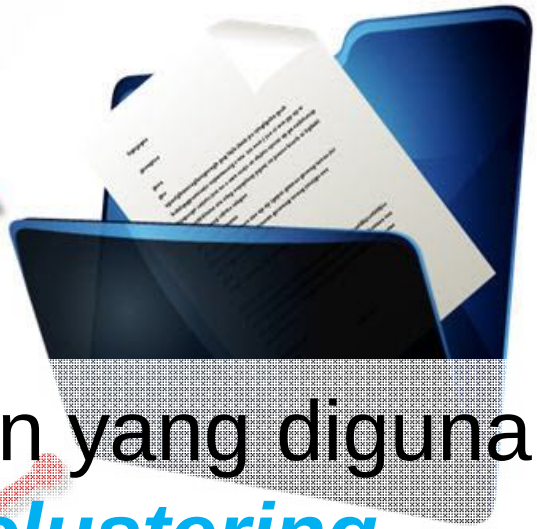
Menghasilkan perangkat lunak



Batasan Masalah

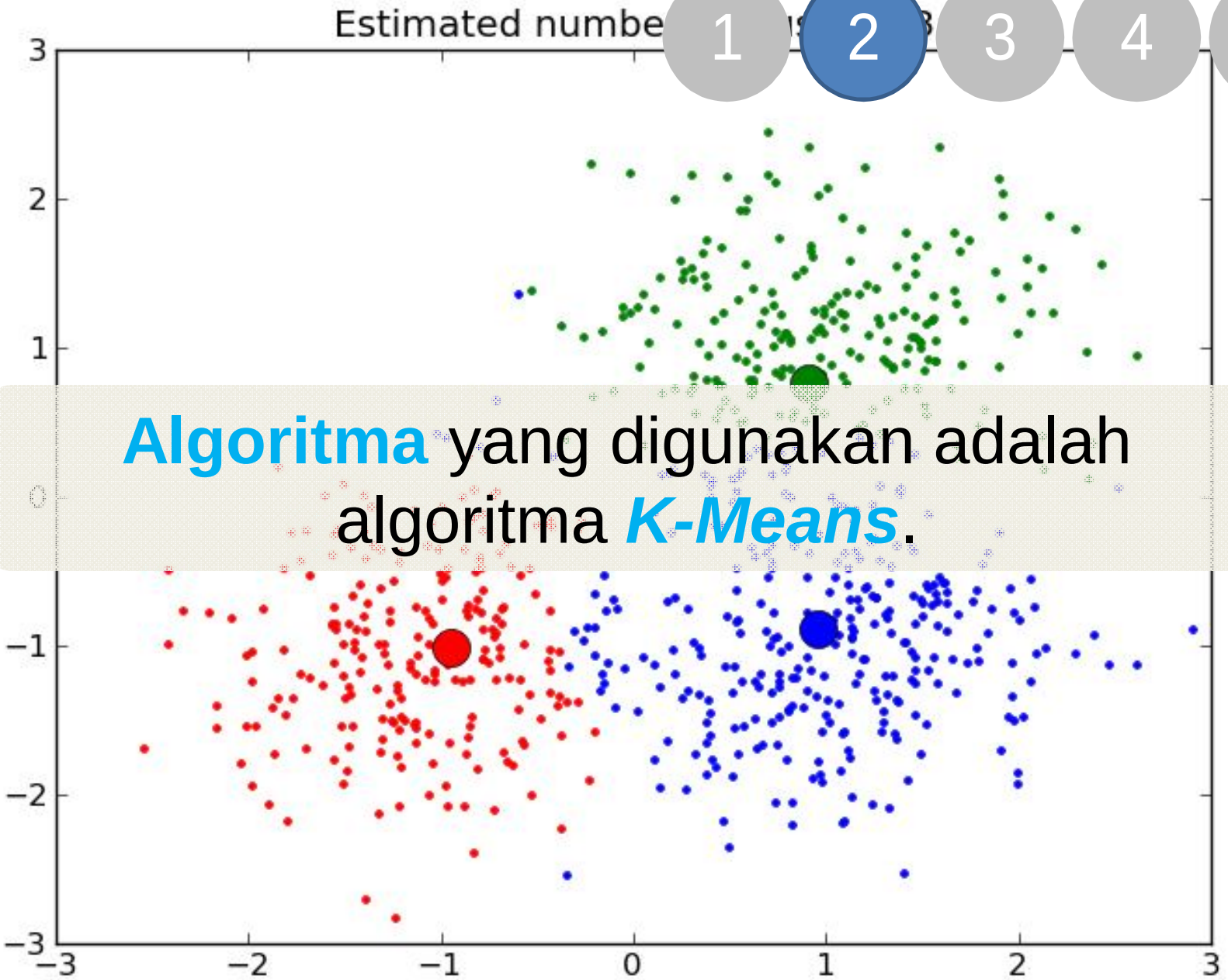
1 2 3 4 5



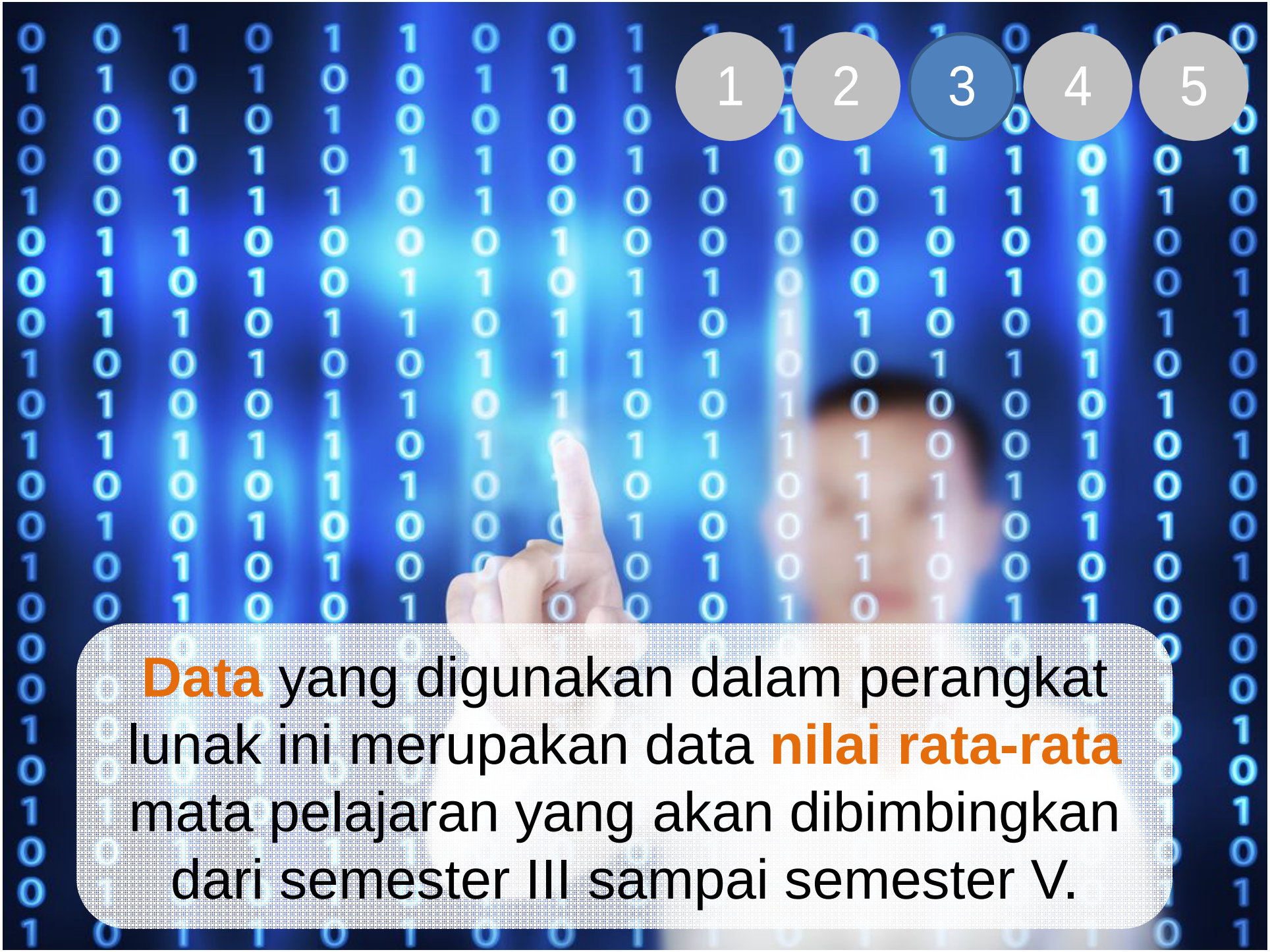




Metode pengelompokan yang digunakan adalah metode ***clustering***.



Algoritma yang digunakan adalah algoritma ***K-Means***.

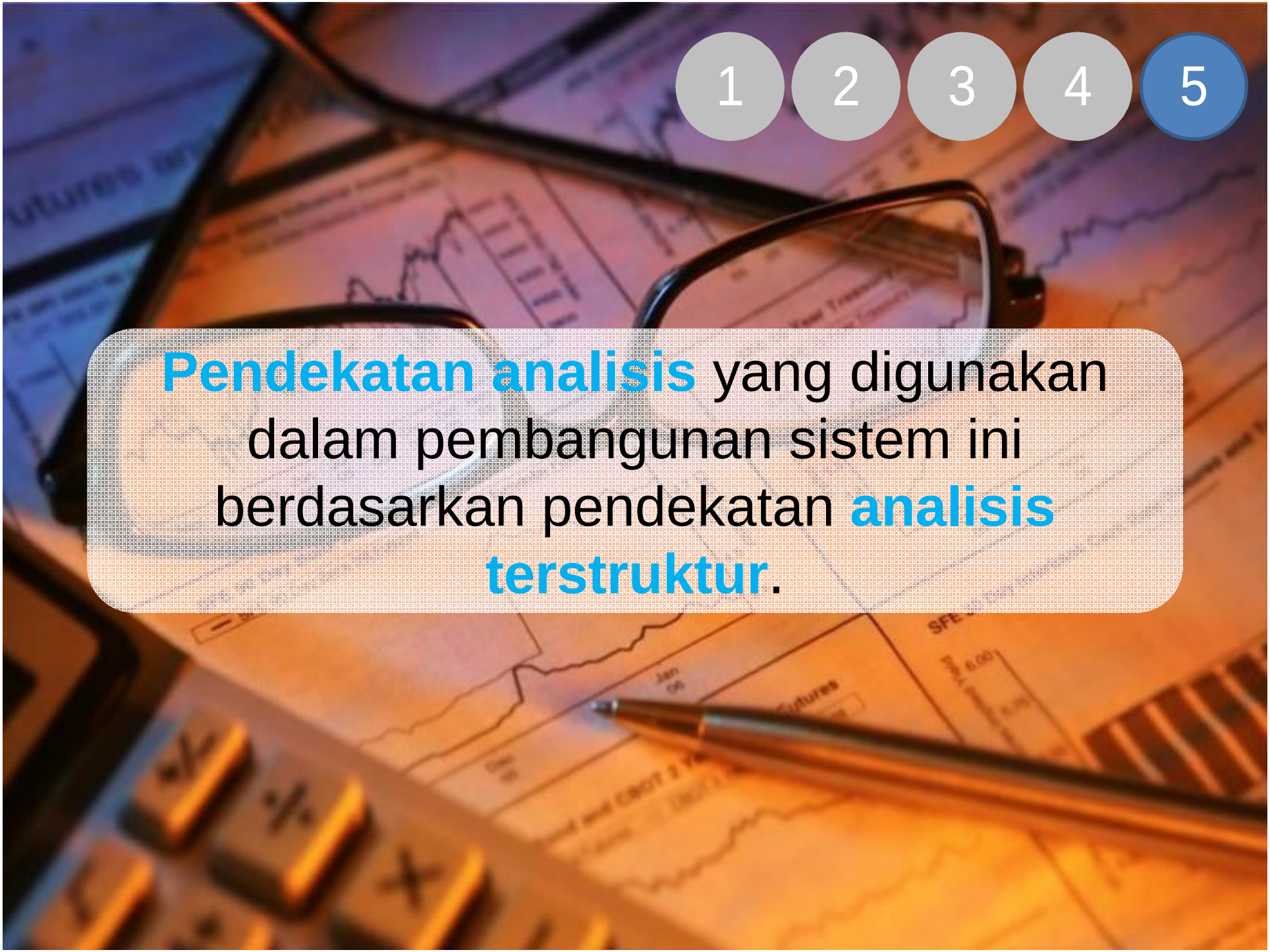


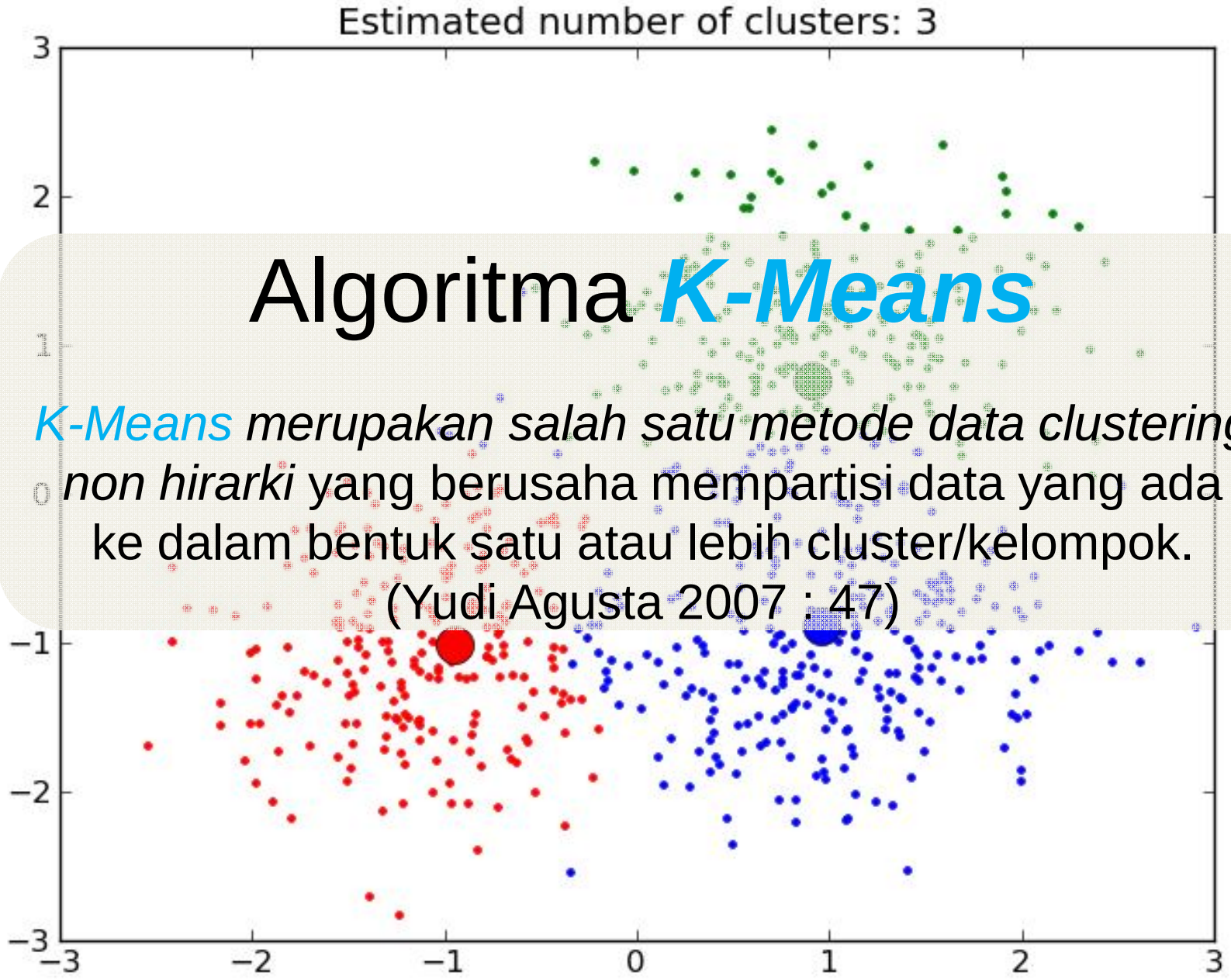
Data yang digunakan dalam perangkat lunak ini merupakan data **nilai rata-rata** mata pelajaran yang akan dibimbingkan dari semester III sampai semester V.

1 2 3 4 5



Pembagian kelompok bimbingan belajar untuk siswa jurusan IPA **dipisahkan** dengan pembagian kelompok untuk siswa jurusan IPS.


1
2
3
4
5
Pendekatan analisis yang digunakan dalam pembangunan sistem ini berdasarkan pendekatan analisis terstruktur.



Algoritma *K-Means*

K-Means merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster/kelompok.

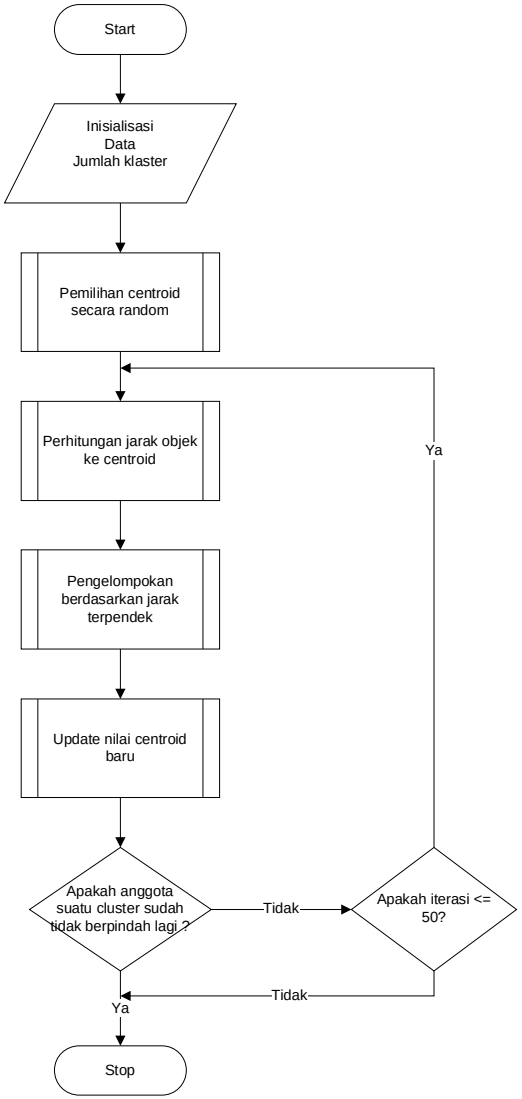
(Yudi Agusta 2007 : 47)

Algoritma ***k-means*** klastering bisa diringkas sebagai berikut :

1. Pilih jumlah klaster ***k***
2. Inisialisasi *k* pusat klaster ini bisa dilakukan dengan berbagai cara. Yang paling sering dilakukan adalah dengan cara **random**. Pusat-pusat klaster diberi nilai awal dengan angka-angka random.
3. Tempatkan setiap data/obyek ke klaster terdekat. Kedekatan dua obyek ditentukan berdasar jarak keedua obyek tersebut. Demikian juga kedekatan suatu data ke klaster tertentu ditentukan jarak antara data dengan pusat klaster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat klaster. **Jarak paling** dekat antara satu data dengan satu klaster tertentu akan menentukan suatu data masuk dalam klaster mana.
4. Hitung kembali pusat klaster dengan keanggotaan klaster yang sekarang. Pusat klaster adalah rata-rata dari semua data/ obyek dalam klaster tertentu. Jika dikehendaki bisa juga memakai median dari klaster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.
5. Tugaskan lagi setiap obyek dengan memakai pusat klaster yang baru. Jika pusat klaster sudah tidak berubah lagi, maka proses pengklasteran selesai. Atau, kembali lagi ke langkah nomor 3 sampai pusat klaster tidak berubah lagi.

(Budi Santosa 2007 : 42)

Flowchart Algoritma *K-means*



Contoh Kasus

Nis	Jurusan	B. Indonesia - smt 3	B. Inggris - smt 3	Matematika - smt 3	Fisika - smt 3	Kimia- smt 3	Biologi - smt 3	B. Indonesia - smt 4	B. Inggris - smt 4	Matematika - smt 4	Fisika - smt 4	Kimia- smt 4	Biologi - smt 4	B. Indonesia - smt 5	B. Inggris - smt 5	Matematika - smt 5	Fisika - smt 5	Kimia - smt 5	Biologi - smt 5
091010 215	A	7.50	7.10	7.20	7.10	7.20	7.80	7.80	7.50	7.20	8.30	7.30	7.80	8.10	8.30	7.80	7.80	8.00	8.50
091010 251	A	7.50	7.20	8.00	7.50	7.30	7.80	8.00	8.10	7.50	8.40	7.40	7.80	8.20	8.20	7.60	8.00	8.20	8.40
091010 072	A	7.50	7.20	7.20	7.80	7.00	7.60	7.80	8.00	7.00	7.80	7.00	7.80	8.00	8.10	8.00	7.60	7.80	8.20
091010 111	A	7.40	7.10	7.00	7.10	7.20	7.70	7.80	7.50	7.00	7.80	7.20	7.50	7.60	8.00	7.50	7.80	7.90	8.40
091010 148	A	7.40	7.00	7.00	7.50	7.10	7.70	7.70	7.50	7.00	7.50	7.10	7.50	8.00	8.20	7.50	7.60	7.60	8.10
091010 182	A	7.40	7.00	7.20	7.60	7.00	7.70	7.50	7.50	7.00	7.70	7.00	7.50	8.10	8.00	7.00	7.50	7.80	8.20
091010 257	A	7.80	7.00	7.20	7.20	7.00	7.80	7.50	7.50	7.00	8.00	7.60	7.50	8.90	8.30	7.80	7.90	7.60	8.50
091010 150	A	7.50	7.10	7.60	8.70	7.90	7.80	8.00	7.50	7.80	8.80	8.10	8.00	8.60	8.00	8.00	8.80	8.20	9.30
091010 077	A	7.80	7.00	7.50	7.50	7.80	7.70	7.80	7.50	8.50	7.60	7.80	7.80	8.40	8.00	9.00	7.70	8.00	8.30
091010 258	A	7.50	7.00	7.20	7.80	7.00	7.70	7.80	7.50	7.20	8.00	7.20	7.70	7.80	8.00	7.80	7.50	7.30	8.50
091010 292	A	7.40	7.00	7.20	7.10	7.00	7.90	7.70	7.50	7.20	7.90	7.10	7.50	8.20	7.50	7.50	7.60	7.50	8.00
091010 045	A	7.40	7.20	7.00	8.20	7.00	7.80	7.80	7.60	7.20	8.00	7.00	7.50	8.00	8.40	7.50	7.50	8.00	8.40

Dari data tersebut akan dikelompokan menjadi **tiga** kelompok berdasarkan **mata pelajaran** yang akan dibimbingkan disetiap jurusan menggunkan algoritma ***K-Means***, Dalam contoh ini pengelompokan akan dilakukan terhadap siswa berdasarkan nilai rata-rata mata pelajaran bahasa indonesia dengan perulangan maximum **tiga** kali perulangan

1

Menghitung nilai rata-rata dari tiap mata pelajaran yang hasilnya terdapat dalam tabel di bawah ini

No	Nis	B.Indonesia	B.Ingris	Matematika	Fisika	Kimia	Biologi
1	091010 215	7.80	7.63	7.40	7.73	7.50	8.03
2	091010 251	7.90	7.83	7.70	7.97	7.63	8.00
3	091010 072	7.77	7.77	7.40	7.73	7.27	7.87
4	091010 111	7.60	7.53	7.17	7.57	7.43	7.87
5	091010 148	7.70	7.57	7.17	7.53	7.27	7.77
6	091010 182	7.67	7.50	7.07	7.60	7.27	7.80
7	091010 257	8.07	7.60	7.33	7.70	7.40	7.93
8	091010 150	8.03	7.53	7.80	8.77	8.07	8.37
9	091010 077	8.00	7.50	8.33	7.60	7.87	7.93
10	091010 258	7.70	7.50	7.40	7.77	7.17	7.97
11	091010 292	7.77	7.33	7.30	7.53	7.20	7.80
12	091010 045	7.73	7.73	7.23	7.90	7.33	7.90

2

Dari tabel tersebut kita tentukan pusat awal *cluster* untuk *clustering* ‘B.Indonesia’ secara acak. Pusat awal kita ambil dari data yang ada pada kolom ‘B.Indonesia’ dengan menggunakan fungsi random, misalnya nilai yang didapat adalah sebagai berikut :

c1 = 7.60 yang diambil dari **data ke-4**

c2 = 7.77 yang diambil dari **data ke-11**

c3 = 8.07 yang diambil dari **data ke-7**

3

Menghitung **jarak** dari semua data ke tiap titik **pusat *cluster*** pertama menggunakan *Euclidean Distace* dengan rumus sebagai berikut:

$$D(x_1, x_2) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^{p} (x_{2j} - x_{1j})^2}$$

1. Menghitung jarak data ke-1 dengan *centroid* pertama (c1)

$$D_{11} = \sqrt{(7.80 - 7.60)^2} = 0.20$$

2. Menghitung jarak data ke-1 dengan *centroid* kedua (c1)

$$D_{12} = \sqrt{(7.80 - 7.77)^2} = 0.03$$

3. Menghitung jarak data ke-1 dengan *centroid* ketiga(c1)

$$D_{13} = \sqrt{(7.80 - 8.07)^2} = 0.27$$

4

Perhitungan jarak setiap data dan posisi *cluster* pada iterasi **pertama**.

No	Nis	D_{i1}	D_{i2}	D_{i3}	C1	C2	C3
1	091010215	0.20	0.03	0.27		*	
2	091010251	0.30	0.13	0.17		*	
3	091010072	0.17	0.00	0.30		*	
4	091010111	0.00	0.17	0.47	*		
5	091010148	0.10	0.07	0.37		*	
6	091010182	0.07	0.10	0.40	*		
7	091010257	0.47	0.30	0.00			*
8	091010150	0.43	0.26	0.04			*
9	091010077	0.40	0.23	0.07			*
10	091010258	0.10	0.07	0.37		*	
11	091010292	0.17	0.00	0.30		*	
12	091010045	0.13	0.04	0.34		*	

Menentukan pusat *cluster* **baru** dengan cara menghitung rata-rata dari data yang ada di masing-masing *cluster.*

5

Perhitungan pusat *cluster* baru pertama
Karena *cluster* pertama (C1) memiliki dua anggota maka :

$$c1 = \left(\frac{7.60 \mid 7.67}{2}\right) = 7.63$$

Perhitungan pusat *cluster* baru kedua
Karena *cluster* kedua (C2) memiliki tujuh anggota maka :

$$c2 = \left(\frac{7.80 + 7.90 + 7.77 + 7.70 + 7.70 + 7.77 + 7.73}{7}\right) = 7.77$$

Perhitungan pusat *cluster* baru ketiga
Karena *cluster* ketiga (C3) memiliki tiga anggota maka :

$$c1 = \left(\frac{8.07 + 8.03 + 8.00}{3}\right) = 8.03$$

6

Mengulangi langkah **tiga**, **empat** dan **lima** sampai pusat *cluster* tidak **berubah** lagi atau perubahan nilai pada jumlah iterasi **melebihi** nilai ***perulangan maksimun*** yang ditentukan.

Perhitungan jarak setiap data dan posisi *cluster* pada iterasi **kedua**.

No	Nis	D_{i1}	D_{i2}	D_{i3}	C1	C2	C3
1	091010215	0.17	0.03	0.23		*	
2	091010251	0.27	0.13	0.13		*	*
3	091010072	0.14	0.00	0.26		*	
4	091010111	0.03	0.17	0.43	*		
5	091010148	0.07	0.07	0.33	*	*	
6	091010182	0.04	0.10	0.36	*		
7	091010257	0.44	0.30	0.04			*
8	091010150	0.40	0.26	0.00			*
9	091010077	0.37	0.23	0.03			*
10	091010258	0.07	0.07	0.33	*	*	
11	091010292	0.14	0.00	0.26		*	
12	091010045	0.10	0.04	0.30		*	

Perhitungan jarak setiap data dan posisi *cluster* pada iterasi **ketiga**.

No	Nis	D_{i1}	D_{i2}	D_{i3}	C1	C2	C3
1	091010215	0.13	0.03	0.20		*	
2	091010251	0.23	0.13	0.10			*
3	091010072	0.10	0.00	0.23		*	
4	091010111	0.07	0.17	0.40	*		
5	091010148	0.03	0.07	0.30	*		
6	091010182	0.00	0.10	0.33	*		
7	091010257	0.40	0.30	0.07			*
8	091010150	0.36	0.26	0.03			*
9	091010077	0.33	0.23	0.00			*
10	091010258	0.03	0.07	0.30	*		
11	091010292	0.10	0.00	0.23		*	
12	091010045	0.06	0.04	0.27		*	

Perulangan **dihentikan** karena dalam perulangan yang **ketiga** sudah tidak ada lagi anggota suatu *cluster* dari perulangan sebelumnya yang **berpindah** *cluster*

ERD

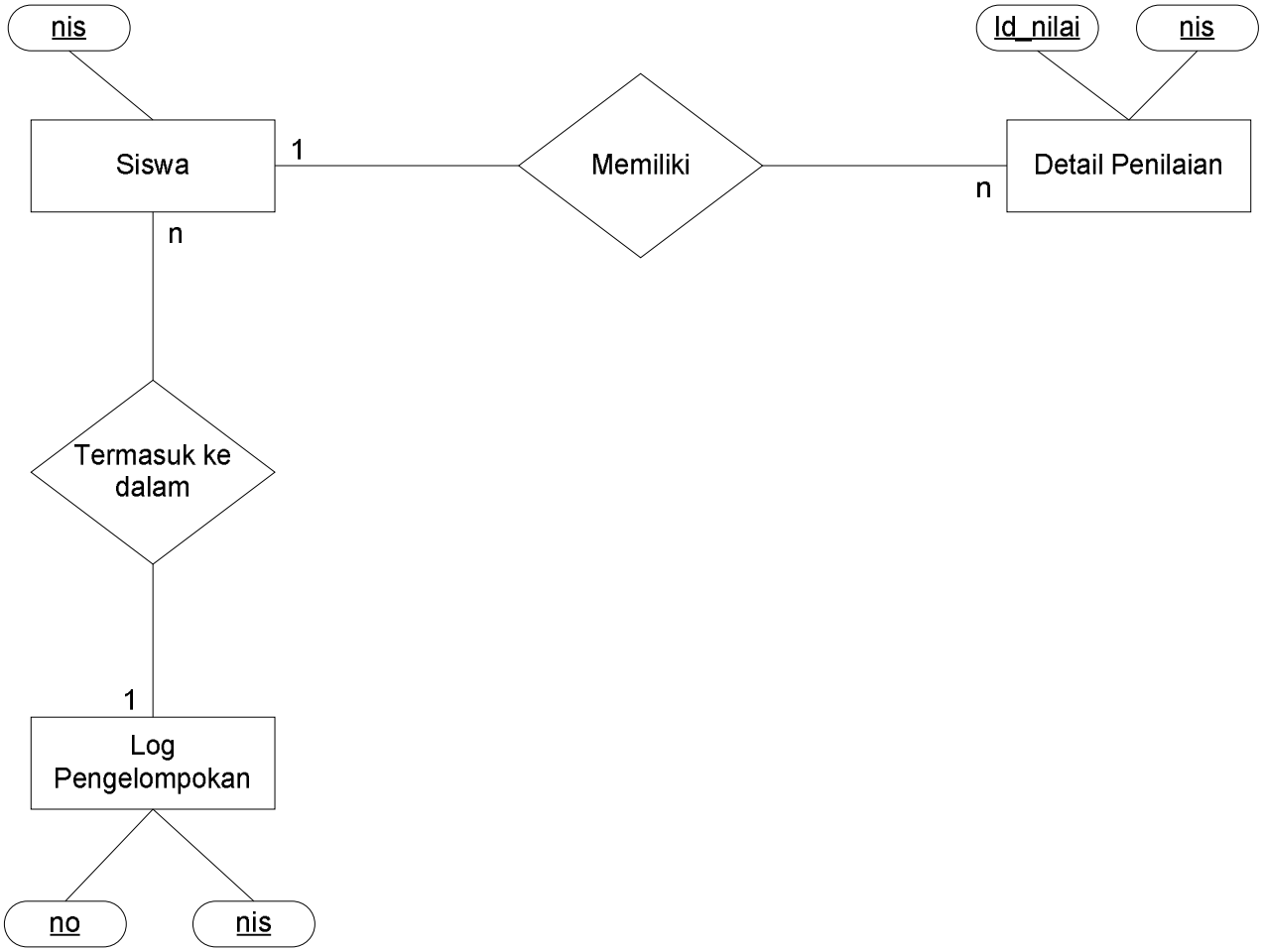
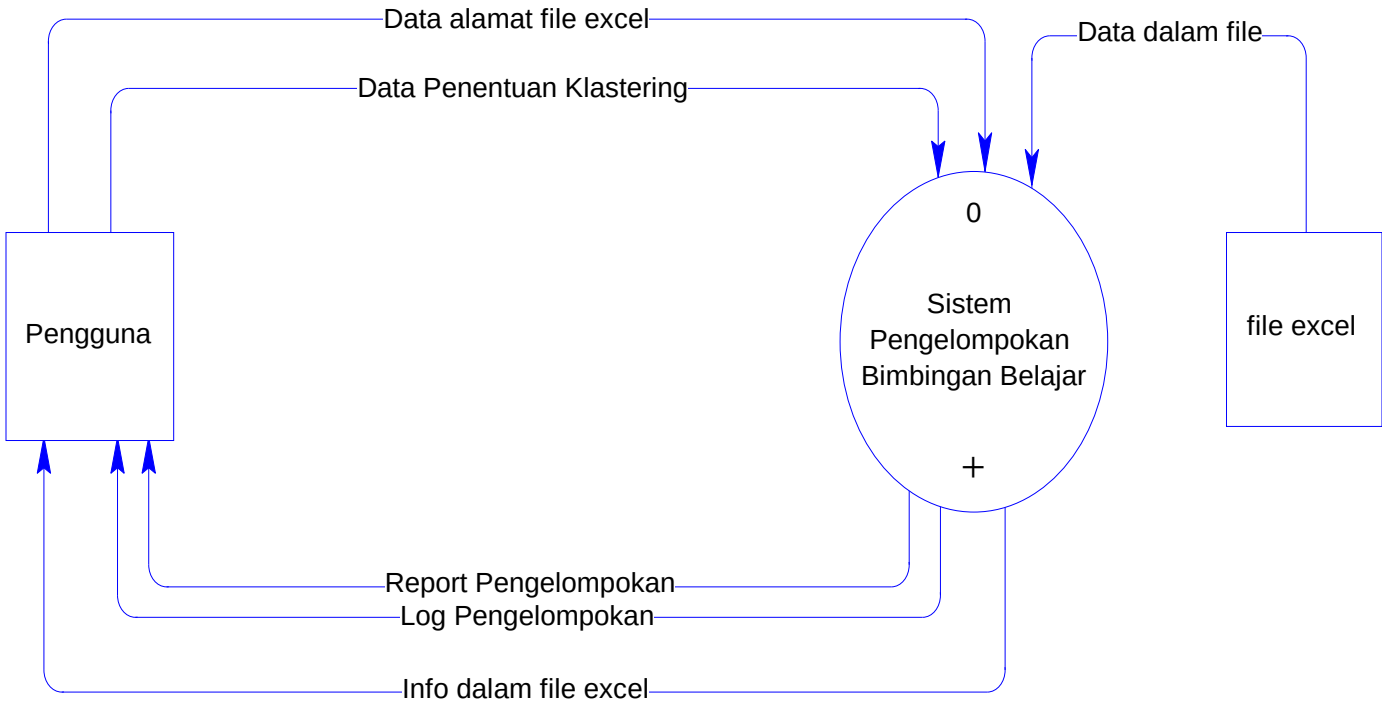
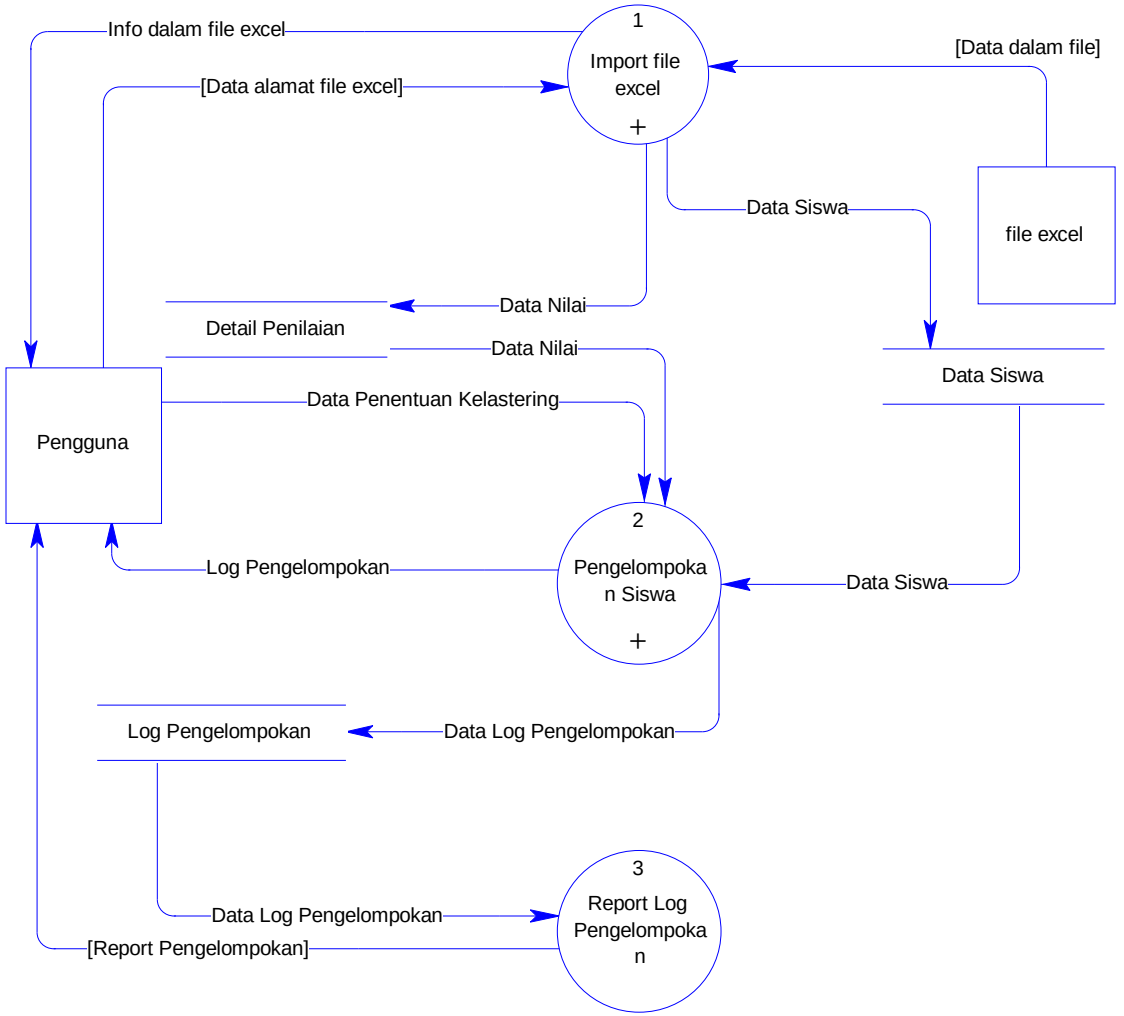

nis
Siswa
1
Memiliki
n
Detail Penilaian
Id_nilai
nis
n
Termasuk ke dalam
1
Log Pengelompokan
no
nis

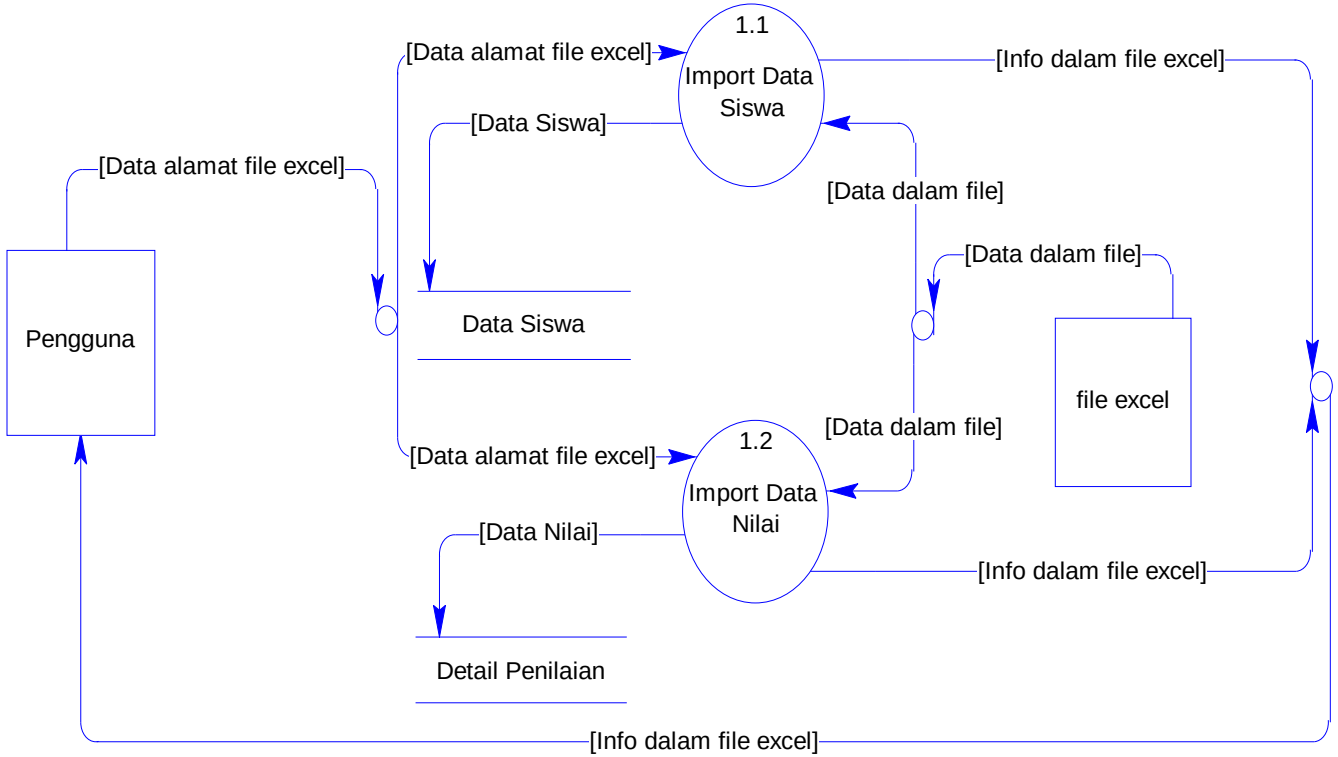
Diagram Konteks



DFD Level 1


1
Import file excel
+
Info dalam file excel
[Data alamat file excel]
[Data dalam file]
file excel
Data Siswa
Data Nilai
Detail Penilaian
Data Nilai
Data Siswa
Pengguna
Data Penentuan Kelastering
2
Pengelompokan Siswa
+
Log Pengelompokan
Data Siswa
Log Pengelompokan
Data Log Pengelompokan
3
Report Log Pengelompokan
Data Log Pengelompokan
[Report Pengelompokan]

DFD Level 2 Import File Excel



DFD Level 2 Pegelompokan Siswa

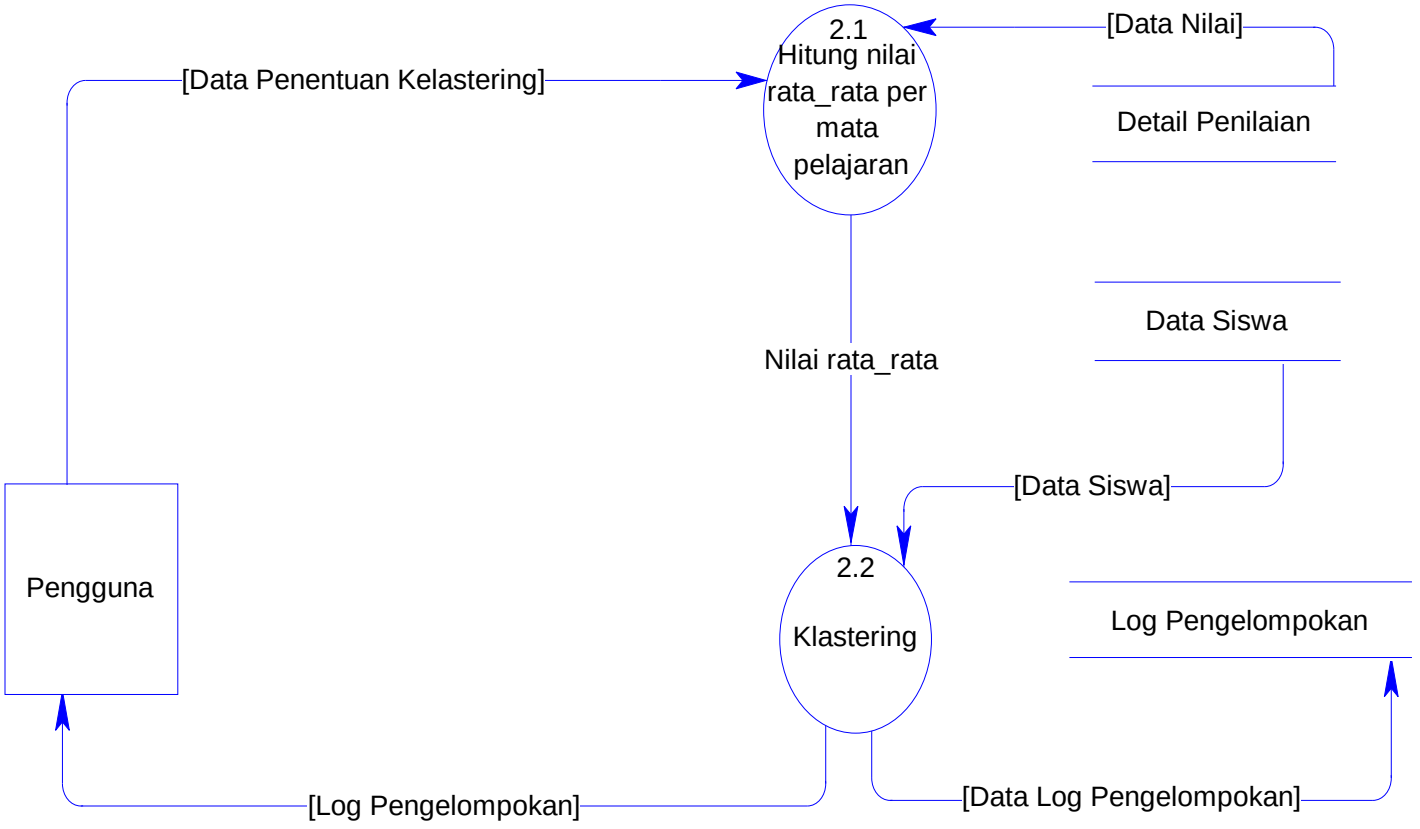
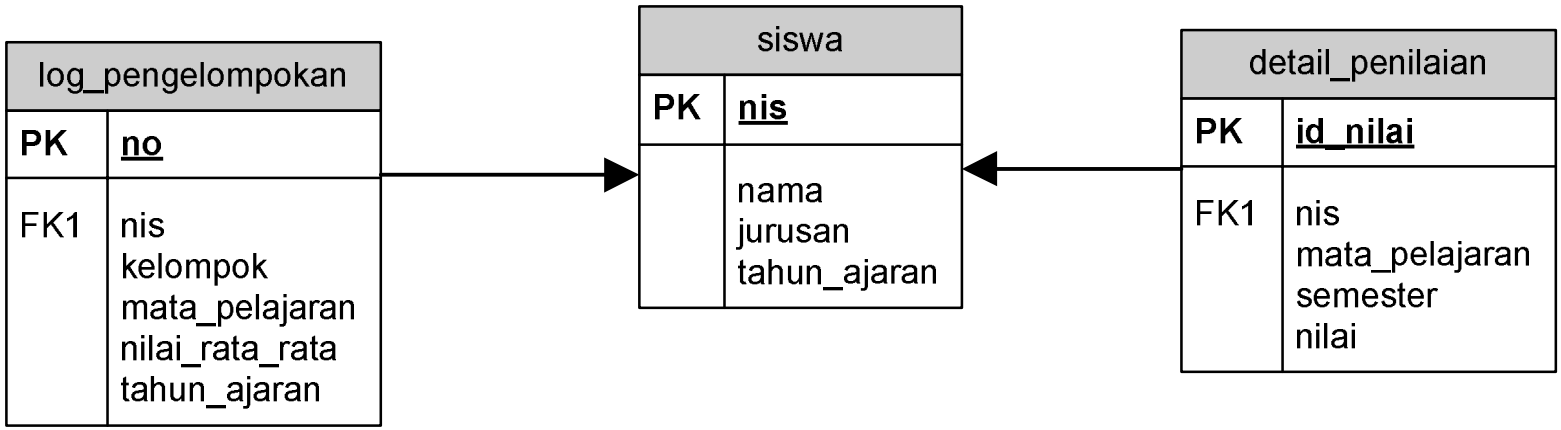
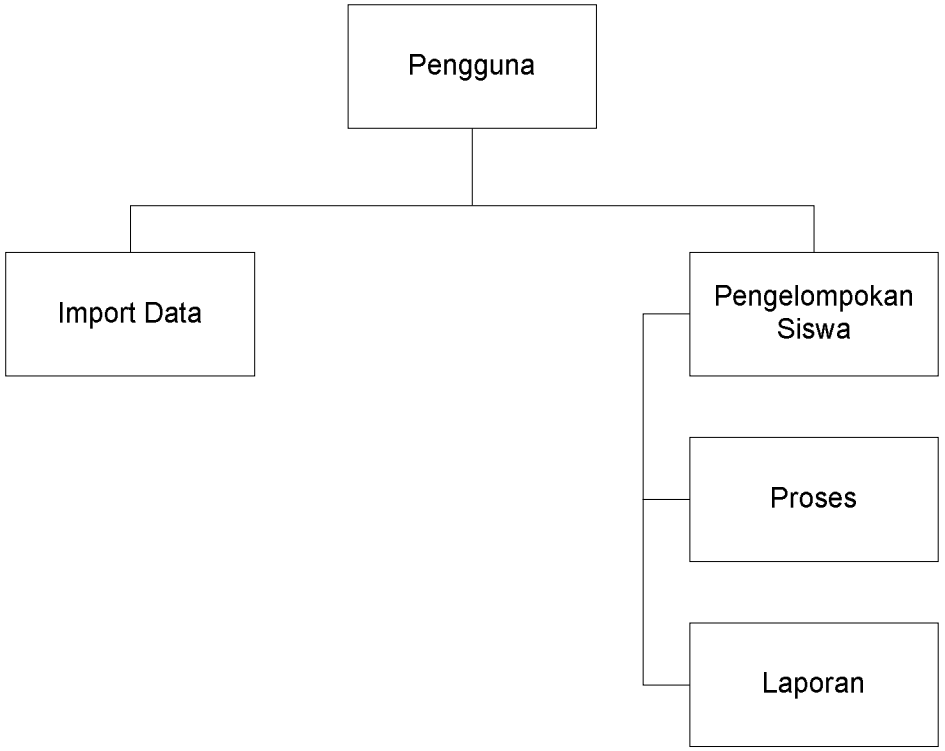


Diagram Relasi

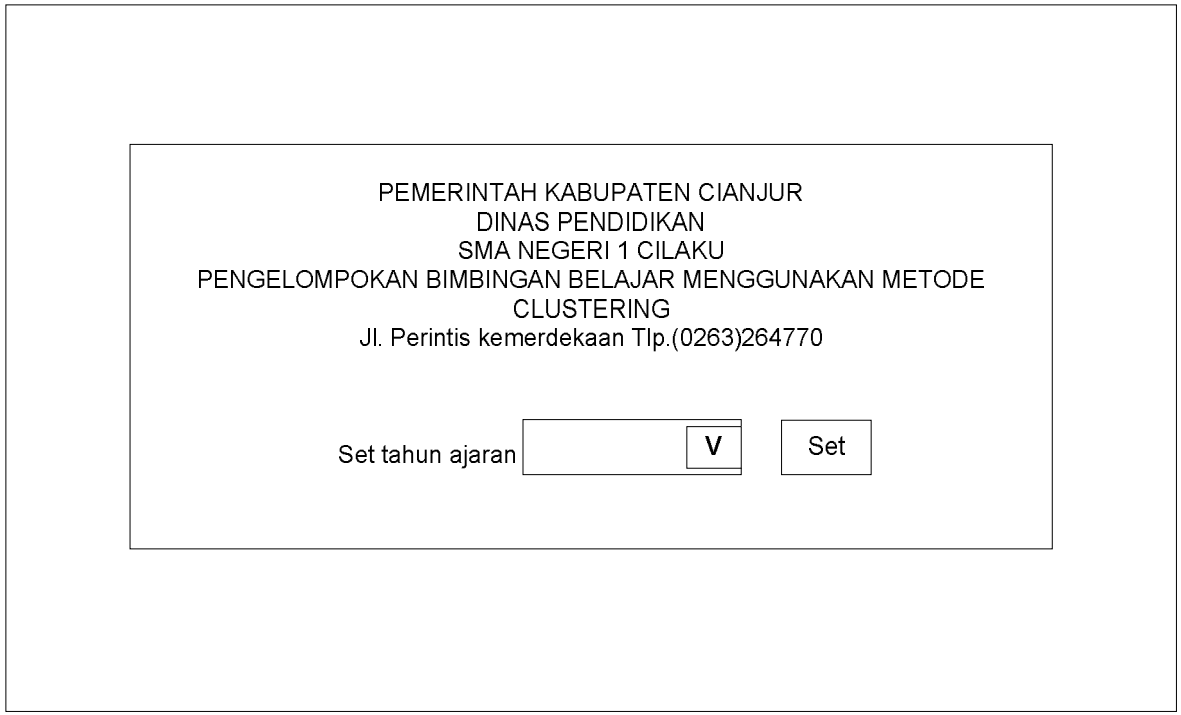


Struktur Menu



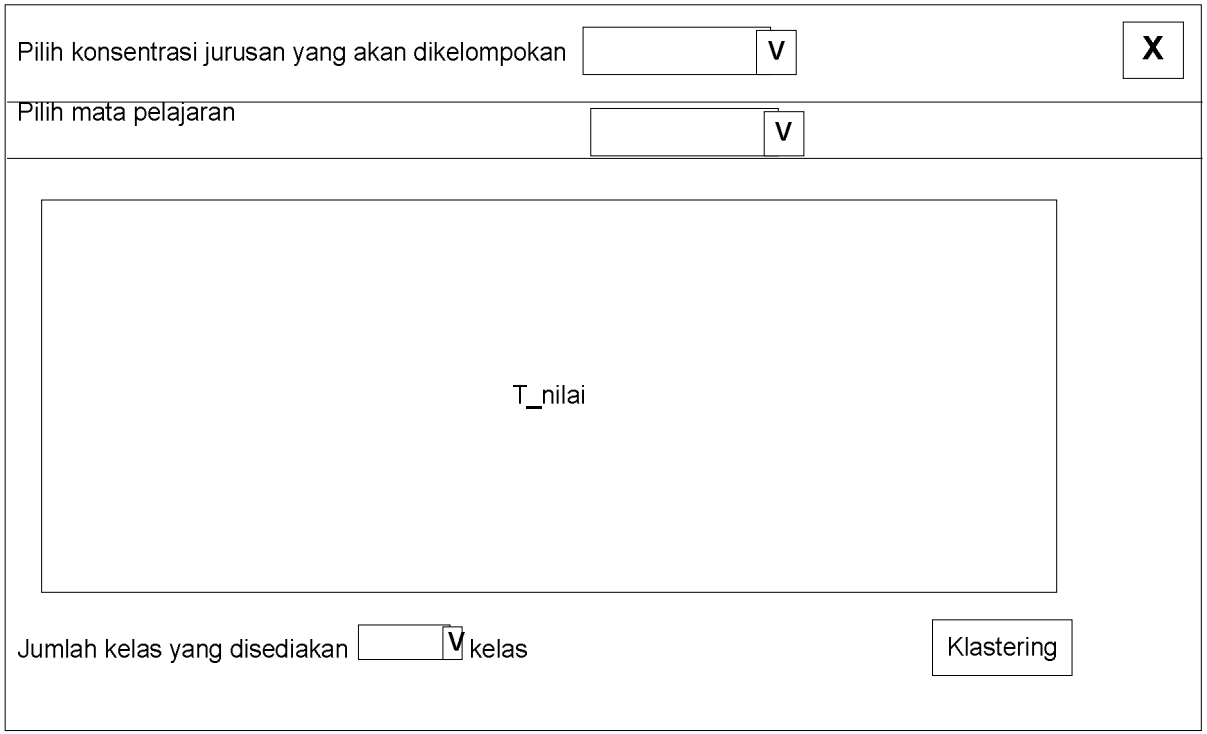


Perancangan Antarmuka



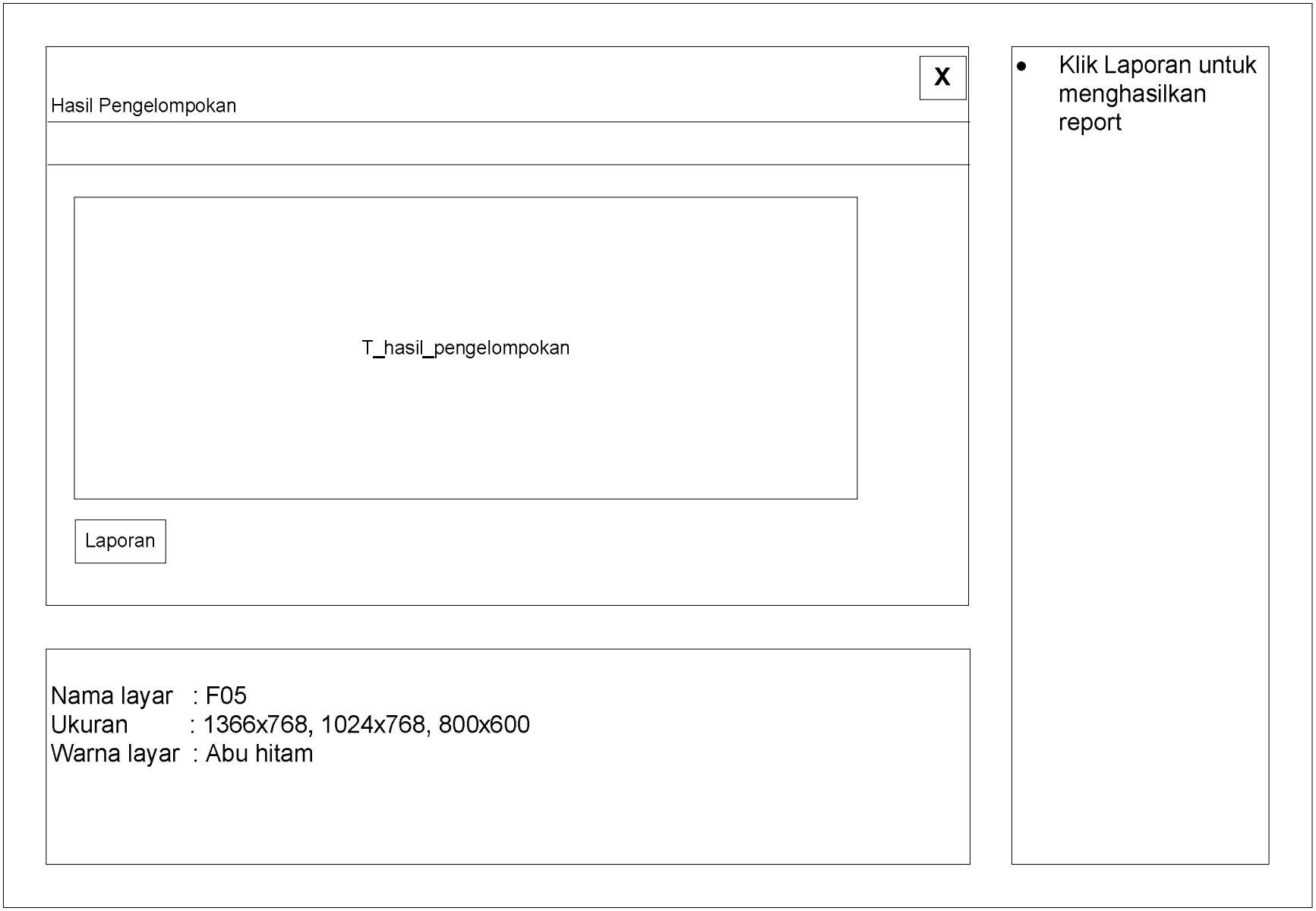
- Klik Set Untuk menentukan tahun ajaran, menuju F02. apabila gagal muncul P01.

Nama layar : F01
Ukuran : 1366x768, 1024x768, 800x600
Warna layar : Abu hitam



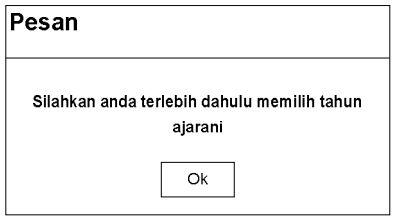
Nama layar : F05
Ukuran : 1366x768, 1024x768, 800x600
Warna layar : Abu hitam

- Klik Combobox pilih jurusan untuk melmilih konsentrasi jurusan siswa yang akan dikelompokan
- Klik Combobox pilih mata pelajaran untuk mensorting data berdasarkan mata nilai mata pelajaran yang akan dikelompokan
- Klik proses menuju F06 apabila nilai belum dirata-ratakan dan jumlah kelas belum diinputkan maka menuju P06.


X
Hasil Pengelompokan
T_hasil_pengelompokan
Laporan
Nama layar : F05
Ukuran : 1366x768, 1024x768, 800x600
Warna layar : Abu hitam
• Klik Laporan untuk menghasilkan report

Perancangan Pesan

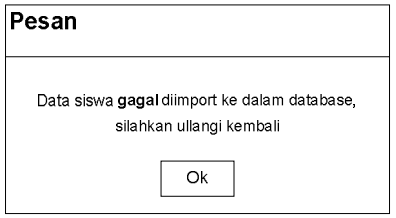
P01



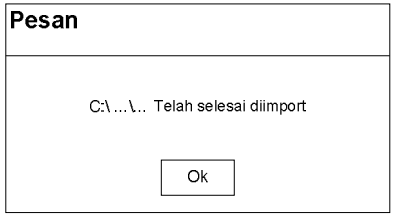
P02



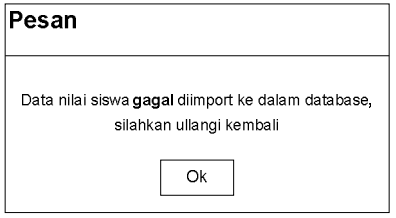
P03



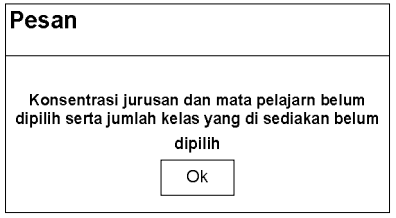
P04



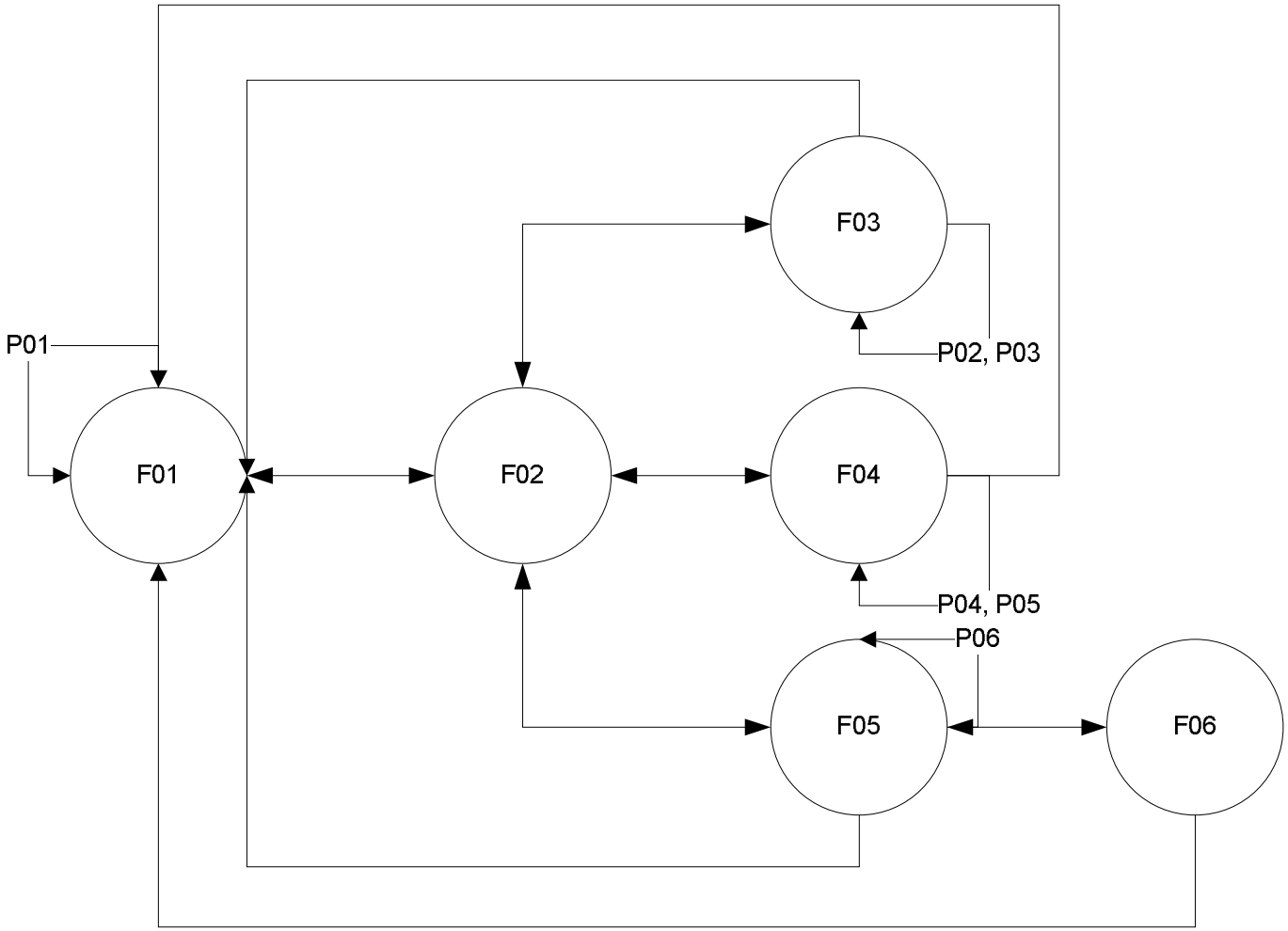
P05

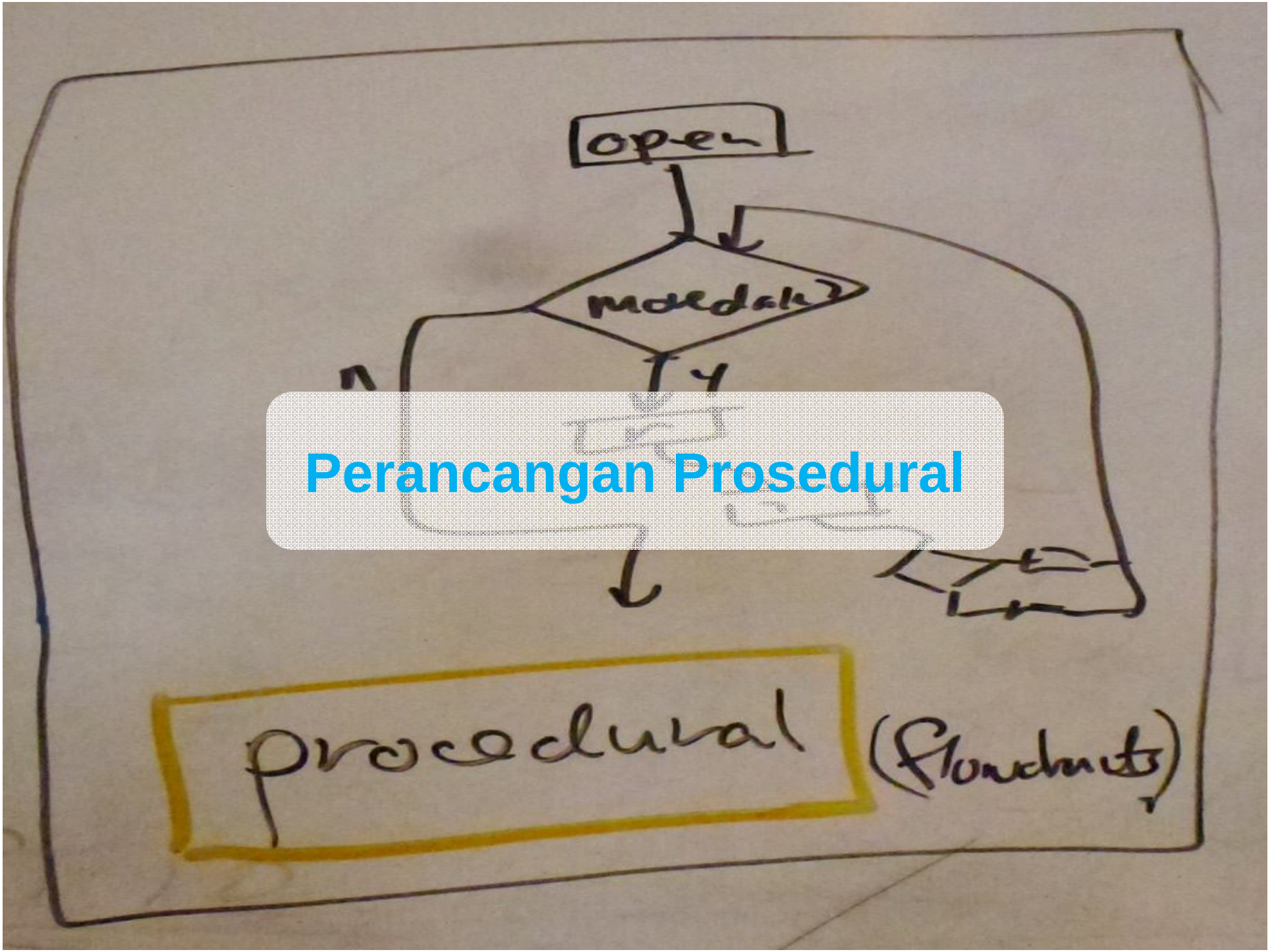


P06



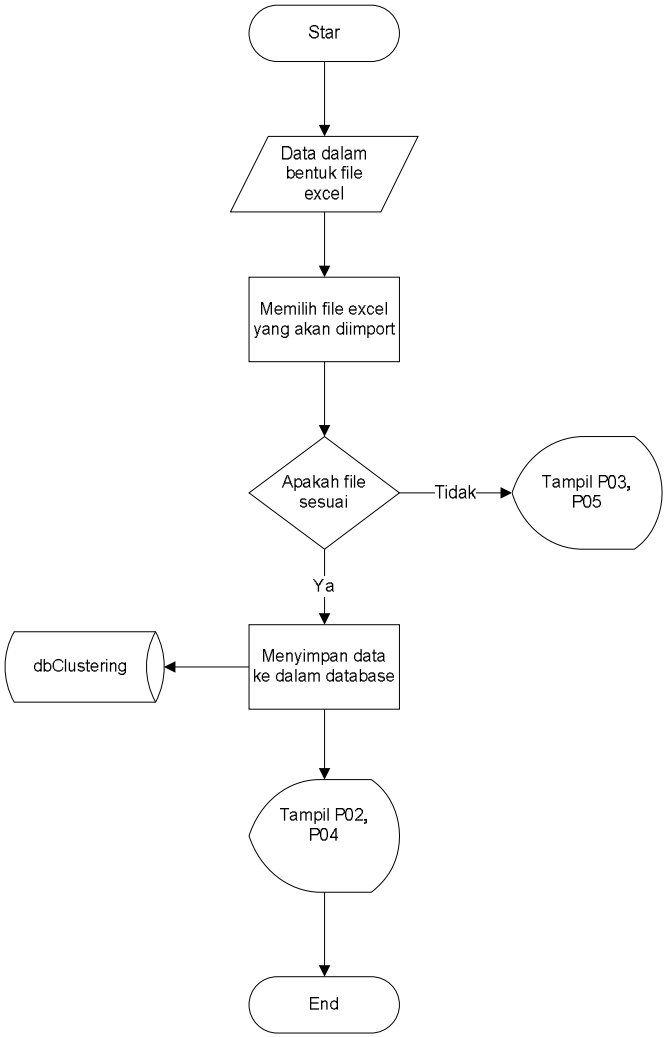
Jaringan Semantik



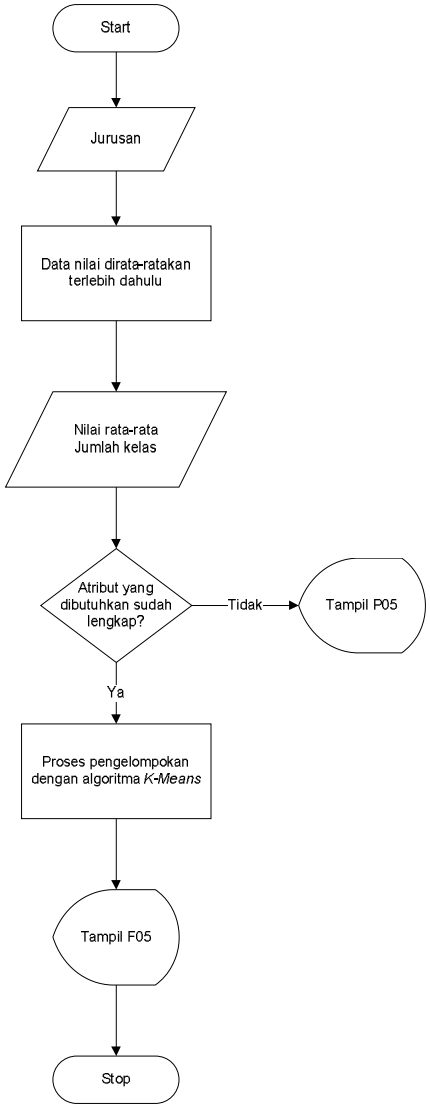


Perancangan Prosedural

Prosedur Import Data



Prosedur Pengelompokan



Algoritma K-means

Equivalence Partitioning Testing

Wawancara

- Membantu sekolah
- Hasil pengelompokan

Kasus sistem pembagian kelas

Terima Kasih