

**IMPLEMENTASI *SIMPLE K-MEDOIDS*
PARTITIONING PADA PENGELOMPOKAN SMA
BERDASARKAN KOMPONEN KESIAPAN
TEKNOLOGI INFORMASI**

(Studi Kasus : Sekolah Menengah Atas (SMA) di Provinsi Yogyakarta)

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Program
Studi Statistika



Disusun Oleh:

Syafira Irsalina

18611105

**PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA
2022**

HALAMAN PERSETUJUAN PEMBIMBING

TUGAS AKHIR

Judul : Implementasi *Simple K-Medoids Partitioning*
Pada Pengelompokan SMA Berdasarkan
Komponen Kesiapan Teknologi Informasi
(Studi Kasus : Sekolah Menengah Atas (SMA) di
Provinsi Daerah Istimewa Yogyakarta)

Nama Mahasiswa : Syafira Irsalina

NIM : 18611105

**TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK
DIUJIKAN**

Yogyakarta, 25 Maret 2022

Dosen Pembimbing

Mengetahui,
Ketua Program Studi



(Kariyam, S.Si., M.Si.)



(Dr. Edy Widodo, S.Si., M.Si.)

HALAMAN PENGESAHAN

TUGAS AKHIR

IMPLEMENTASI *SIMPLE K-MEDOIDS PARTITIONING* PADA PENGELOMPOKAN SMA BERDASARKAN KOMPONEN KESIAPAN TEKNOLOGI INFORMASI

(Studi Kasus : Sekolah Menengah Atas (SMA) di Provinsi Yogyakarta)

Nama Mahasiswa : Syafira Irsalina

NIM : 18611105

TUGAS AKHIR INI TELAH DIUJIKAN
PADA TANGGAL : 29 MARET 2022

Nama Penguji

Tanda Tangan

1. Dr. Atina Ahdika, S.Si., M.Si.
2. Muhammad Muhajir, S.Si., M.Sc.
3. Kariyam, S.Si., M.Si.

.....
.....
.....

Mengetahui,

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

.....
(Prof. Riyanto, S.Pd., M.Si., Ph.D.)

KATA PENGANTAR



Assalamu'alaikum Warahmatullaahi Wabarakatuh

Alhamdulillahirabbil'alamiin, puji dan syukur penulis panjatkan kehadiran Allah SWT yang telah memberikan rahmat dan hidayah-Nya yang besar dalam menyelesaikan amanah yang menjadi tanggung jawab penulis. Shalawat serta salam senantiasa tercurahkan kepada junjungan kita Nabi Muhammad SAW, berkat beliau kita bisa menjadi manusia yang berada di jalan yang diridhoi oleh Allah SWT.

Atas karunia-Nya penulis telah menyelesaikan Tugas Akhir yang merupakan langkah terakhir untuk memperoleh gelar Sarjana Statistika di Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia. Tugas akhir ini berjudul **“Implementasi *Simple K-Medoids Partitioning* Pada Pengelompokan SMA Berdasarkan Komponen Kesiapan Teknologi Informasi (Studi Kasus : Sekolah Menengah Atas (SMA) di Provinsi Yogyakarta)”**.

Penulis menyadari bahwa selama penyusunan Tugas Akhir tidak lepas dari bantuan, bimbingan, arahan, masukan dan dukungan baik secara langsung maupun tidak langsung, sehingga dapat berjalan dengan lancar dan tepat waktu. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Riyanto, S.Pd., M.Si., Ph.D., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia.
2. Bapak Dr. Edy Widodo, S.Si., M.Si., selaku Ketua Program Studi Statistika, Universitas Islam Indonesia beserta seluruh jajarannya.
3. Ibu Kariyam S.Si., M.Si., selaku Dosen Pembimbing yang selalu bersedia meluangkan waktu untuk membimbing, memberikan arahan, masukan selama penyusunan Tugas Akhir ini.
4. Bapak Ibu Dosen dan Staff Prodi Statistika Universitas Islam Indonesia yang selalu berbagi ilmu baik dalam bidang akademik maupun non akademik.

5. Ayah dan Ibu yang selalu memberikan do'a, dukungan, kasih sayang dan beliau lah yang menjadi semangat terbesar penulis. Serta Adek dan Keluarga Besar penulis yang selalu mendo'akan dan memberi semangat.
6. Bapak Dr. Weksi Budiaji, S.Si., M.Sc., yang bersedia membantu dan membagi ilmunya terkait metode yang digunakan oleh penulis.
7. Rekan – rekan CAPTAIN 18 (Statistika Angkatan 2018), terimakasih untuk kebersamaan, bantuan dan dukungan moril, pengalaman dan segalanya. Kalian hebat!.
8. Teman – teman saya, Rizky Normelia, Ellen Prihana Putri, Titian Dewi Fortuna, yang selalu berbagi ilmu dan berbagi cerita. Semangat!
9. Teman seperbimbingan Tugas Akhir, Sheilta Alphenia yang sama-sama berjuang untuk menyelesaikan Tugas Akhir. Semangat!.
10. Seluruh pihak yang telah memberikan bantuan baik materiil dan non materiil yang penulis tidak dapat sebutkan.

Penulis menyadari bahwa Tugas Akhir ini jauh dari kata sempurna. Oleh karenanya, penulis mengharapkan kritik serta saran yang dapat membangun untuk kesempurnaan Tugas Akhir ini. Penulis berharap Tugas Akhir ini dapat memberi manfaat bagi seluruh pihak. *Aamiin ya Rabbal'Alamin.*

Yogyakarta, 28 Januari 2022



Syafira Irsalina

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN PEMBIMBING TUGAS AKHIR.....	ii
HALAMAN PENGESAHAN TUGAS AKHIR	iii
KATA PENGANTAR.....	iv
DAFTAR ISI	vi
DAFTAR TABEL	viii
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN	xiii
PERNYATAAN	xiv
INTISARI.....	xv
ABSTRACT	xvi
BAB I PENDAHULUAN	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah	6
1.3. Batasan Masalah.....	6
1.4. Jenis Penelitian dan Metode Analisis.....	7
1.5. Tujuan Penelitian	7
1.6. Manfaat Penelitian	8
BAB II TINJAUAN PUSTAKA.....	9
2.1. Penelitian Terdahulu	13
BAB III LANDASAN TEORI.....	15
3.1. Asesmen Nasional.....	15
3.1.1 AKM Literasi – Numerasi	16
3.1.2 Survei Karakter.....	16
3.1.3 Survei Lingkungan Belajar.....	16
3.2. Sekolah Menengah Atas (SMA)	16
3.2.1 Kesiapan TIK untuk ANBK	17
3.3. Sekolah Kita	19
3.4. Data	20
3.4.1 Jenis Data.....	20
3.4.2 Skala Pengukuran	21
3.5. Analisis Deskriptif	21
3.6. Analisis Multivariat.....	22
3.6.1 <i>One-Way</i> MANOVA	22
3.7. Analisis Klaster	23
3.7.1 Ukuran Kedekatan (<i>Similarity</i>) Antar Objek.....	23
3.7.2 Metode Pengelompokan Non Hirarki.....	27
3.7.3 Validasi Klaster	30
3.8. Pengujian Rata – Rata Dua Populasi Independen	34
3.8.1 Teorema Limit Pusat	34
3.8.2 Uji Rata – Rata Dua Populasi Independen	34
BAB IV METODOLOGI PENELITIAN	36
4.1. Populasi dan Sampel Penelitian	36
4.2. Tempat dan Waktu Penelitian	36
4.3. Variabel Penelitian	36

4.4.	Alat dan Cara Organisir Data	38
4.5.	Tahapan Penelitian	38
4.5.1	Tahapan Analisis <i>Simple K-Medoids Partitioning</i>	41
4.5.2	Tahapan Menghitung Nilai <i>Silhouette Index</i>	43
4.5.3	Tahapan Menghitung Nilai <i>Medoid Shadow Value</i> (MSV) ..	44
	BAB V HASIL DAN PEMBAHASAN	46
5.1.	Gambaran Data Penelitian	46
5.2.	Statistika Deskriptif	46
5.3.	Pengukuran Kedekatan Antar Objek	56
5.4.	Pengelompokan SMA di Provinsi Yogyakarta	60
5.4.1	Pengelompokan dengan <i>Simple K-Medoids Partitioning</i>	60
5.4.2	Validasi Pengelompokan <i>Simple K-Medoids Partitioning</i>	64
5.4.3	Profilisasi Pengelompokan <i>Simple K-Medoids Partitioning</i> ..	70
5.5.	Pengujian Rata – Rata Dua Populasi Independen	83
5.6.	Validasi Metode	84
	BAB VI PENUTUP	86
6.1.	Kesimpulan	86
6.2.	Saran	87
	DAFTAR PUSTAKA	88
	RINGKASAN TA	92
	LAMPIRAN	104

DAFTAR TABEL

Tabel 2.1 Tabel Penelitian Sebelumnya	13
Tabel 3.1 Nilai Kaufman Karakteristik <i>Silhouette Index</i>	33
Tabel 4.1 Variabel Penelitian	36
Tabel 4.2 Struktur Data Penelitian	37
Tabel 4.3 Definisi Operasional Variabel	37
Tabel 5.1 Gambaran Data Penelitian	46
Tabel 5.2 Statistika Deskriptif	47
Tabel 5.3 Matriks Kedekatan Objek Data Numerik	57
Tabel 5.4 Matriks Kedekatan Objek Data Biner dan Kategorik	58
Tabel 5.5 Matriks Kedekatan Objek Data Campuran	59
Tabel 5.6 Matriks Jarak Data Campuran	60
Tabel 5.7 Nilai ai	61
Tabel 5.8 Medoid Akhir Hasil Pengelompokan	62
Tabel 5.9 Jarak Objek dengan Medoid Akhir	62
Tabel 5.10 Ringkasan Hasil Pengelompokan dengan $seeding = 20$	63
Tabel 5.11 Nilai SI: Percobaan dengan 2 Klaster	64
Tabel 5.12 Nilai SI: Percobaan dengan 3 Klaster	65
Tabel 5.13 Nilai SI: Percobaan dengan 4 Klaster	66
Tabel 5.14 Nilai SI: Percobaan dengan 5 Klaster	66
Tabel 5.15 Nilai <i>Silhouette Index</i>	67
Tabel 5.16 Nilai MSV: Percobaan dengan 2 Klaster	68
Tabel 5.17 Nilai MSV: Percobaan dengan 3 Klaster	68
Tabel 5.18 Nilai MSV: Percobaan dengan 4 Klaster	69
Tabel 5.19 Nilai MSV: Percobaan dengan 5 Klaster	70
Tabel 5.20 Nilai <i>Medoid Shadow Value</i> (MSV)	70
Tabel 5.21 Hasil Pengelompokan dengan SKM	71
Tabel 5.22 Profilisasi Anggota Klaster Berdasarkan Variabel Numerik	81
Tabel 5.23 Karakteristik Klaster Hasil Pengelompokan	82
Tabel 5.24 Statistik Uji: Pengujian Rata – Rata Dua Populasi Independen Variabel Numerik Penelitian	83

Tabel 5.25 Statistik Uji: Pengujian Rata – Rata Dua Populasi Independen Variabel Numerik Non Penelitian	84
Tabel 5.26 Validasi Metode: KM, SFKM dan SKM	85
Tabel 27 Nilai Nilai Kaufman Karakteristik <i>Silhouette Index</i>	97
Tabel 28 Variabel dan Jenis Data Penelitian	98
Tabel 29 Gambaran Data Penelitian	99
Tabel 30 Matriks Jarak Data Campuran	99
Tabel 31 Ringkasan Hasil Pengelompokan dengan <i>seeding</i> = 20	99
Tabel 32 Hasil Pengelompokan SMA di Provinsi Yogyakarta.....	100
Tabel 33 Nilai Validasi SI dan MSV	100
Tabel 34 Profilisasi Klaster Berdasarkan Variabel Numerik.....	101
Tabel 35 Profilisasi Klaster Berdasarkan Variabel Kategorik	101
Tabel 36 Statistik Uji Rata – Rata Dua Populasi Independen Variabel Numerik.....	102
Tabel 37 Statistik Uji Rata – Rata Dua Populasi Independen Variabel Numerik Non Penelitian.....	102
Tabel 1 Data Contoh Studi Kasus	166
Tabel 2 Contoh: Matriks Kedekatan Antar Objek Data Numerik.....	167
Tabel 3 Contoh: Matriks Kedekatan Antar Objek Data Biner & Kategorik	167
Tabel 4 Contoh: Kedekatan Antar Objek Data Campuran.....	168
Tabel 5 Contoh: Matriks Jarak Antar Objek	169
Tabel 6 Contoh: Nilai <i>ai</i> Matriks Jarak	169
Tabel 7 Contoh: Medoid Awal	170
Tabel 8 Contoh: Iterasi 1	170
Tabel 9 Contoh: Matriks Jarak <i>Reduce</i> Klaster 1 Iterasi 1	171
Tabel 10 Contoh: Matriks Jarak <i>Reduce</i> Klaster 2 Iterasi 1	171
Tabel 11 Contoh: Medoid Baru Iterasi 2.....	171
Tabel 12 Contoh: Iterasi 2	172
Tabel 13 Contoh: Hasil Pengelompokan <i>Seeding</i> 1	172
Tabel 14 Contoh: Hasil Perhitungan dengan <i>Seeding</i> = 20.....	173
Tabel 15 Perhitungan SI : Nilai <i>ai</i> Klaster 1 dan Klaster 2	174
Tabel 16 Perhitungan SI : Hasil Jarak Objek Klaster 1 dengan Klaster 2.....	175
Tabel 17 Perhitungan SI : Hasil Jarak Objek Klaster 2 dengan Klaster 1.....	175

Tabel 18 Perhitungan SI : Nilai <i>bi</i> Klaster 1 dan Klaster 2	175
Tabel 19 Perhitungan SI : Nilai <i>Silhouette Index</i> Setiap Objek	176
Tabel 20 Perhitungan SI: Nilai <i>Silhouette Index</i> Setiap Klaster	176
Tabel 21 Nilai <i>Medoid Shadow Value</i> (MSV)	178

DAFTAR GAMBAR

Gambar 1.1 Tingkat Penyelesaian Pendidikan Jenjang SM/Sederajat	2
Gambar 1.2 Data Sekolah, Guru dan Siswa SMA Provinsi Yogyakarta	3
Gambar 1.3 Grafik Jumlah Sekolah Menurut Status Sekolah	4
Gambar 3.1 Elemen Kebijakan dan Dampak Asesmen Nasional	15
Gambar 3.2 Infrastruktur Moda Pelaksanaan TIK	18
Gambar 4.1 Diagram Alir Penelitian	39
Gambar 4.2 Diagram Alir Algoritma <i>Simple K-Medoids Partitioning</i>	42
Gambar 4.3 Diagram Alir Perhitungan Nilai <i>Silhouette Index</i>	43
Gambar 4.4 Diagram Alir Perhitungan Nilai <i>Medoid Shadow Value</i>	44
Gambar 5.1 Diagram Batang Jumlah Infrastruktur Sekolah SMA	47
Gambar 5.2 Diagram Batang Jumlah Guru Sekolah SMA	48
Gambar 5.3 Diagram Batang Jumlah Siswa Sekolah SMA	49
Gambar 5.4 Persentase Sekolah SMA Menurut Status Sekolah	49
Gambar 5.5 Diagram Batang Guru Sekolah SMA	50
Gambar 5.6 Diagram Batang Siswa Sekolah SMA	50
Gambar 5.7 Diagram Batang Laboratorium Komputer Sekolah SMA	51
Gambar 5.8 Diagram Batang Komputer Utama Sekolah SMA	51
Gambar 5.9 Diagram Batang Komputer Spesifikasi (Milik) Sekolah SMA	52
Gambar 5.10 Diagram Batang Rata-Rata <i>Bandwidth</i> Sekolah SMA	52
Gambar 5.11 Diagram Batang <i>Swicth Hub</i> Sekolah SMA	53
Gambar 5.12 Diagram Batang WIFI Sekolah SMA	54
Gambar 5.13 Diagram Batang Akreditasi Sekolah SMA	54
Gambar 5.14 Diagram Batang Status Pelaksanaan TIK Sekolah SMA	55
Gambar 5.15 Diagram Batang Sumber Listrik Sekolah SMA	56
Gambar 5.16 Diagram Batang Moda Pelaksanaan TIK Sekolah SMA	56
Gambar 5.17 Grafik SI: Percobaan dengan 2 Klaster	65
Gambar 5.18 Grafik SI: Percobaan dengan 3 Klaster	65
Gambar 5.19 Grafik SI: Percobaan dengan 4 Klaster	66
Gambar 5.20 Grafik SI: Percobaan dengan 5 Klaster	67
Gambar 5.21 Grafik MSV: Percobaan dengan 2 Klaster	68

Gambar 5.22 Grafik MSV: Percobaan dengan 3 Klaster	69
Gambar 5.23 Grafik MSV: Percobaan dengan 4 Klaster	69
Gambar 5.24 Grafik MSV: Percobaan dengan 5 Klaster	70
Gambar 5.25 Profil Klaster 1: Akreditasi Berdasarkan Wilayah	71
Gambar 5.26 Profil Klaster 1: Persentase Status Sekolah	72
Gambar 5.27 Profil Klaster 1: Persentase Moda Pelaksanaan TIK	72
Gambar 5.28 Profil Klaster 1: Persentase Sumber Listrik	73
Gambar 5.29 Profil Klaster 1: Persentase Status Pelaksanaan TIK	73
Gambar 5.30 Profil Klaster 1: Guru	73
Gambar 5.31 Profil Klaster 1: Siswa	73
Gambar 5.32 Profil Klaster 1: Laboratorium Komputer	74
Gambar 5.33 Profil Klaster 1: Komputer Utama	74
Gambar 5.34 Profil Klaster 1: Komputer Sesuai Spesifikasi (Milik)	75
Gambar 5.35 Profil Klaster 1: Rata-Rata <i>Bandwidth</i>	75
Gambar 5.36 Profil Klaster 1: <i>Switch Hub</i>	76
Gambar 5.37 Profil Klaster 1: WIFI	76
Gambar 5.38 Profil Klaster 2: Akreditasi Berdasarkan Wilayah	77
Gambar 5.39 Profil Klaster 2: Persentase Status Sekolah	77
Gambar 5.40 Profil Klaster 2: Persentase Moda Pelaksanaan TIK	77
Gambar 5.41 Profil Klaster 2: Persentase Sumber Listrik	78
Gambar 5.42 Profil Klaster 2: Persentase Status Pelaksanaan TIK	78
Gambar 5.43 Profil Klaster 2: Guru	79
Gambar 5.44 Profil Klaster 2: Siswa	79
Gambar 5.45 Profil Klaster 2: Laboratorium Komputer	79
Gambar 5.46 Profil Klaster 2: Komputer Utama	79
Gambar 5.47 Profil Klaster 2: Komputer Sesuai Spesifikasi (Milik)	80
Gambar 5.48 Profil Klaster 2: Rata-Rata <i>Bandwidth</i>	80
Gambar 5.49 Profil Klaster 2: <i>Switch Hub</i>	81
Gambar 5.50 Profil Klaster 2: WIFI	81
Gambar 1 Grafik <i>Silhouette Index</i>	177
Gambar 2 Grafik <i>Medoid Shadow Value</i>	178

DAFTAR LAMPIRAN

Lampiran 1	Data Kode Sekolah Menengah Atas (SMA)	104
Lampiran 2	Data Penelitian.....	109
Lampiran 3	Data Penelitian untuk <i>R Studio</i>	117
Lampiran 4	Ukuran Kedekatan Antar Objek Data Numerik	125
Lampiran 5	Ukuran Kedekatan Antar Objek Data Kategorik.....	126
Lampiran 6	Ukuran Kedekatan Antar Objek Data Campuran	127
Lampiran 7	Ukuran Jarak Antar Objek Data Campuran.....	128
Lampiran 8	Ringkasan Hasil <i>Seeding</i> = 20 ($k = 3, 4, 5$)	129
Lampiran 9	Validasi Nilai <i>Silhouette Index</i> ($k = 2, 3, 4, 5$)	132
Lampiran 10	Validasi Nilai <i>Medoid Shadow Value</i> (MSV) ($k = 2, 3, 4, 5$)..	136
Lampiran 11	Hasil Pengelompokan Algoritma SKM ($k = 2, 3, 4, 5$)	140
Lampiran 12	<i>Syntax Simple K-Medoids</i> Perbaikan.....	144
Lampiran 13	<i>Syntax Nested Function</i> SFKM untuk Modifikasi SKM.....	146
Lampiran 14	<i>Syntax</i> SKM untuk Memunculkan Proses Iterasi	150
Lampiran 15	<i>Syntax</i> Analisis Data Menggunakan SKM	152
Lampiran 16	Pengujian Hipotesis: Rata – Rata Dua Populasi Independen ..	158
Lampiran 17	Contoh Perhitungan Manual Algoritma SKM.....	166

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya karya yang sebelumnya pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali yang diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 20 Maret 2022



(Syafira Irsalina)

INTISARI

IMPLEMENTASI *SIMPLE K-MEDOIDS PARTITIONING* PADA PENGELOMPOKAN SMA BERDASARKAN KOMPONEN KESIAPAN TEKNOLOGI INFORMASI

(Studi Kasus : Sekolah Menengah Atas (SMA) di Provinsi Yogyakarta)

Syafira Irsalina

Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Islam Indonesia

Evaluasi pendidikan menjadi hal yang tidak dapat dipisahkan dengan komponen pendidikan. Secara nasional, evaluasi dilakukan dengan sistem Ujian Nasional Berbasis (UNBK). Namun, pada tahun 2021, UNBK diganti menjadi Asesmen Nasional Berbasis Komputer (ANBK). Provinsi Yogyakarta memiliki tingkat penyelesaian sekolah jenjang SM/ sederajat tertinggi dan tingkat kesenjangan dalam kecakapan digital paling rendah. Persiapan Teknologi Informasi dan Komunikasi (TIK) menjadi hal penting pada pelaksanaan ANBK. Tujuan penelitian yakni mengelompokkan sekolah SMA di Provinsi Yogyakarta berdasarkan data komponen kesiapan TIK untuk ANBK menggunakan algoritma *Simple K-Medoids Partitioning* (SKM) dengan jarak *Gower*, validasi kluster menggunakan *Silhouette Index* dan *Medoid Shadow Value* serta karakteristik kelompok yang terbentuk dari hasil pengelompokan. Data berasal dari website Sekolah Kita milik Kemendikbud yang merupakan data campuran yakni terdiri dari 8 variabel numerik dan 5 variabel kategorik. Data dikelompokkan menjadi dua kelompok. Hasil pengelompokan dengan *seeding* = 20 dan *iterate* = 10 menunjukkan bahwa kluster 1 terdiri dari 83 sekolah yang merupakan kelompok dengan jumlah SDM dan infrastruktur TIK yang tinggi dan kluster 2 terdiri dari 82 sekolah yang merupakan kelompok dengan jumlah SDM dan infrastruktur TIK yang rendah serta validasi kluster diperoleh jumlah k kluster yang optimal adalah ketika $k = 2$.

Kata Kunci : Komponen Kesiapan TIK untuk ANBK, SMA, *Simple K-Medoids Partitioning*

ABSTRACT

IMPLEMENTATION OF SIMPLE K-MEDOIDS PARTITIONING IN SMA GROUPING BASED ON INFORMATION TECHNOLOGY READINESS COMPONENTS

(Case Study : Senior High School (SMA) in Yogyakarta Province)

Syafira Irsalina

*Department of Statistics, Faculty of Mathematics and Natural Sciences
Islamic University of Indonesia*

Educational evaluation is something that cannot be separated from the education component. Nationally, the evaluation is carried out using the National Based Examination (UNBK) system. However, in 2021, UNBK was changed to the Computer-Based National Assessment (ANBK). Yogyakarta province has the highest secondary school completion rate and the lowest disparity in digital skills. Preparation of Information and Communication Technology (ICT) is an important thing in the implementation of ANBK. The research objectives are to group high school schools in Yogyakarta Province based on the component data of ICT readiness for ANBK using the Simple K-Medoids Partitioning (SKM) algorithm with Gower distance, cluster validation using Silhouette Index and Medoid Shadow Value and the characteristics of the groups formed from the grouping results. The data comes from the Sekolah Kita website belonging to the Ministry of Education and Culture which is mixed data consisting of 8 numerical variables and 5 categorical variables. The data are grouped into two groups. The results of grouping with seeding = 20 and iterate = 10 indicate that cluster 1 consists of 83 schools which are a group with a high number of human resources and ICT infrastructure and cluster 2 consists of 82 schools which are a group with a low number of human resources and ICT infrastructure. In cluster validation, the optimal number of k clusters is when $k = 2$.

Keywords: *ICT Readiness Components for ANBK, Senior High School, Simple K-Medoids Partitioning*

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

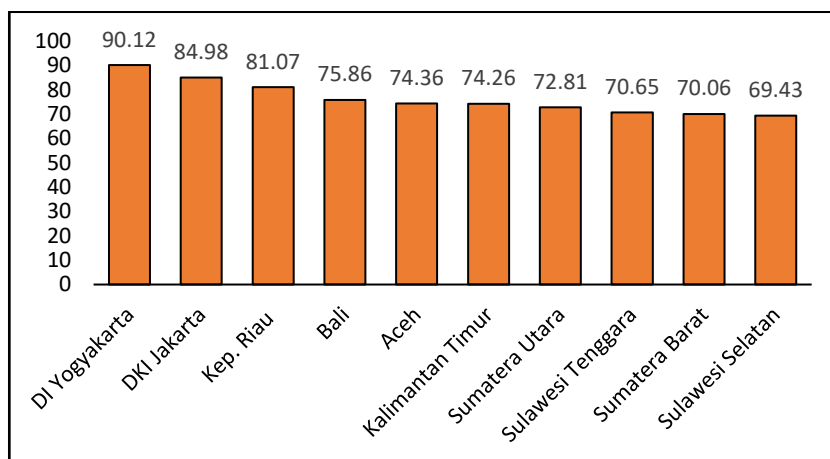
Evaluasi pendidikan menjadi hal yang tidak dapat dipisahkan dengan komponen pendidikan. Namun, tidak seluruh bentuk evaluasi dapat digunakan dalam mengukur pencapaian tujuan pendidikan yang sudah ditentukan. Apabila alat ukur yang digunakan untuk mengevaluasi sesuai dan dapat mengukur setiap tujuan, maka keberhasilan pendidikan dapat dicapai (L, 2019).

Hasil evaluasi menjadi indikator keberhasilan terhadap proses pembelajaran yang dilakukan pengajar dan peserta didiknya. Saat ini proses evaluasi yang diselenggarakan secara nasional untuk mengetahui ketercapaiannya pendidikan nasional dilakukan melalui Ujian Nasional (UN). Keberadaan dan penerapan UN menuai berbagai pro kontra yang semakin berlarut. Terkhusus sekarang, sistem UN yang pemerintah terapkan adalah berbasis komputer atau Ujian Nasional Berbasis Komputer (UNBK). Tahun 2014 merupakan penyelenggaraan UNBK pertama kalinya dan terbatas pada sebagian sekolah di Indonesia (Ayang Sari, 2020).

Pada akhirnya Menteri Pendidikan dan Kebudayaan Nadiem Makarim secara resmi mengatakan bahwa tahun 2020 adalah tahun terakhir dalam pelaksanaan UNBK (11/12/2020). Tahun 2021, UNBK akan diganti menjadi Asesmen Nasional Berbasis Komputer (ANBK). Menurutny, ANBK tidak dilakukan hanya berdasarkan pada penguasaan materi kurikulum seperti selama ini yang diterapkan pada UNBK, melainkan berupa pemetaan pada dua kompetensi minimum peserta didik yakni literasi dan numerasi (Ayang Sari, 2020). ANBK diselenggarakan dengan memanfaatkan perkembangan teknologi.

Asesmen Nasional terdiri dari tiga instrumen yaitu Asesmen Kompetensi Minimum (AKM) Literasi-Numerasi, Survei Karakter, dan Survei Lingkungan Belajar (Kemendikbud, 2021). Secara umum, ANBK adalah bentuk penerapan ujian yang berbasiskan komputer dari Asesmen Nasional. Namun, ANBK hanya ditujukan untuk sampel siswa dari setiap jenjang pendidikan, bukan untuk keseluruhan siswa. Pada jenjang SD/ sederajat, ANBK akan dilaksanakan kepada siswa kelas 5. Sementara, pada jenjang SMP/ sederajat pada siswa kelas 8 dan

jenjang SMA/ sederajat untuk siswa kelas 11. Pelaksanaan per jenjang secara bertahap memungkinkan sekolah untuk berbagi sumber daya. Sekolah/ madrasah yang belum mempunyai infrastruktur memadai dapat mengikuti ANBK di satuan pendidikan lain. Kemendikbud Ristek melaksanakan AKM dengan moda asesmen berbasis komputer baik secara online maupun semi online (Maharani, 2019).



Gambar 1.1 Tingkat Penyelesaian Pendidikan Jenjang SM/ Sederajat

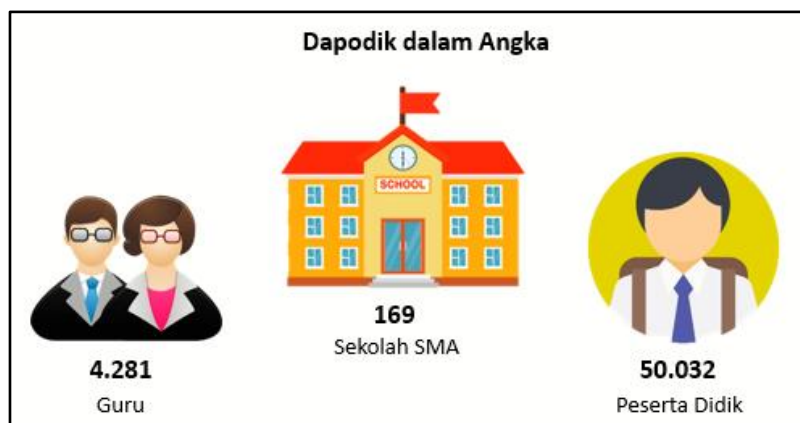
Provinsi Yogyakarta menjadi provinsi pertama dengan tingkat penyelesaian pendidikan pada jenjang SM/ Sederajat yang tertinggi yaitu dengan persentase sebesar 90.12% (Rakyat, 2021). Dengan persentase yang tinggi tersebut, maka banyak indikator yang menunjang siswa untuk dapat menyelesaikan jenjang pendidikan sesuai dengan rentang usianya. Salah satu indikatornya adalah berhasilnya kegiatan evaluasi belajar yang diselenggarakan pada sekolah SMA di Provinsi Yogyakarta baik dari infrastruktur maupun sumber daya.

Selain itu, Provinsi Yogyakarta memiliki tingkat kesenjangan dalam kecakapan digital paling rendah dibandingkan provinsi lain karena di wilayah Yogyakarta mempunyai infrastruktur *wireless* terbesar setelah provinsi DKI Jakarta. Sehingga kemampuan mengakses TIK juga besar (Kemendikbudristek, 2021). Kesiapan sekolah dalam bidang penggunaan teknologi untuk melaksanakan ANBK menjadi salah satu tolak ukur keberhasilan dari penyelenggaraan ANBK.

Pada klasterisasi sekolah di Provinsi Yogyakarta tahun 2021 oleh Kemendikbud pada indikator status kesiapan TIK suatu sekolah. Indikator pertama yaitu apakah sekolah melaksanakan UNBK pada tahun sebelumnya secara mandiri, indikator kedua yaitu apakah sekolah memiliki laboratorium komputer, indikator ketiga yaitu apakah sekolah pernah mendapat bantuan TIK, indikator keempat

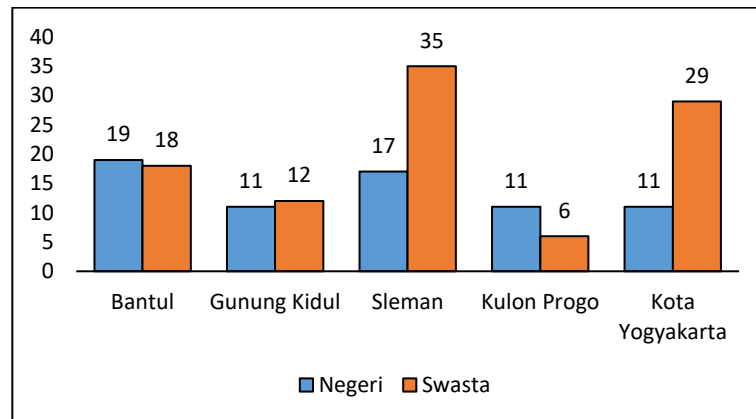
apakah sekolah memiliki minimal 15 komputer, indikator kelima yaitu apakah sekolah memiliki aliran listrik dan indikator keenam yaitu apakah sekolah memiliki jaringan internet. Klasterisasi Kemendikbud mengelompokkan sekolah menjadi 4 kelompok. Kelompok 1 merupakan sekolah dengan status Siap A sebanyak 134 sekolah. Kelompok 2 merupakan sekolah dengan status Siap B sebanyak 9 sekolah. Kelompok 3 merupakan sekolah dengan status Potensial 1 sebanyak 19 sekolah. Kelompok 4 merupakan sekolah dengan status Potensial 2 sebanyak 3 sekolah.

Pemerintah melalui Kemendikbud telah memberikan kemudahan bagi masyarakat maupun pihak terkait untuk mengakses data tersebut melalui website Sekolah Kita Kemendikbud. Website tersebut menjadi pusat data untuk melaporkan data sekolah meliputi profil sekolah, profil sanitasi sekolah dan komponen kesiapan TIK sekolah. Data tersebut diinputkan langsung oleh masing – masing sekolah.



Gambar 1.2 Data Sekolah, Guru dan Siswa SMA Provinsi Yogyakarta

Berdasarkan pada **Gambar 1.2** diperoleh informasi mengenai jumlah sekolah SMA sederajat, guru dan siswa pada tahun 2021 di Provinsi Yogyakarta. Total ada sebanyak 169 sekolah SMA dengan 4.281 guru yang mengajar dan peserta didik sebanyak 50.032. Apabila dipetakan kembali, terdapat dua status sekolah yaitu negeri atau swasta berdasarkan kabupaten/kota disajikan pada **Gambar 1.2**.



Gambar 1.3 Grafik Jumlah Sekolah Menurut Status Sekolah

Pada **Gambar 1.3** memberikan informasi bahwa jumlah SMA/ sederajat di Provinsi Yogyakarta didominasi oleh sekolah dengan status swasta sebanyak 100 sekolah. Kabupaten Sleman menjadi kabupaten dengan jumlah sekolah swasta terbanyak sedangkan Kabupaten Kulon Progo menjadi kabupaten dengan jumlah sekolah swasta terendah. Terdapat tiga kabupaten/kota dengan jumlah sekolah negeri yang sama yakni sebanyak 11 sekolah, wilayah tersebut adalah Kabupaten Gunung Kidul, Kabupaten Kulon Progo dan Kota Yogyakarta.

Analisis kluster adalah alat eksplorasi penting dalam suatu penyelidikan struktur data. Setiap objek yang berada dalam suatu kelompok bersifat homogen satu dengan lainnya, dan objek yang berada dalam kelompok lain bersifat heterogen satu sama lain (Kaufman, 1990). Terdapat beberapa metode yang dapat digunakan untuk pengelompokan yaitu metode hirarki dan non-hirarki. Pemilihan metode non-hirarki dikarenakan metode ini lebih efektif dalam segi waktu komputasi untuk data yang besar. Selain itu, penentuan jumlah k kluster ditentukan di awal oleh peneliti serta metode ini memiliki sedikit kelemahan pada data *outlier*, ukuran jarak yang digunakan serta variabel yang tidak relevan. Sedangkan metode hirarki kurang efisien apabila digunakan untuk data yang besar serta memiliki kelemahan pada beberapa permasalahan seperti sering terdapat kesalahan pada data *outlier*, perbedaan ukuran jarak yang digunakan dan penggunaan variabel yang kurang relevan. Beberapa metode non-hirarki yang sering digunakan di antaranya meliputi metode *k-means* dan *k-medoids*.

Salah satu algoritma *k-medoids* yang dikembangkan untuk data dengan variabel campuran yaitu ordinal, biner dan kategorik adalah *Simple K-Medoids*

Partitioning (SKM). Algoritma ini menggunakan medoid sebagai pusat klasternya. Algoritma ini dikembangkan untuk mengatasi permasalahan *local optima* dan klaster kosong pada algoritma *K-Medoids* (KM) dan *Simple and Fast K-Medoids* (SFKM). Permasalahan *local optima* yang dimaksudkan adalah *local optima* minimal yakni dengan meminimalkan jarak antar objek dan pusat klaster. Sehingga banyak *local optima* yang dapat muncul bergantung pada pemilihan nilai awal. Adanya inisialisasi pusat klaster secara acak dapat memberikan hasil pembentukan klaster yang berubah – ubah (tidak konsisten) yang sering terjadi pada algoritma KM dan SFKM mengakibatkan sering terjadinya permasalahan *local optima*. Salah satu cara mengatasi masalah tersebut adalah dengan menerapkan algoritma SFKM untuk pengelompokan data bertipe campuran (Weksi Budiaji, 2019). Selain masalah *local optima*, permasalahan lainnya adalah dapat mengakibatkan terjadinya klaster kosong. Masalah klaster kosong yang dimaksudkan adalah tidak adanya anggota di dalam salah satu klaster yang ingin dibentuk. Kumpulan medoid dapat menjadi objek non-unik, misalnya ketika terdapat dua medoid yang memiliki nilai yang sama di semua variabelnya. Ketika himpunan medoid merupakan objek yang non-unik, maka klaster kosong dapat terjadi karena medoid non-unik dan objek terdekat akan memiliki label keanggotaan yang sama. Sehingga untuk menghindari klaster kosong, medoid non-unik dibatasi untuk dapat mempertahankan labelnya (Weksi Budiaji, 2019).

Setelah dilakukan pemaparan mengenai latar belakang sebelumnya, peneliti ingin melakukan penelitian mengenai analisis pengelompokan pada data Sekolah Kita milik Kemendikbud tentang sekolah SMA di Provinsi Yogyakarta. Hal yang dilakukan berupa melakukan pengelompokan variabel campuran berdasarkan kesamaan karakteristik yang dimiliki oleh masing – masing objek. Variabel numerik yang digunakan adalah jumlah guru, siswa, laboratorium komputer, komputer utama, komputer milik sesuai spesifikasi, rata – rata *bandwidth*, *switch hub* dan WIFI. Variabel kategorik yang digunakan adalah status sekolah, moda pelaksanaan TIK, sumber listrik, akreditasi sekolah dan status pelaksanaan TIK. Setiap objek yang ada akan dihitung ukuran kedekatannya dengan jarak *Gower* serta validasi hasil klaster akan menggunakan *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV).

Klasterisasi yang dilakukan oleh Kemendikbud tentang status kesiapan TIK pada seluruh jenjang sekolah menggunakan 6 indikator, akan tetapi proses klasterisasi masih memiliki kekurangan pada kejelasan metode yang digunakan. Penelitian ini berbeda dengan klasterisasi yang dilakukan Kemendikbud. Perbedaan tersebut ada pada pemilihan variabel serta jenjang pendidikan pada suatu wilayah tertentu, karena untuk efisiensi waktu dan adanya keterbatasan data yang ada pada sumber data yaitu sekolah.data.kemendikbud.go.id. Selanjutnya, perbedaan yang lain yakni penelitian ini akan berfokus pada jenjang SMA saja bukan SMA/ sederajat. Alasannya karena setelah melihat data pada SMA/ sederajat masih terdapat beberapa sekolah yang belum mengisi kelengkapan komponen data sehingga penelitian hanya berfokus pada jenjang SMA saja.

1.2. Rumusan Masalah

Berdasarkan uraian latar belakang yang telah dipaparkan, maka rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana gambaran umum data sekolah SMA di Provinsi Yogyakarta berdasarkan komponen kesiapan TIK untuk ANBK 2021?
2. Bagaimana hasil pengelompokan sekolah SMA di Provinsi Yogyakarta berdasarkan komponen kesiapan TIK untuk ANBK 2021 menggunakan algoritma *Simple K-Medoids Partitioning* (SKM) serta validasi kluster menggunakan *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV)?
3. Bagaimana karakteristik yang terbentuk dari hasil pengelompokan sekolah SMA di Provinsi Yogyakarta berdasarkan komponen kesiapan TIK untuk ANBK 2021?

1.3. Batasan Masalah

Agar pembahasan dalam penelitian ini tidak meluas, maka terdapat batasan masalah dalam penulisan ini yaitu:

1. Data yang digunakan adalah data sekunder tentang komponen kesiapan TIK sekolah SMA dalam penyelenggaraan ANBK di Provinsi Yogyakarta yang didapatkan dari website sekolah.data.kemendikbud.go.id untuk semester data 2021/2022-1.

2. Metode analisis yang digunakan menggunakan kluster non hirarki yaitu algoritma *Simple K-Medoids Partitioning* (SKM) dengan *set.seed(1)*, *seeding* = 20 dan iterasi maksimum adalah 10.
3. Hasil pengelompokan dan kluster optimal akan divalidasi dengan menggunakan *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV).
4. Perangkat lunak yang digunakan untuk membantu dalam proses analisis adalah *software R Studio* dan *Microsoft Excel*. Sedangkan untuk visualisasi dibantu dengan *software Tableau* dan *Geoda*.

1.4. Jenis Penelitian dan Metode Analisis

Jenis penelitian yang digunakan dalam Tugas Akhir ini adalah penelitian aplikatif yang mengacu pada penelitian yang dilakukan oleh Weksi Budiaji dan Friedrich Leisch dengan topik yang berjudul “*Simple K-Medoids Partitioning Algorithm for Mixed Variable Data*”. Pada penelitian tersebut membahas tentang suatu algoritma pengelompokan untuk data dengan variabel numerik dan campuran yaitu ordinal, biner, kategorik dan numerik dengan memiliki parameter *s* (atau *seeding*) sebagai jumlah inisialisasi untuk proses pembaruan medoid awal dengan adanya *initial* medoid yang berbeda – beda. Metode tersebut adalah algoritma *Simple K-Medoids Partitioning* (SKM). Berdasarkan penjelasan tersebut, peneliti akan menerapkan data komponen kesiapan TIK Sekolah SMA untuk ANBK Tahun 2021 yang selanjutnya akan dikelompokkan menggunakan algoritma SKM.

1.5. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditetapkan, maka penelitian ini memiliki tujuan di antaranya:

1. Mengetahui gambaran umum data sekolah SMA di Provinsi Yogyakarta berdasarkan komponen kesiapan TIK untuk ANBK 2021.
2. Mengetahui hasil pengelompokan sekolah SMA di Provinsi Yogyakarta berdasarkan komponen kesiapan TIK untuk ANBK 2021 menggunakan algoritma *Simple K-Medoids Partitioning* (SKM) serta validasi kluster menggunakan *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV).
3. Mengetahui karakteristik yang terbentuk dari hasil pengelompokan sekolah SMA di Provinsi Yogyakarta berdasarkan komponen kesiapan TIK untuk ANBK 2021.

1.6. Manfaat Penelitian

Penelitian ini diharapkan memiliki manfaat bagi pemerintah khususnya Kementerian Pendidikan dan Kebudayaan (Kemendikbud) untuk memberikan bantuan dan perbaikan infrastruktur pada sekolah yang memiliki fasilitas penunjang pembelajaran dan evaluasi belajar yang masih rendah. Selain itu, menjadi bahan evaluasi bagi sekolah SMA di Provinsi Yogyakarta untuk senantiasa meningkatkan kualitas dan kuantitas baik dari segi sumber daya manusia (SDM), fasilitas serta sarana dan prasarana yang mampu menunjang proses kegiatan pembelajaran dan evaluasi belajar siswa.

BAB II

TINJAUAN PUSTAKA

Peneliti sangat membutuhkan dukungan dari hasil penelitian yang telah ada sebelumnya untuk mengetahui hubungan antara penelitian yang sedang dilakukan dengan penelitian terdahulu sehingga terhindar dari adanya duplikasi. Setelah melakukan telaah pada beberapa penelitian, terdapat beberapa penelitian yang mempunyai keterkaitan dengan penelitian yang sedang dilakukan oleh peneliti.

Pada tahun 1997, (Huang, 1997) dengan penelitiannya yang berjudul “*Clustering Large Data Sets With Mixed Numeric And Categorical Values*” menyajikan algoritma *k-prototype* yang didasarkan pada paradigma *k-means*. Pada algoritma ini, sekumpulan objek akan dikelompokkan terhadap *k-prototype*. Dalam penelitian ini mengembangkan suatu metode untuk memperbarui *k-prototype* secara dinamis untuk memaksimalkan kesamaan antar objek dalam kluster. Jika diterapkan dalam data numerik, algoritma ini mirip dengan *k-means*.

Menurut penelitian (Hae-Sang Park, 2009) yang membahas suatu algoritma baru untuk *k-medoids* untuk pengelompokan yang mirip dengan *k-means* dan menguji beberapa metode dalam pemilihan medoid awal. Algoritma ini diberi nama *Simple and Fast K-Medoids* (SFKM). Algoritma SFKM ini menghitung matrik jarak dan digunakan untuk mendapatkan medoid baru pada setiap langkah iterasi. Dalam mengevaluasi SFKM, peneliti memakai beberapa data set nyata dan buatan lalu akan dibandingkan dengan hasil algoritma lain. Penelitian ini memberikan hasil bahwa algoritma SFKM memiliki kinerja yang lebih baik daripada algoritma *k-means* dan membutuhkan waktu yang lebih sedikit dibandingkan dengan algoritma *Partitioning Around Medoid* (PAM).

Empat tahun kemudian, di tahun 2013 (P. IndiraPriya, 2013) mengadakan penelitian tentang “*A Survey on Different Clustering Algorithms in Data Mining Technique*”. Penelitian ini dilakukan untuk mengumpulkan lebih banyak konsep dan teknik untuk analisis kluster pada data yang besar. Pada penelitian ini memberikan hasil bahwa teknik pengelompokan *soft* dan metode pengelompokan hirarki menjadi metode yang terbaik untuk pengelompokan data yang besar.

Penelitian berjudul “*K-Medoids Clustering for Heterogeneous Data Sets*” oleh (Sandhya Harikumar, 2015) mengusulkan ukuran kesamaan baru. Ukuran

kesamaan tersebut dalam bentuk triplet guna menemukan jarak antara dua objek data dengan tipe atribut yang heterogen. Selain itu, juga diusulkan algoritma *k-medoids* baru dengan memanfaatkan ukuran kesamaan bentuk vektor. Hasil penelitian menunjukkan bahwa algoritma *k-medoids* baru dengan ukuran kesamaan yang diusulkan memiliki nilai validasi indeks *purity* dan indeks *davies bouldin* yang tinggi dibandingkan dengan algoritma *k-means* untuk sekumpulan data campuran.

Tahun 2016 penelitian oleh (Nooraeni, 2016) dengan judul “Metode *Cluster* Menggunakan Kombinasi Algoritma *K-Prototype* dan Algoritma Genetika untuk Data Bertipe Campuran”. Pada penelitian ini metode penggabungan *k-prototype* dengan algoritma genetika mampu menghasilkan inisial pusat klaster yang optimal. Hasil nilai *index clustering criterion* model *K-Prototype-Genetika* untuk *center* klaster optimal merupakan model terbaik dikarenakan memiliki nilai yang lebih tinggi dari yang lainnya yakni 0.0054 dibandingkan 0.0031. Hal tersebut menunjukkan bahwa tingkat kesamaan ciri pada setiap kelompok yang terbentuk pada model *K-Prototype-Genetika* untuk optimasi inisial *cente cluster* lebih mirip.

Penelitian di tahun 2017 oleh (Ikasakti, 2017) yang berjudul “Efektivitas Algoritma *K-Means*, *K-Medoids*, Dan Optimal Jumlah *Cluster* Untuk Modifikasi Algoritma *K-Medoids* Pada Kumpulan Data Besar”. Penelitian tersebut membahas tentang perbandingan rata-rata waktu komputasi pada algoritma *k-means*, *k-medoids* serta modifikasi *k-medoids*. Hasilnya adalah ketiga algoritma tersebut berbeda pada rata-rata waktu komputasi untuk sekumpulan data yang besar. Algoritma modifikasi *k-medoids* memiliki waktu yang lebih cepat dibandingkan dengan algoritma *k-means* dan *k-medoids*.

Masih di tahun yang sama yaitu 2017, (Putri, 2017) dengan penelitian berjudul “Pengelompokan Perguruan Tinggi Negeri di Indonesia Menggunakan Metode *Ensemble Robust Clustering Using Links*”. Pada penelitian tersebut, pengelompokan dilakukan pada data campuran menggunakan metode *ensemble ROCK (Robust Clustering using linKs)*. Pada metode tersebut data numerik dikelompokkan menggunakan metode hirarki dan didapatkan metode terbaik berupa *single linkage* dengan jumlah *k* klaster optimal sebanyak 3. Pada data kategorik menggunakan metode *ROCK* dan memberikan hasil bahwa jumlah *k* klaster optimal sebanyak 2 dengan nilai θ sebesar 0.5. Pengelompokan gabungan

yaitu data campuran menggunakan metode *ensemble ROCK* dengan nilai θ sebesar 0.3 yang didapatkan ketika k klaster sebanyak 2 yakni kelompok kualitas rendah sebanyak 42 PTN dan kualitas tinggi sebanyak 25 PTN.

Pada tahun 2019, (Weksi Budiaji, 2019) melakukan penelitian dengan judul “*Simple K-Medoids Partitioning Algorithm for Mixed Variable Data*” di mana penelitian tersebut bertujuan untuk mengusulkan algoritma *k-medoids* yang disebut dengan *Simple K-Medoids Partitioning* (SKM) yang mampu meningkatkan kinerja algoritma *Simple and Fast K-Medoids* (SFKM). Perhitungan jarak menggunakan jarak *gower*, *wishart*, *ahmad*, *harikumar*, *podani* dan *huang*. Algoritma SKM lalu didemonstrasikan pada dataset buatan dan nyata. Adanya parameter s atau *seeding* sebagai jumlah inisialisasi medoid awal, di mana ketika jumlah $s = 20$ dapat memberikan hasil yang serupa dengan PAM. Hasil pengujian menunjukkan bahwa kinerja algoritma SKM memiliki hasil yang lebih baik dan konsisten dibandingkan dengan SFKM. Selain itu, kinerja SKM membutuhkan waktu yang lebih sedikit dibandingkan PAM. Namun, SKM dan PAM setara untuk n dan k yang kecil.

Masih ditahun yang sama, penelitian berjudul “Penerapan Algoritma *K-Medoids Clustering* Untuk Pengelompokan Penyebaran Diare Di Kota Medan (Studi Kasus: Kantor Dinas Kesehatan Kota Medan)” yang dilakukan oleh (Riyanto, 2019) memiliki tujuan untuk mengelompokkan data penyebaran diare di kota Medan yang tersedia dalam suatu sistem. Hasil penelitian ini yaitu klaster 1 merupakan wilayah dengan penyakit diare yang tidak terlalu banyak. Sedangkan klaster 2 adalah wilayah dengan penyebaran diare paling banyak di Kota Medan.

Selain Weksi dan Riyanto, (Ivana Indriani Putri Damanik, 2019) melakukan penelitian yang berjudul “Algoritma *K-Medoids* untuk Mengelompokkan Desa yang Memiliki Fasilitas Sekolah di Indonesia”. Metode *k-medoids* diterapkan pada data persentase fasilitas sekolah berdasarkan provinsi. Hasil penelitian menunjukkan bahwa pengelompokan dilakukan dalam 3 klaster yaitu klaster rendah 15 provinsi, klaster sedang 16 provinsi dan klaster tinggi 3 provinsi.

Penelitian yang dilakukan (Hariyani, 2021) dengan judul “Pengelompokan Perguruan Tinggi Negeri Di Indonesia Menggunakan Metode Hirarki dan *K-Medoids* Dengan Ukuran Jarak Modifikasi Data Campuran (Studi Kasus : Perguruan Tinggi Negeri Di Indonesia Tahun 2020)”. Penelitian ini menggunakan

klaster hirarki *complete linkage* dan *k-medoids*. Hasil penelitian ini adalah perbedaan antara kedua metode yang digunakan yaitu pada jumlah k klaster yang optimum. Metode *complete linkage* menghasilkan 2 klaster sedangkan *k-medoids* menghasilkan 4 klaster. Sehingga didapatkan klaster yang terbentuk dari metode yang terpilih yaitu metode *complete linkage*.

2.1. Penelitian Terdahulu

Penelitian yang mendukung pada penulisan ini disajikan pada **Tabel 2.1**.

Tabel 2.1 Tabel Penelitian Sebelumnya

Nama	Judul	Hasil Penelitian
Zhexue Huang (1997)	<i>Clustering Large Data Sets With Mixed Numeric And Categorical Values</i>	Mengembangkan metode untuk memperbarui <i>k-prototype</i> secara dinamis untuk memaksimalkan kesamaan antar objek dalam klaster.
Hae-Sang Park, Chi-Hyuck Jun (2009)	<i>A Simple and Fast Algorithm for K-Medoids Clustering</i>	Algoritma SFKM merupakan algoritma yang membutuhkan waktu yang lebih sedikit dibandingkan dengan algoritma PAM.
P. IndiraPriya, Dr.D.K.Ghosh (2013)	<i>A Survey on Different Clustering Algorithms in Data Mining Technique</i>	Teknik pengelompokan <i>soft</i> dan metode pengelompokan hirarki menjadi metode terbaik untuk pengelompokan data yang besar.
Sandhya Harikumar, Surya PV (2015)	<i>K-Medoid Clustering for Heterogeneous DataSets</i>	Ukuran kesamaan dan algoritma <i>k-medoids</i> baru untuk data campuran memberikan hasil validasi yang tinggi jika dibandingkan dengan algoritma <i>k-means</i> .
Rani Nooraeni (2016)	Metode <i>Cluster</i> Menggunakan Kombinasi Algoritma <i>Cluster K-Prototype</i> dan Algoritma Genetika untuk Data Bertipe Campuran	Metode gabungan yaitu algoritma <i>K-Prototype-Genetika</i> mampu memberikan hasil inisial pusat <i>cluster</i> yang optimal jika dibandingkan dengan algoritma <i>K-Prototype</i> tanpa Genetika. Model <i>K-Prototype-GA</i> pada pengelompokan desa menggunakan dataset PODES 2011 memiliki tingkat akurasi yang lebih baik dibandingkan model <i>cluster</i> dengan <i>K-Prototype</i> tanpa Genetika.
Nur Aini Ikasakti (2017)	Efektivitas Algoritma <i>K-Means</i> , <i>K-Medoids</i> , Dan Optimal Jumlah <i>Cluster</i> Untuk Modifikasi Algoritma <i>K-Medoids</i> Pada Kumpulan Data Besar	Rata – rata waktu komputasi pada sekumpulan data yang besar untuk modifikasi algoritma <i>k-medoids</i> lebih cepat dibandingkan algoritma <i>k-means</i> dan <i>k-medoids</i> .
Indriyana Permata Putri (2017)	Pengelompokan Perguruan Tinggi Negeri di Indonesia Menggunakan	Pengelompokan dilakukan pada data campuran menggunakan metode <i>ensemble ROCK (Robust Clustering using linKs)</i> . Pada metode tersebut data numerik dikelompokkan menggunakan

Nama	Judul	Hasil Penelitian
	Metode <i>Ensemble Robust Clustering Using Links</i>	metode hirarki dan diperoleh metode terbaik berupa <i>single linkage</i> dengan jumlah k klaster optimal sebanyak 3. Pada data kategorik menggunakan metode <i>ROCK</i> dan memberikan hasil bahwa jumlah k klaster optimal sebanyak 2 dengan nilai θ sebesar 0.5. Pengelompokan gabungan yaitu data campuran menggunakan metode <i>ensemble ROCK</i> dengan nilai θ sebesar 0.3 yang didapatkan ketika k klaster sebanyak 2 yakni kelompok kualitas rendah sebanyak 42 PTN dan kualitas tinggi sebanyak 25 PTN
Bambang Riyanto (2019)	Penerapan Algoritma <i>K-Medoids Clustering</i> Untuk Pengelompokan Penyebaran Diare Di Kota Medan (Studi Kasus: Kantor Dinas Kesehatan Kota Medan)	Pengelompokan dilakukan menjadi 2 bagian yaitu klaster pertama merupakan wilayah dengan penyakit diare yang tidak terlalu banyak. Sedangkan klaster kedua adalah wilayah dengan penyebaran diare yang paling banyak di Kota Medan.
Ivana Indriani Putri Damanik, dkk (2019)	Algoritma <i>K-Medoids</i> untuk Mengelompokkan Desa yang Memiliki Fasilitas Sekolah di Indonesia	Pengelompokan dilakukan dalam 3 klaster yaitu klaster rendah sebanyak 15 provinsi, klaster sedang 16 provinsi dan klaster tinggi 3 provinsi.
Weksi Budiaji, Friedrich Leisch (2019)	<i>Simple K-Medoids Partitioning Algorithm for Mixed Variable Data</i>	Kinerja dari algoritma SKM memiliki hasil yang lebih baik dan konsisten dibandingkan dengan SFKM. Selain itu, kinerja SKM membutuhkan waktu yang lebih sedikit dibandingkan dengan PAM. Namun, SKM dan PAM setara untuk n dan k yang kecil.
Nabila Puspa Hariani (2021)	Pengelompokan Perguruan Tinggi Negeri Di Indonesia Menggunakan Metode Hirarki dan <i>K-Medoids</i> Dengan Ukuran Jarak Modifikasi Data Campuran (Studi Kasus: Perguruan Tinggi Negeri Di Indonesia Tahun 2020)	Perbedaan antara metode <i>complete linkage</i> dan <i>k-medoids</i> terletak pada jumlah k klaster yang optimum. Pada metode <i>complete linkage</i> menghasilkan 2 klaster sedangkan metode <i>k-medoids</i> menghasilkan 4 klaster. Didapatkan komponen klaster yang terbentuk dari metode yang terpilih yaitu metode <i>complete linkage</i> .

BAB III

LANDASAN TEORI

3.1. Asesmen Nasional

Asesmen Nasional atau AN dirancang untuk membantu dalam pemantauan dan evaluasi pada sistem pendidikan untuk jenjang dasar dan menengah. Pendidik serta satuan pendidikan akan mengevaluasi prestasi dari murid. Sesuai dengan bunyi Undang – Undang Sistem Pendidikan Nasional (UU Sisdiknas) pada pasal 57 ayat 1 yakni “Evaluasi dilakukan dalam rangka pengendalian mutu pendidikan secara nasional ...” serta pasal 59 ayat 1 yang berbunyi “Pemerintah dan Pemerintah Daerah akan melakukan evaluasi terhadap pengelola, satuan, jalur, jenjang dan jenis pendidikan”.

Elemen kebijakan Asesmen Nasional	Hasil dan/atau dampak yang diharapkan
AN dilaksanakan di semua sekolah/madrasah dan program pendidikan kesetaraan .	Pemetaan dan potret mutu SD/MI, SMP/MTs, dan SMA/MA di semua daerah.
AN dilaksanakan setiap tahun dan dilaporkan pada setiap sekolah/madrasah dan pemda.	(a) Kinerja sistem terantau secara berkala , dan (b) hasil AN digunakan untuk evaluasi diri .
AN 2021 digunakan sebagai <i>baseline</i> , tanpa konsekuensi pada guru, sekolah, dan pemda.	Mengurangi kecemasan pemangku kepentingan dan menghilangkan tekanan untuk curang.
Evaluasi kinerja tidak hanya berdasarkan skor rerata tapi juga perubahan skor atau trend dari satu tahun ke tahun berikutnya.	(a) Evaluasi kinerja diyakini lebih adil karena memperhitungkan posisi awal yang beragam, dan (b) mendorong orientasi pada perbaikan , bukan pada perbandingan antar sekolah/daerah.
AN hanya diikuti sebagian (sampel) murid yang dipilih secara acak dari kelas 5, 8, dan 11 di setiap sekolah/madrasah.	(a) Menegaskan bahwa AN bukan evaluasi individu murid , dan (b) tidak menambah beban murid kelas 6, 9 dan 12.

Gambar 3.1 Elemen Kebijakan dan Dampak Asesmen Nasional

Dalam pelaksanaan AN harus didukung dengan elemen – elemen kebijakan yang dapat memudahkan dan mendukung terselenggaranya AN. Salah satu elemen kebijakan yang diambil adalah peserta AN hanya diikuti oleh sebagian atau sampel dari para murid yang dipilih secara random dari kelas 5, 8 dan 11 pada setiap sekolah/ madrasah. Sehingga AN bukan merupakan evaluasi individu murid serta tidak akan menjadi beban untuk murid kelas 6, 9 dan 12.

AN sendiri terdiri dari tiga penilaian yaitu Asesmen Kompetensi Minimum (AKM) Literasi – Numerasi, Survei Karakter dan Survei Lingkungan Belajar. Sehingga diharapkan informasi dari ketiga penilaian tersebut mampu memberikan dorongan perbaikan pada mutu pembelajaran. Dengan adanya AN maka akan

memberikan potret yang komprehensif bagi sekolah/ madrasah, Pemerintah Daerah dalam mengevaluasi diri dan merencanakan perbaikan mutu pendidikan (Kemendikbud, 2021).

3.1.1 AKM Literasi – Numerasi

Literasi membaca dan numerasi merupakan kompetensi paling dasar yang dibutuhkan oleh seluruh murid agar dapat belajar sepanjang hidupnya serta memberikan kontribusi kepada masyarakat. Pengukuran kemampuan literasi dan numerasi memberi dorongan pada guru untuk fokus mengembangkan daya nalar dibandingkan pengetahuan konten yang luas namun dangkal (Kemendikbud, 2021).

3.1.2 Survei Karakter

Pengukuran karakter yang mendalam sangat sulit dilakukan pada asesmen dengan skala yang besar. Namun, dengan adanya survei karakter akan memberikan informasi penting mengenai sikap, nilai dan juga kebiasaan yang mencerminkan pada komponen pelajar pancasila. Adanya survei karakter, mendorong sekolah untuk memberi perhatian terhadap tumbuh dan kembang murid secara utuh, yang meliputi dimensi kognitif, afektif dan spiritual (Kemendikbud, 2021).

3.1.3 Survei Lingkungan Belajar

Adanya survei lingkungan belajar bertujuan untuk mengukur lima elemen yaitu (a) kualitas kegiatan pembelajaran, (b) iklim keamanan dan inklusivitas sekolah, (c) refleksi guru, (d) perbaikan pada praktik mengajar, dan (e) latar belakang dari keluarga murid. Hasil informasi survei lingkungan belajar berguna dalam mendiagnosis permasalahan dan rencana perbaikan kegiatan pembelajaran oleh guru, kepala sekolah serta dinas pendidikan (Kemendikbud, 2021).

3.2. Sekolah Menengah Atas (SMA)

Pendidikan menengah merupakan pendidikan lanjutan dari pendidikan dasar. Pendidikan ini ditempuh dalam jangka waktu tiga sampai empat tahun. Jenis pendidikan menengah meliputi Sekolah Menengah Atas (SMA), Sekolah Menengah Kejuruan (SMK), dan Sekolah Keagamaan (Madrasah Aliyah) serta sederajat (Paket C dan lainnya) (Umi Wahyuningsih Muhadi, 2017).

Sekolah Menengah Atas yang disingkat dengan SMA merupakan salah satu bentuk satuan pendidikan formal yang menyelenggarakan pendidikan dari proses lanjutan hasil belajar pada tingkat SMP atau MTs. Peserta didik SMA adalah warga

negara dengan usia antara 16 sampai 18 tahun yang mana pada usia tersebut sedang mengalami fase remaja yaitu antara 10 sampai 19 tahun.

SMA tidak termasuk dalam program wajib belajar yang ditetapkan oleh pemerintah. Program wajib belajar ini meliputi SD (atau sederajat) selama 6 tahun dan SMP (atau sederajat) selama 3 tahun. Meski sejak pada tahun 2005, program wajib belajar 12 tahun telah diberlakukan terdapat beberapa daerah yang mengikutsertakan program ini di antaranya adalah Kota Yogyakarta dan Kabupaten Bantul.

3.2.1 Kesiapan TIK untuk ANBK

Adapun beberapa komponen kesiapan dalam menyelenggarakan ANBK, dengan penjelasan sebagai berikut:

a. Guru

Menurut Peraturan Pemerintah Nomor 74 Tahun 2008 pada Bab I Pasal 1, menjelaskan pengertian guru adalah pendidik profesional yang memiliki tugas utama untuk mendidik, mengajar, membimbing, mengarahkan, melatih, menilai serta mengevaluasi peserta didik. Peserta didik berada dalam pendidikan anak usia dini dengan jalur pendidikan formal, dasar dan menengah.

b. Siswa

Dalam Undang – Undang Nomor 20 Tahun 2003 menjelaskan bahwa siswa atau peserta didik merupakan anggota masyarakat yang mempunyai usaha untuk mengembangkan potensi diri melalui proses pembelajaran yang ada pada jenjang dan jenis pendidikan tertentu.

c. Status Sekolah

Pada umumnya kedudukan sekolah di Indonesia dibagi menjadi dua yaitu sekolah negeri dan sekolah swasta. Sekolah negeri merupakan sekolah yang dipersiapkan oleh pemerintah dengan berbagai fasilitas gratis, seperti kelas, hingga guru yang digaji oleh pemerintah (Wikipedia, 2021). Sedangkan sekolah swasta adalah sekolah yang tidak dioperasikan oleh pemerintah, melainkan oleh yayasan, individu, maupun yang lainnya (Wikipedia, 2018).

d. Akreditasi Sekolah

Akreditasi adalah suatu proses penilaian sekolah dengan indikator tertentu yang berbasis fakta. Tujuan akreditasi yaitu untuk mendapatkan gambaran keadaan

kinerja sekolah dalam menyelenggarakan pendidikan. Akreditasi menjadi dasar untuk alat pembinaan serta pengembangan dalam rangka peningkatan mutu pendidikan. Pemeringkatan akreditasi terbagi dalam tiga klasifikasi yaitu sangat baik (A), baik (B) dan cukup (C) (Asmani, 2011).

e. Moda Pelaksanaan TIK

Pelaksanaan AKM diselenggarakan dalam dua moda yaitu online dan semi online. Pada moda online, pelaksanaannya harus mempunyai akses internet yang stabil untuk komputer *client*. *Bandwidth* yang disarankan untuk digunakan yakni 12 Mbps per 15 komputer *client*. Selain komputer *client*, diharuskan mempunyai komputer proktor yang berguna untuk membuka token.

Sedangkan untuk moda semi online, komputer *client* tidak perlu mempunyai akses internet secara langsung. Nantinya komputer *client* akan terhubung pada komputer proktor yang mempunyai akses internet. Pada moda semi online hanya dibutuhkan *bandwidth* minimal 1 Mbps (SMP, 2021).

Adapun infrastruktur TIK pada satuan pendidikan yang disesuaikan dengan moda pelaksanaan TIK yang diambil disajikan pada **Gambar 3.2**.

	Semi-online	Online
Komputer Proktor	Berbentuk <i>desktop PC/all in one</i> bukan laptop, CPU 4 core, RAM 8 GB, HD Free 250 GB, OS: Windows 10 Perbandingan 1 server diakses 15 komputer client	Berbentuk <i>desktop PC/all in one /laptop</i> Terhubung Internet
Komputer klien	Berbentuk <i>desktop PC/all in one/laptop/Thin Client</i> CPU dual Core, Monitor 11,6", RAM 2 GB, Resolusi minimal 1024 x 720, HD Free minimal 10 GB, OS: Windows, Linux, Chrome OS, Mac OS, Rasberry	Berbentuk <i>desktop PC/all in one/laptop</i> , CPU dual Core, Monitor 11,6", RAM 2 GB, Resolusi minimal 1024 x 720, HD Free minimal 10 GB, web kamera (optional). OS: Windows, Linux, Chrome OS, Mac OS, Rasberry
Aplikasi	VHD dijalankan virtualbox (Akan dibagikan pada Rakor) Browser proktor Exam browser klien	(tidak ada) Browser proktor Exam browser klien
Topologi	Bandwith internet 1 Mbps (untuk sinkronisasi) IP statik dengan segmen 192.168.0.XXX LAN (bukan wifi)	Bandwith internet 12 Mbps/15 komputer klien (stabil selama asesmen berlangsung dan dedicated) IP dinamis LAN/wifi

Gambar 3.2 Infrastruktur Moda Pelaksanaan TIK

Moda pelaksanaan TIK yang ditetapkan oleh pemerintah terdiri dari dua kategori yaitu online dan semi online. Perbedaan antara kedua moda tersebut terdapat pada empat hal yang meliputi komputer proktor, komputer klien, aplikasi dan topologi.

f. Status Pelaksanaan TIK

Persiapan dalam rangka penyelenggaraan ANBK Tahun 2021, pemerintah akan memetakan data kesiapan satuan pendidikan. Salah satunya adalah status

pelaksanaan TIK dari suatu sekolah. Pengkategorian ini dibagi menjadi tiga dengan penjelasan sebagai berikut (Pendidikan, 2021):

- 1) Mandiri adalah status pelaksanaan TIK di mana suatu sekolah mempunyai kesiapan serta kemampuan individu untuk dapat menyelenggarakan ANBK di sekolah tersebut.
- 2) Mandiri dan Ditumpangi adalah status pelaksanaan TIK di mana suatu sekolah mempunyai kesiapan dan kemampuan individu serta bersedia memberikan kesempatan pada sekolah lain untuk ikut dapat menyelenggarakan ANBK di sekolah tersebut.
- 3) Menumpang adalah status pelaksanaan TIK di mana suatu sekolah belum mempunyai kesiapan dan kemampuan individu sehingga harus turut ikut melaksanakan ANBK di sekolah yang mempunyai kemampuan individu untuk menyelenggarakan ANBK.

3.3. Sekolah Kita

Sekolah Kita merupakan aplikasi yang dikembangkan oleh Kementerian Pendidikan dan Kebudayaan (Kemendikbud). Kemendikbud menjadi lembaga pemerintah yang mempunyai tugas untuk mengurus seluruh perihal yang berkaitan dengan pendidikan serta kebudayaan di Indonesia. Adapun tujuan pengembangan aplikasi Sekolah Kita yakni memberikan kemudahan bagi masyarakat yang mempunyai kebutuhan dan kepentingan dalam mengakses komponen sekolah di Indonesia. Seluruh sekolah, baik negeri maupun swasta dapat ditemukan dalam aplikasi ini, apabila sudah terdaftar di Kemendikbud.

Sekolah Kita bukan aplikasi yang berbasis *mobile* ataupun desktop, tetapi aplikasi ini berbasiskan website. Sehingga, aplikasi ini tidak dapat dipasang dalam komputer maupun ponsel. Akan tetapi, dapat mengaksesnya melalui browser yang ada pada perangkat misalnya *Google Chrome*, *Mozilla Firefox*, *Microsoft Edge* dan lainnya. Aplikasi ini beralamatkan di <https://sekolah.data.kemdikbud.go.id/>. Data yang ada dalam aplikasi Sekolah Kita adalah hasil dari sinkronisasi yang dikelola oleh masing – masing sekolah. Data tersebut merupakan kumpulan data dari penyelenggaraan seluruh jenjang sekolah yang terintegrasi secara nasional.

Dalam penyelenggaraan AKM, setiap sekolah akan dilakukan pendataan mengenai kesiapan perlengkapan TIK yang akan menunjang pelaksanaan tersebut.

Pada klasterisasi yang dilakukan oleh Kemendikbud mengenai kesiapan TIK sekolah terbagi dalam 4 definisi yang menentukan kesiapan dalam penyelenggaraan AKM yaitu (Kemendikbud Ristek, 2020):

1) Siap

Kategori Siap ini terbagi dalam 3 bagian yakni:

- a) Tipe A, merupakan klasifikasi sekolah yang sudah menyelenggarakan UNBK secara mandiri (pada tahun sebelumnya).
- b) Tipe B, merupakan klasifikasi sekolah yang mempunyai komputer minimal 15 unit dan mempunyai aliran listrik serta akses internet.
- c) Tipe C, merupakan klasifikasi sekolah yang mempunyai komputer kurang dari 15 unit, mempunyai aliran listrik, akses internet serta memperoleh Bantuan Pusat tahun 2020 dan Bantuan DAK 2021.

2) Potensial 1, merupakan klasifikasi sekolah yang mempunyai komputer kurang dari 15 unit, dan mempunyai aliran listrik serta akses internet.

3) Potensial 2, merupakan klasifikasi sekolah yang mempunyai komputer kurang dari 15 dan hanya memiliki aliran listrik.

4) Tidak Siap, apabila sekolah tidak mempunyai fasilitas atau prasarana yang mendukung pelaksanaan AKM, seperti komputer, aliran listrik dan jaringan internet.

3.4. Data

3.4.1 Jenis Data

Pada analisis statistik terdapat dua jenis data yang dapat digunakan. Jenis data tersebut adalah data numerik (atau kuantitatif) dan data kategorik (atau kualitatif). Adapun penjelasan tentang data numerik dan kategorik sebagai berikut:

a. Data Numerik

Data kuantitatif atau numerik merupakan suatu data yang diperoleh dari hasil pengukuran suatu objek (Muhyidin, 2020). Cirinya data berupa angka. Misalnya tinggi badan, berat badan, penghasilan dan lainnya.

b. Data Kategorik

Data kualitatif atau kategorik merupakan suatu data dari hasil pengklasifikasian atau penggolongan suatu data (Muhyidin, 2020). Cirinya data berupa kata – kata. Misalnya jenis kelamin, jenis pekerjaan, pendidikan dan lainnya.

3.4.2 Skala Pengukuran

a. Skala Nominal

Skala pengukuran yang paling sederhana yaitu skala pengukuran nominal. Skala nominal ini menempatkan objek atau individu ke dalam kategori-kategori yang memiliki perbedaan kualitatif, bukan kuantitatif. Penggolongan mahasiswa berdasarkan jenis kelaminnya yaitu laki – laki (1) atau perempuan (0) misalnya merupakan contoh pengukuran dengan skala nominal (Syafri, 2010).

b. Skala Ordinal

Persyaratan pokok dalam pengukuran skala ordinal yaitu adanya kriteria empiris untuk menyusun objek atau kejadian – kejadian dalam hubungannya dengan atribut tersebut. Skala pengukuran ini digunakan dalam menentukan ranking suatu kelompok tertentu. Pada ranking hanya dipertimbangkan urutan objek dari yang terbesar sampai terkecil. Misalnya pada pemeringkatan juara kelas yaitu dari ranking 1, 2 dan 3 (Syafri, 2010).

c. Skala Interval

Skala interval yakni skala yang memberikan jarak interval yang sama dari suatu titik asal yang tidak tetap. Pada skala interval hubungan tata urutan dan jarak antara angka – angka memiliki makna. Contohnya, perbedaan antara IQ 90 dan 91 sama dengan perbedaan antara IQ 100 dan 101. Tetapi, tidak menyatakan bahwa orang dengan IQ 120 merupakan dua kali orang yang ber IQ 60 (Syafri, 2010) .

d. Skala Rasio

Skala rasio adalah skala pengukuran tertinggi yang mana skala ini memiliki nilai nol mutlak. Pada skala rasio terdapat seluruh karakteristik skala nominal, ordinal dan interval. Nilai *no* mutlak mempunyai makna bahwa nilai dasar yang tidak dapat diubah meskipun menggunakan skala yang lain. Misalnya berat benda A 30 kg, sedangkan berat benda B 60 kg. maka dapat dikatakan bahwa berat benda B dua kali lebih berat dibandingkan benda A (Syafri, 2010).

3.5. Analisis Deskriptif

Sebelum melakukan analisis yang lebih jauh, dalam suatu penelitian pada umumnya akan dilakukan analisis deskriptif terlebih dahulu. Menurut Sugiyono (2017:147) analisis deskriptif adalah suatu metode yang digunakan dalam analisis data dengan cara memberikan deskripsi atau gambaran mengenai data yang sudah

dikumpulkan sebagaimana adanya tanpa ada maksud untuk membuat suatu kesimpulan yang berlaku secara umum. Umumnya, penyajian data dalam analisis deskriptif dapat berupa tabel, grafik serta ukuran statistik. Visualisasi grafik sangat memudahkan untuk memahami karakteristik data dapat berupa *bar plot*, *pie chart*, *boxplot* dan lainnya. Ukuran statistik yang umum digunakan yaitu nilai maksimum, minimum, ukuran tendensi sentral seperti *mean*, *median* serta ukuran penyebaran data berupa kuartil, *range*, variansi dan standar deviasi (Yuvalianda, 2019).

3.6. Analisis Multivariat

Analisis multivariat merupakan metode untuk mengolah variabel dalam jumlah yang banyak. Tujuannya yaitu untuk mengetahui pengaruh antar variabel yang ada terhadap objek secara serentak (Hidayat, 2016). Analisis ini menjadi analisis yang melibatkan suatu variabel dengan jumlah lebih dari atau sama dengan 3 variabel. Terdapat dua macam teknik analisis multivariat yaitu:

a) Teknik Dependen

Didefinisikan sebagai metode yang digunakan ketika variabel diketahui sebagai variabel terikat yang dijelaskan oleh variabel bebas (Hidayat, 2016). Contoh analisisnya adalah *Multiple Regression*, *Multiple Discriminant Analysis*, *Logistic Regression*, MANOVA, *Conjoint Analysis* dan lainnya.

b) Teknik Interdependen

Merupakan suatu metode yang mana tidak terdapat variabel yang didefinisikan sebagai variabel bebas atau variabel terikat (Hidayat, 2016). Contoh analisisnya adalah *Principal Components & Common Factor Analysis*, *Cluster Analysis*, *Multidimensional Scaling* dan lainnya.

3.6.1 One-Way MANOVA

Salah satu analisis multivariat yang masuk dalam teknik dependen adalah MANOVA. *Multivariate Analysis of Variance* (MANOVA) merupakan uji statistik yang berfungsi untuk mengukur seberapa besar pengaruh antara variabel terikat berskala kategorik dengan beberapa variabel bebas secara simultan. Jika asumsi normalitas multivariat dan matriks kovarian antar populasi homogen telah terpenuhi maka uji MANOVA dapat dilakukan (Johnson, 2007).

3.7. Analisis Klaster

Analisis klaster yakni suatu alat eksplorasi penting dalam suatu penyelidikan struktur data. Setiap objek yang berada dalam suatu kelompok bersifat homogen satu dengan lainnya, dan objek yang berada dalam kelompok lain bersifat heterogen (Kaufman, 1990). Pada pengelompokan berbasis jarak, kuantifikasi kesamaan antar objek didasarkan pada jaraknya, di mana benda yang lebih mirip mempunyai jarak yang lebih dekat dan sebaliknya. Setelah jarak didefinisikan, baik algoritma hirarki maupun non hirarki diterapkan (Weksi Budiaji, 2019).

Salah satu algoritma non hirarki (partisi) yang populer adalah *k-means*, yang merupakan salah satu dari sepuluh algoritma teratas dalam data mining. Algoritma *k-means* tidak relevan ketika data yang digunakan adalah data variabel campuran, karena “*means*” sebagai pusat klaster (*centroid*) tidak tersedia serta perhitungan jarak *Euclidean* tidak berlaku. Pada praktiknya, algoritma yang umum digunakan untuk data dengan variabel campuran adalah *Partitioning Around Medoids (PAM)* yang menggantikan *centroid* dengan medoid (Weksi Budiaji, 2019).

Pendekatan lain untuk menghitung jarak data dengan variabel campuran adalah melalui fungsi jarak gabungan antara jarak numerik dan kategorik. Fungsi jarak gabungan mengasumsikan bahwa setiap variabel memiliki kontribusi yang sama. Namun, ketika variabel berkontribusi secara berbeda, maka dapat menerapkan bobot yang tidak sama, jarak *Gower* merupakan jarak yang paling umum untuk data dengan variabel campuran.

Setelah menentukan jarak untuk data dengan variabel campuran, maka baik algoritma *k-prototype* atau PAM dapat diterapkan. Definisi pusat klaster dalam algoritma PAM yaitu mengambil lokasi paling sentral (medoid) sebagai pusat klaster untuk setiap fungsi jarak. Sehingga PAM menjadi algoritma berbasis medoid yang lebih kuat daripada algoritma berbasis *centroid*.

3.7.1 Ukuran Kedekatan (*Similarity*) Antar Objek

Pada analisis klaster atau pengelompokan, ukuran kedekatan antar objek menjadi hal yang paling dasar. Ketika proses dalam memisahkan objek dari sekumpulan data, diperlukan ukuran kemiripan serta ketidakmiripan dalam analisis pengelompokan. Pada pengelompokan data kategorik diperlukan ukuran kemiripan sedangkan ukuran ketidakmiripan digunakan untuk data numerik. Apabila ukuran

jarak bernilai kecil maka semakin besar kemiripan di antara dua objek sehingga cenderung berada pada kelompok yang sama (Johnson, 2007).

Secara umum, jarak digunakan dalam perhitungan ukuran kemiripan atau *similarity* dan ketidakmiripan atau *dissimilarity*. Kemiripan secara umum dituliskan dengan s atau *similarity* menunjukkan ukuran numerik dari kesamaan diantara dua objek. Apabila nilai *similarity* semakin tinggi menandakan jika dua objek lebih mirip. Sedangkan ketidakmiripan atau *dissimilarity* merupakan ukuran numerik dari perbedaan diantara dua objek. Jika nilai jarak (*dissimilarity*) semakin rendah maka kedua objek tersebut lebih mirip.

Hal yang paling mempengaruhi hasil dari analisis pengelompokan adalah jarak antara satu objek dengan objek lainnya (Sharma, 1996). Maka, diperlukan alat ukur untuk menghitung jarak antar objek. Terdapat tiga metode pengukuran jarak berdasarkan jenis variabel data yaitu numerik, biner dan kategorik serta campuran.

a. Pengukuran Kedekatan Objek Variabel Numerik

Perhitungan kedekatan antar objek menggunakan rentang *manhattan* yang terbobot jika variabel yang diteliti merupakan variabel numerik atau variabel yang memiliki nilai angka (Weksi Budiaji, 2019). Ukuran kedekatan variabel numerik merupakan kemiripan (*similarity*) antar objek dituliskan pada **Persamaan 3.1**.

$$s_{ijl_{numerical}} = 1 - \frac{|x_{il} - x_{jl}|}{R_l} \quad (3.1)$$

dimana:

$s_{ijl_{numerical}}$ = kemiripan antara objek ke- i dan ke- j variabel numerik ke- l

x_{il} = nilai objek ke- i pada variabel numerik ke- l

x_{jl} = nilai objek ke- j pada variabel numerik ke- l

R_l = *range* variabel numerik ke- l

b. Pengukuran Kedekatan Objek Variabel Biner dan Kategorik

Pada variabel dengan dua kategori atau biner serta variabel dengan lebih dari dua kategori, perhitungan kedekatan antar objek akan menggunakan *simple matching*. Pada *simple matching* menggunakan variabel yang memiliki kontribusi sama dan tanpa adanya data *missing* (Weksi Budiaji, 2019). Ukuran kedekatan variabel biner dan kategorik merupakan kemiripan (*similarity*) antar objek ke- i dan ke- j yang dituliskan pada **Persamaan 3.2**.

$$S_{ijl_{binary/categorical}} \in \{0,1\} \quad (3.2)$$

dimana:

$S_{ijl_{binary/categorical}} = 1$; apabila objek ke- i dan ke- j mempunyai nilai sama pada variabel biner atau kategorik ke- l

$S_{ijl_{binary/categorical}} = 0$; apabila objek ke- i dan ke- j mempunyai nilai yang berbeda pada variabel biner atau kategorik ke- l

c. Pengukuran Kedekatan *Gower* Objek Variabel Campuran

Ukuran kemiripan (*similarity*) *Gower* pada variabel campuran antara objek ke- i dan objek ke- j dituliskan pada **Persamaan 3.3** (Weksi Budiaji, 2019).

$$s_{ij} = \frac{s_{ij}^{numerical} + s_{ij}^{binary} + s_{ij}^{categorical}}{p_n + p_b + p_c} \quad (3.3)$$

dimana:

s_{ij} = kesamaan *Gower* objek ke- i dan ke- j

$s_{ij}^{numerical}$ = kemiripan numerik objek ke- i dan ke- j

s_{ij}^{binary} = kemiripan biner objek ke- i dan ke- j

$s_{ij}^{categorical}$ = kemiripan kategorik objek ke- i dan ke- j

p_n, p_b, p_c = jumlah variabel numerik, biner, kategorik

d. Pengukuran Jarak *Gower* Objek Variabel Campuran

Kemiripan *Gower* (1971) digunakan untuk menghitung data dengan variabel campuran yaitu variabel numerik, biner dan kategorik (Weksi Budiaji, 2019). Jarak *Gower* (*dissimilarity*) antara objek ke- i dan ke- j didefinisikan pada **Persamaan 3.4**.

$$d_{ij} = 1 - s_{ij} \cdot \omega_{ijl} \quad (3.4)$$

Jarak *Gower* mengatur kontribusi dari masing-masing variabel. Jika terdapat data *missing* maka ω_{ijl} bernilai nol. Perhitungan jarak *Gower* mengasumsikan tidak adanya data *missing* dan setiap variabel memiliki kontribusi yang sama sehingga $\omega_{ijl} = 1$ (Weksi Budiaji, 2019). Sehingga d_{ij} (*dissimilarity*) dengan variabel tunggal dari variabel numerik, biner dan kategorik dituliskan pada **Persamaan 3.5**.

$$\begin{aligned} d_{ij} &= 1 - s_{ij} \\ &= 1 - \frac{s_{ij}^{numerical} + s_{ij}^{binary} + s_{ij}^{categorical}}{1+1+1} \end{aligned} \quad (3.5)$$

$$\begin{aligned}
&= 1 - \frac{s_{ij}^{numerical} + s_{ij}^{binary} + s_{ij}^{categorical}}{3} \\
&= \frac{3 - s_{ij}^{numerical} - s_{ij}^{binary} - s_{ij}^{categorical}}{3} \\
&= \frac{(1 - s_{ij}^{numerical}) + (1 - s_{ij}^{binary}) + (1 - s_{ij}^{categorical})}{3} \\
&= \frac{\left(1 - \left(1 - \frac{|x_i - x_j|}{R}\right)\right) + d_{ij}^{binary} + d_{ij}^{categorical}}{3} \\
&= \frac{\frac{|x_i - x_j|}{R} + d_{ij}^{binary} + d_{ij}^{categorical}}{3}
\end{aligned}$$

e. Jarak Generalized Distance Function (GDF)

Generalized Distance Function (GDF) dikembangkan untuk meningkatkan variasi campuran variabel. GDF terdiri atas kombinasi bobot serta jarak. Bobot tertentu digunakan untuk kombinasi jarak numerik, biner dan kategorik. Penggunaan GDF, dapat memformulasikan ulang jarak *Gower* yang merupakan kasus khusus dari GDF. Pendekatan GDF serupa dengan Huang (1997), Ahmad dan Dey (2007), McCane dan Albert (2008) serta Harikumar dan PV (2015) yang menggabungkan jarak numerik, biner dan kategorik yang didefinisikan dalam **Persamaan 3.6** (Weksi Budiaji, 2019).

$$d_{ij} = \left(\alpha \sum_{r=1}^{p_n} \delta_n(x_{ir}, x_{jr}) + \beta \sum_{t=1}^{p_b} \delta_b(x_{it}, x_{jt}) + \gamma \sum_{s=1}^{p_c} \delta_c(x_{is}, x_{js}) \right)^{\omega} \quad (3.6)$$

dimana:

- ω = bobot untuk seluruh fungsi
- α, β, γ = bobot masing-masing variabel numerik, biner dan kategorik
- δ_n = jarak variabel numerik
- δ_b = jarak variabel biner
- δ_c = jarak variabel kategorik
- (x_{ir}, x_{jr}) = jarak numerik antara objek ke- i dan ke- j pada variabel r
- (x_{it}, x_{jt}) = jarak biner antara objek ke- i dan ke- j pada variabel t
- (x_{is}, x_{js}) = jarak kategorik antara objek ke- i dan ke- j pada variabel s

Penerapan yang dilakukan dengan menggunakan persamaan GDF, jarak dapat disesuaikan untuk semua jenis data dan lebih fleksibel dengan kombinasi

jarak dan bobot apapun. Modifikasi GDF untuk jarak *Gower* apabila banyaknya variabel numerik, biner dan kategorik adalah p_n, p_b dan p_c maka jarak *Gower* (*dissimilarity*) antara objek ke- i dan ke- j didefinisikan pada **Persamaan 3.7**.

$$d_{ij} = \left(\frac{1}{p_n + p_b + p_c} \left(\sum_{r=1}^{p_n} \frac{1}{R_r} |x_{ir} - x_{jr}| + p_b \sum_{t=1}^{p_b} \delta_b(x_{it}, x_{jt}) + p_c \sum_{s=1}^{p_c} \delta_c(x_{is}, x_{js}) \right) \right) \quad (3.7)$$

dimana:

p_n = jumlah variabel numerik

p_b = jumlah variabel biner

p_c = jumlah variabel kategorik

R_l = *range* variabel numerik ke- r

$\delta_b(x_{it}, x_{jt}) = \delta_c(x_{is}, x_{js}) = 0$ = jika objek ke- i dan objek ke- j adalah sama

$\delta_b(x_{it}, x_{jt}) = \delta_c(x_{is}, x_{js}) = 1$ = jika objek ke- i dan objek ke- j adalah tidak sama

Pada **Persamaan 3.7** menyiratkan bahwa jarak untuk variabel numerik merupakan rentang terbobot *Manhattan*, sedangkan jarak untuk variabel biner dan kategorik merupakan *simple matching* (Weksi Budiaji, 2019). Adapun contoh studi kasus untuk perhitungan manual kedekatan dan jarak disajikan pada **Lampiran 17**.

3.7.2 Metode Pengelompokan Non Hirarki

Metode non hirarki adalah salah satu metode pengelompokan kumpulan data yang mana jumlah kelompok yang terbentuk sudah ditentukan dahulu pada awal proses pengelompokan. Metode ini dapat digunakan pada sekumpulan data yang jumlahnya lebih besar. Adapun penjelasan beberapa metode yang masuk dalam kategori non hirarki adalah sebagai berikut:

a. Algoritma *K-Medoids*

Hadirnya algoritma *k-medoids*, sehubungan untuk mengatasi kelemahan pada algoritma *k-means* yang sensitif dengan adanya data *outlier* (J. Han, 2006). Algoritma berbasis medoid yang paling populer adalah algoritma PAM. Secara umum, algoritma ini terdiri dari tiga langkah yaitu pemilihan medoid awal, pertukaran medoid serta pelabelan objek.

Selain PAM, banyak algoritma yang telah dikembangkan seperti algoritma *K-Medoids* (KM) dan *Simple and Fast K-Medoids* (SFKM). Namun, kedua algoritma tersebut membatasi pada langkah pertukaran medoid. Algoritma tersebut hanya bekerja di dalam kluster saja sehingga mirip dengan *k-means* pada tahap pembaruan *centroid*.

Algoritma KM mengadaptasi algoritma *K-Means* dengan pusat kluster terwakili oleh satu set medoid sebagai gantinya. Pada algoritma SFKM hampir sama dengan algoritma KM. Akan tetapi, keduanya juga memiliki perbedaan pada saat pemilihan medoid awal, yang mana algoritma KM menggunakan medoid awal yang dipilih secara random. Sedangkan, pada SFKM memilih medoid awal dengan menggunakan urutan dari satu set v_j (Budiaji, 2019). Nilai dari v_j didefinisikan sebagai jumlah baris atau jumlah kolom dari matriks jarak yang dihitung dengan **Persamaan 3.18**.

$$v_j = \sum_{i=1}^m \frac{d_{ij}}{\sum_{l=1}^n d_{il}}, j = 1, 2, 3, \dots, n \quad (3.8)$$

d_{ij} merupakan jarak antara benda ke- i dan benda ke- j . Ketika jumlah kluster yang ditentukan sebelumnya adalah k maka k set terkecil pertama dipilih dari v_j sebagai medoid awal. Akan tetapi, algoritma KM akan berhenti ketika medoid sebelumnya identik sedangkan SFKM berhenti ketika jumlah jarak objek ke medoidnya sama dengan langkah sebelumnya. Jarak antara medoid dengan objek dalam sebuah kluster didefinisikan dalam **Persamaan 3.9**.

$$E = \sum_{g=1}^k \sum_{x_n \in C_g} d(x_n, m_g) \quad (3.9)$$

dimana:

C_g = Grup ke- g yang mempunyai medoid m_g

Ketika jumlah kluster yang ditentukan sebelumnya adalah k , set terkecil k pertama dipilih dari v_j yang dijadikan sebagai medoid awal. Karena kesamaan dengan pembaruan *centroid* dalam *k-means*, algoritma KM akan mengalami kemungkinan masalah pada *local* optima dan kluster kosong tergantung nilai awal. Algoritma SFKM dapat diterapkan pada kumpulan data numerik dan kategorik (Budiaji, 2019).

b. Algoritma Simple K-Medoids Partitioning (SKM)

Algoritma SKM dikembangkan untuk mengelola *local* optima dan kluster kosong yang dapat terjadi dalam algoritma KM dan SFKM. Cara untuk menghindari *local* optima, rekomendasi dari Steinley and Brusco (2009) dengan menerapkan beberapa inisialisasi. Proses dalam mempertahankan label pada saat medoid non unik diinisialisasi merupakan alternatif untuk mengatasi kluster yang kosong. Adapun tahapan dalam algoritma SKM adalah sebagai berikut (Weksi Budiaji, 2019):

1. Pilih satu set *intial* medoid awal M_k dengan memilih k objek dari X_n .

Medoid awal M_k adalah objek yang terletak paling sentral. Inisialisasi acak yang diulang beberapa kali adalah pendekatan terbaik. Namun, pembatasan diterapkan pada satu titik dari himpunan M_k . Salah satu M_k yang dipilih secara acak digantikan oleh objek yang paling sentral. Medoid yang optimal adalah ketika $k = 1$ merupakan objek yang terletak pada pusat dengan meminimalkan jumlah jarak menuju objek lain. Kemudian, lokasi objek yang paling sentral dapat dipilih berdasarkan jumlah baris atau kolom dari matriks jarak yang dihitung menggunakan rumus pada **Persamaan 3.10**.

$$a_i = \sum_{j=1}^n d_{ij} = \sum_{j=1}^n d_{ji}, \quad i = 1, 2, 3, \dots, n \quad (3.10)$$

dimana:

a_i = jumlah jarak dari objek ke- i menuju objek ke- j

Apabila vektor a_i diurutkan dengan M_k pertama dipilih dari nilai yang terkecil, yang merupakan objek yang terletak paling sentral.

2. Tetapkan label atau keanggotaan dari setiap objek $x_n \in X_n, l(x_n)$ menuju ke medoid M_k yang paling dekat.
3. Perbarui set medoid M_k baru dengan mempertahankan label kluster $l(x_n)$ yang tetap dengan **Persamaan 3.11**.

$$m_p := \underset{p \in K}{\operatorname{argmin}} \sum_{n: l(x_n)=m_p} d(x_n, m_p), \quad p = 1, 2, 3, \dots, k \quad (3.11)$$

dimana:

m_p = objek yang meminimalkan jumlah jarak antar objek yang berada dalam satu kluster.

4. Hitung jumlah jarak setiap klasternya (E) yang merupakan jarak antara objek dengan medoidnya yang dituliskan dalam **Persamaan 3.12**.

$$E = \sum_{p=1}^k \sum_{n:l(x_n)=m_p} d(x_n, m_p) \quad (3.12)$$

5. Ulangi langkah 2 – 4 hingga nilai E bernilai sama dengan nilai E sebelumnya atau himpunan medoid tidak berubah atau jumlah iterasi yang ditentukan telah terpenuhi.
6. Ulangi langkah 1 – 5, untuk memperoleh beberapa *seeding* secara acak dari medoid awal. Perhatikan bahwa objek yang berada pada pusat kluster (paling sentral) akan selalu menjadi salah satu medoid awal. Penentuan medoid akhir (final) dipilih berdasarkan nilai E yang paling minimum.
7. Tetapkan keanggotaan dari setiap objek menuju ke medoid akhir terdekat.

Adapun contoh studi kasus untuk perhitungan manual dan tahapan dari algoritma SKM disajikan pada **Lampiran 17**.

3.7.3 Validasi Kluster

Validasi kluster adalah tahapan yang digunakan dalam melakukan evaluasi hasil analisis pengelompokan. Validasi kluster dapat diklasifikasikan menjadi tiga jenis yaitu kriteria eksternal, internal dan relatif (Weksi Budiaji, 2019). Penjelasan mengenai tiga pendekatan tersebut sebagai berikut:

- 1) Kriteria eksternal: memvalidasi hasil pengelompokan dengan memeriksa apakah objek terstruktur secara acak atau tidak (Halkidi, 2001).
- 2) Kriteria relatif: memvalidasi hasil pengelompokan dengan tidak memerlukan asumsi kekompakan. Hal tersebut bergantung pada set sampel yang direplikasi melalui teknik replikasi seperti metode validasi silang.
- 3) Kriteria internal: memvalidasi hasil pengelompokan yang didasarkan pada informasi intrinsik data. Biasanya untuk mengukur kekompakan dan pemisahan kluster. Kekompakan digunakan untuk menguji homogenitas kluster, yakni dalam kluster sedangkan pemisahan menilai derajat pemisahan antar kluster. Misalnya yaitu *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV) (Budiaji, Weksi, 2019).

a. *Silhouette Index* (SI)

Silhouette Index (SI) adalah ukuran validasi yang digunakan dalam pengujian keoptimalan atau ketepatan suatu klaster yang terbentuk dari proses pengelompokan. SI masuk dalam kategori kriteria internal. Pada kriteria internal ini menggunakan informasi dari dalam data (intrinsik) untuk mengukur kualitas pengelompokan (Budiaji, 2019).

Untuk menghitung nilai SI dari objek i , terdapat dua komponen yakni nilai kohesi yang disimbolkan dengan a_i dan nilai separasi yang disimbolkan dengan b_i . Nilai a_i adalah jarak rata – rata antara objek ke- i dengan seluruh objek lainnya yang berada dalam satu klaster, sedangkan b_i adalah hasil dari perhitungan jarak rata rata objek ke- i terhadap seluruh objek yang berada dalam klaster terdekat yang kemudian diambil nilai terkecilnya. Berikut adalah rumus perhitungan dari nilai SI (Puspitasari, 2016):

Rumus untuk menghitung nilai a_i^j dituliskan dalam **Persamaan 3.13**.

$$a_j^i = \frac{1}{m_j - 1} \sum_{\substack{r=1 \\ r \neq i}}^{m_j} d(x_i^j, x_r^j) \quad i = 1, 2, \dots m_j \quad (3.13)$$

dimana:

- i = indeks data
- j = klaster
- m_j = banyaknya data dalam klaster ke- j
- $d(x_i^j, x_r^j)$ = jarak data ke- i dengan data ke- r dalam satu klaster ke- j
- a_j^i = jarak rata – rata antara objek ke- i dengan semua objek yang ada dalam satu klaster

Rumus untuk menghitung nilai b_i^j dituliskan dalam **Persamaan 3.14**.

$$b_i^j = \min_{\substack{n=1, \dots, k \\ n \neq j}} \left\{ \frac{1}{m_n} \sum_{\substack{r=1 \\ r \neq i}}^{m_j} d(x_i^j, x_r^n) \right\} \quad i = 1, 2, \dots m_n \quad (3.14)$$

dimana:

- i = indeks data
- j = klaster
- n = klaster yang lain
- $d(x_i^j, x_r^n)$ = jarak data ke- i dengan data ke- r dalam satu klaster j terhadap data ke- r dalam klaster n

b_i^j = nilai terkecil dari jarak rata – rata antara objek ke- i dengan semua objek yang ada dalam klaster lain, tidak berada dalam satu klaster dengan data ke- i

Rumus untuk menghitung nilai SI data ke- i dituliskan dalam **Persamaan 3.15**.

$$SI_i^j = \frac{b_i^j - a_i^j}{\max(a_i^j, b_i^j)} \quad (3.15)$$

dimana:

SI_i^j = *Silhouette Index* dari data ke- i dalam satu klaster

b_i^j = nilai terkecil dari jarak rata – rata antara objek ke- i dengan semua objek yang ada dalam klaster lain, tidak berada dalam satu klaster dengan data ke- i

a_i^j = jarak rata – rata antara objek ke- i dengan semua objek yang ada dalam satu klaster

Nilai SI suatu klaster diperoleh dari rata – rata lebar *Silhouette* untuk seluruh data yang berada di dalam klaster tersebut, dengan rumus pada **Persamaan 3.16**.

$$SI_j = \frac{1}{m_j} \sum_{i=1}^{m_j} SI_i^j \quad (3.16)$$

dimana:

SI_j = rata – rata nilai *Silhouette Index* pada klaster j

m_j = banyaknya data dalam klaster ke- j

SI_i^j = nilai *Silhouette Index* dari data ke- i pada satu klaster j

Sedangkan untuk nilai SI global didapatkan dengan cara menghitung rata – rata dari nilai *Silhouette* untuk seluruh objek dengan rumus pada **Persamaan 3.17**.

$$SI = \frac{1}{N} \sum_{i=1}^{m_j} SI_i^j \quad (3.17)$$

dimana:

SI = nilai *Silhouette Index* global

SI_j = nilai *Silhouette Index* dari data ke- i pada satu klaster j

N = banyaknya data

Nilai SI berada dalam rentang -1 sampai 1. Ketika nilainya 1 menunjukkan klaster yang dihasilkan semakin “*well-separated*”. Sebaliknya, jika nilainya -1

menunjukkan bahwa klaster yang dihasilkan semakin “*poor-separated*”. Hasil rata – rata nilai SI pada suatu klaster untuk setiap objek merupakan tingkat kemiripan objek yang berada dalam klaster yang sama (Rousseeuw, 1984). Nilai Kaufman untuk struktur nilai SI disajikan pada **Tabel 3.1**.

Tabel 3.1 Nilai Kaufman Karakteristik *Silhouette Index*

<i>Silhouette Index</i>	Struktur
$0.7 < SC \leq 1$	Kuat
$0.5 < SC \leq 0.7$	Baik
$0.25 < SC \leq 0.5$	Lemah
$SC \leq 0.25$	Tidak Mempunyai Struktur

Adapun contoh studi kasus untuk perhitungan manual validasi klaster menggunakan *Silhouette Index* (SI) disajikan pada **Lampiran 17**.

b. Medoid Shadow Value (MSV)

Teknik validasi berbasis medoid dengan meniru *Silhouette Index* (SI) serta karakter nilai bayangan berbasis *centroid* dikembangkan oleh (Budiaji, Weksi, 2019). Teknik ini disebut dengan *Medoid Shadow Value* (MSV). Yang mana teknik ini akan menghasilkan bentuk yang hampir sama dengan SI. Perhitungan nilai MSV dapat dilakukan dengan menggunakan rumus pada **Persamaan 3.18**.

$$msv = \frac{d(x, m'(x)) - d(x, m(x))}{d(x, m'(x))} \quad (3.18)$$

dimana:

$d(x, m(x))$ = jarak antara objek ke-*i* dengan medoid terdekat pertama

$d(x, m'(x))$ = jarak antara objek ke-*i* dengan medoid terdekat kedua

Dikarenakan medoid yang menjadi pusat suatu klaster, maka dengan menggunakan metode jarak apapun dapat berlaku dalam nilai bayangan berbasis medoid (MSV). Adapun karakteristik dari nilai MSV di antaranya adalah:

- 1) Nilai batas bawah dan atas adalah 0 dan 1
- 2) Untuk klaster yang terpisah secara buruk (atau *poor-separated*) nilai MSV bernilai 0, sedangkan terpisah secara baik (atau *well-separated*) bernilai 1.
- 3) Nilai 0 akan berlaku apabila jarak yang sama antara medoid terdekat pertama dan kedua.
- 4) Nilai 1 dapat dicapai jika objeknya merupakan objek medoid.

Adapun contoh studi kasus untuk perhitungan manual validasi klaster menggunakan *Medoid Shadow Value* (MSV) disajikan pada **Lampiran 17**.

3.8. Pengujian Rata – Rata Dua Populasi Independen

Serangkaian tahapan dalam analisis pengelompokan yang sudah dilakukan, akan terbentuk hasil pengelompokan. Masing – masing variabel numerik pada setiap klaster akan dilakukan pengujian untuk dibandingkan antar klaster guna mengetahui rata – rata variabel secara independen berbeda signifikan.

3.8.1 Teorema Limit Pusat

Teorema limit pusat atau *Central Limit Theorem* merupakan suatu pernyataan yang mengatakan apabila suatu distribusi sampel dari nilai *mean* akan membentuk distribusi normal jika ukuran sampel semakin besar (Insani, 2020). Teorema ini menyatakan bahwa bentuk distribusi sampling dari *mean* sampel selalu cenderung mendekati distribusi normal.

Jika sampel acak berukuran n diambil dari suatu populasi dengan rata – rata μ serta varians σ^2 , maka nilai rata – rata sampel $\bar{X}$ akan memiliki distribusi yang mendekati distribusi normal dengan rata – rata μ dan variansi σ^2/n . Sehingga distribusi fungsinya adalah (Long, 2020):

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad (3.19)$$

3.8.2 Uji Rata – Rata Dua Populasi Independen

Uji rata-rata dua populasi independen adalah salah satu uji statistik parametrik yang digunakan dalam melakukan perbandingan dua kelompok independen untuk mengetahui apakah rata – rata populasi secara statistik signifikan berbeda. Terdapat beberapa kondisi yang perlu diperhatikan pada pengujian rata – rata dua populasi. Salah satu kondisinya yaitu ketika varians sama dan diketahui.

Misalkan dipunyai dua sampel acak independen masing – masing memiliki ukuran n_1 dan n_2 diambil dari populasi dengan rata – rata μ_1 dan μ_2 dan varians σ_1^2 dan σ_2^2 dengan peubah acak berdistribusi normal standar (Dani Hermawan, 2020):

$$Z = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (3.20)$$

dimana:

$\bar{X}_1$ = rata – rata sampel 1

$\bar{X}_2$ = rata – rata sampel 2

n_1 = ukuran sampel 1

n_2 = ukuran sampel 2

σ_1^2 = varians sampel 1

σ_2^2 = varians sampel 2

Diasumsikan bahwa n_1 dan n_2 cukup besar yang menjadikan Teorema Limit Pusat dalam hal ini dapat berlaku. Adapun pengujian hipotesisnya adalah sebagai berikut (Long, Tju Ji, 2020):

i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak ada perbedaan yang signifikan pada rata – rata kelompok 1 dan kelompok 2)

$H_1 : \mu_1 \neq \mu_2$ (Ada perbedaan yang signifikan pada rata – rata kelompok 1 dan kelompok 2)

ii. Daerah Kritis

Tolak H_0 jika $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}}$

iii. Statistik Uji

Perhitungan statistik uji untuk kondisi σ_1^2 dan σ_2^2 sama dan diketahui dapat menggunakan **Persamaan 3.20**.

iv. Keputusan

Jika nilai $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}}$ maka akan tolak H_0

Jika nilai $-Z_{\frac{\alpha}{2}} \leq Z_{hitung} \leq Z_{\frac{\alpha}{2}}$ maka gagal tolak H_0

v. Kesimpulan

Apabila tolak H_0 maka terdapat perbedaan yang signifikan pada rata – rata kelompok 1 dan kelompok 2

Apabila gagal tolak H_0 maka tidak terdapat perbedaan yang signifikan pada rata – rata kelompok 1 dan kelompok 2

BAB IV

METODOLOGI PENELITIAN

4.1. Populasi dan Sampel Penelitian

Adapun populasi dalam penelitian ini adalah data komponen kesiapan TIK untuk ANBK pada seluruh sekolah jenjang SMA di Provinsi Yogyakarta yakni sebanyak 169 sekolah. Sampel yang digunakan pada penelitian ini adalah data sekolah SMA yang telah melengkapi dan memperbaiki data komponen kesiapan TIK untuk ANBK serta bukan sekolah SMA dengan jenjang SPK (Satuan Pendidikan Kerjasama) SMA yakni sebanyak 165 sekolah.

4.2. Tempat dan Waktu Penelitian

Penelitian dilakukan mulai bulan November 2021 sampai Februari 2022 bertempat di Universitas Islam Indonesia. Proses pengambilan dan penginputan data sekunder dilakukan pada tanggal 16 November 2021.

4.3. Variabel Penelitian

Penelitian ini, menggunakan 13 variabel yang berkaitan dengan komponen kesiapan TIK sekolah dalam penyelenggaraan ANBK 2021 yang tersedia dalam website Sekolah Kita milik Kemendikbud. Terdiri dari 8 variabel tipe numerik serta 5 tipe kategorik. Adapun variabel tersebut dijelaskan pada **Tabel 4.1**.

Tabel 4.1 Variabel Penelitian

Variabel	Keterangan	Jenis Data	Kategori dan/atau Satuan
X_1	Guru	Numerik	Orang
X_2	Siswa	Numerik	Orang
X_3	Laboratorium Komputer	Numerik	Ruang
X_4	Komputer Utama / Proktor	Numerik	Unit
X_5	Komputer Sesuai Spesifikasi (Milik)	Numerik	Unit
X_6	Rata-Rata <i>Bandwidth</i>	Numerik	Mbps
X_7	<i>Switch Hub</i>	Numerik	Buah
X_8	WIFI	Numerik	Buah
X_9	Status Sekolah	Biner	0: Negeri 1: Swasta
X_{10}	Moda Pelaksanaan TIK	Biner	0: Online 1: Semi Online
X_{11}	Sumber Listrik	Biner	0: PLN 1: PLN dan Diesel

Variabel	Keterangan	Jenis Data	Kategori dan/atau Satuan
X_{12}	Akreditasi Sekolah	Kategorik (Ordinal)	0: Terakreditasi A 1: Terakreditasi B 2: Terakreditasi C 3: Tidak/ Belum Terakreditasi
X_{13}	Status Pelaksanaan TIK	Kategorik	0: Mandiri dan Ditumpangi 1: Mandiri 2: Menumpang

Tabel 4.2 menunjukkan struktur data yang akan digunakan pada penelitian ini:

Tabel 4.2 Struktur Data Penelitian

SMA ke- i	X_1	X_2	X_3	X_4	X_5	X_6	...	X_{13}
1	$x_{(1,1)}$	$x_{(2,1)}$	$x_{(3,1)}$	$x_{(4,1)}$	$x_{(5,1)}$	$x_{(6,1)}$	...	$x_{(13,1)}$
2	$x_{(1,2)}$	$x_{(2,2)}$	$x_{(3,2)}$	$x_{(4,2)}$	$x_{(5,2)}$	$x_{(6,2)}$	...	$x_{(13,2)}$
3	$x_{(1,3)}$	$x_{(2,3)}$	$x_{(3,3)}$	$x_{(4,3)}$	$x_{(5,3)}$	$x_{(6,3)}$	...	$x_{(13,3)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
165	$x_{(1,165)}$	$x_{(2,165)}$	$x_{(3,165)}$	$x_{(4,165)}$	$x_{(5,165)}$	$x_{(6,165)}$	...	$x_{(13,165)}$

Adapun definisi operasional dari masing – masing variabel yang ada pada **Tabel 4.1** dijelaskan dalam **Tabel 4.3**.

Tabel 4.3 Definisi Operasional Variabel

Variabel	Keterangan	Definisi Operasional Variabel
X_1	Guru	Banyaknya guru yang mengajar di suatu sekolah.
X_2	Siswa	Banyaknya siswa yang menempuh pendidikan di suatu sekolah.
X_3	Laboratorium Komputer	Banyaknya ruang laboratorium komputer di suatu sekolah.
X_4	Komputer Utama / Proktor	Banyaknya komputer utama atau proktor yang berfungsi sebagai server dalam pelaksanaan ANBK.
X_5	Komputer Sesuai Spesifikasi (Milik)	Banyaknya komputer yang dimiliki oleh suatu sekolah dan memenuhi spesifikasi untuk digunakan dalam pelaksanaan ANBK.
X_6	Rata – Rata Bandwidth	Rata-rata volume maksimum data untuk proses transmisi melalui koneksi internet pada saat proses <i>upload</i> dan <i>download</i> . $\text{Rata – rata Bandwidth} = \frac{\text{Bandwidth Upload} + \text{Bandwidth Download}}{2}$
X_7	Switch Hub	Banyaknya komponen jaringan komputer yang digunakan untuk menghubungkan komputer yang mempunyai kebutuhan pada <i>bandwidth</i> yang lumayan besar.

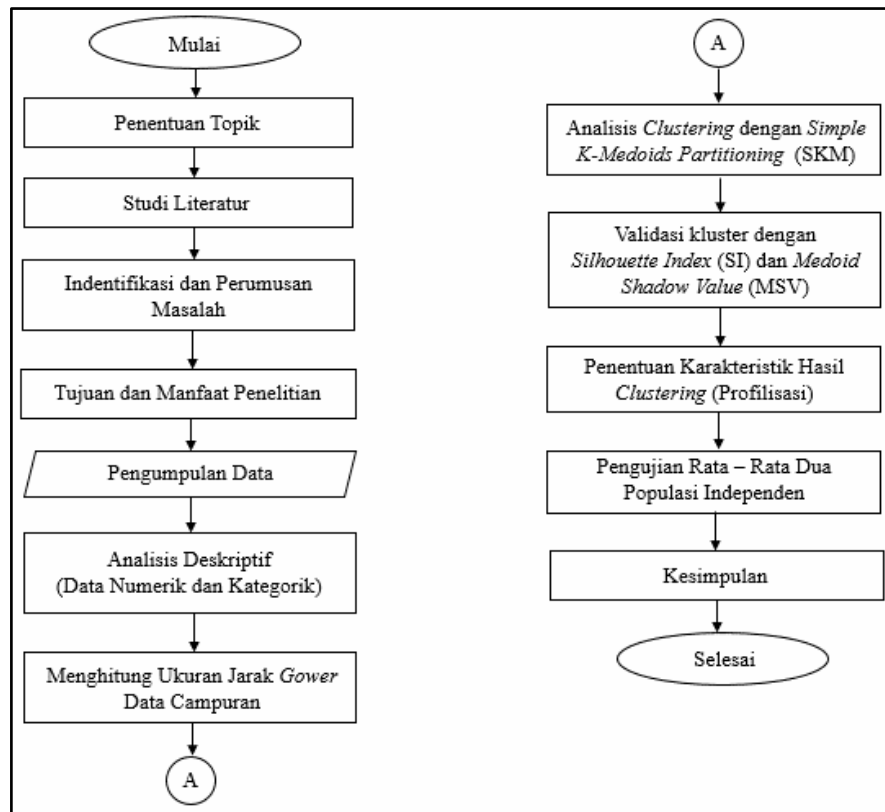
Variabel	Keterangan	Definisi Operasional Variabel
X_8	WIFI	Banyaknya jaringan nirkabel yang digunakan oleh komputer untuk mengakses internet.
X_9	Status Sekolah	Keberadaan serta kualitas dari suatu sekolah yang telah diakui oleh pemerintah.
X_{10}	Moda Pelaksanaan TIK	Bentuk atau jenis pelaksanaan ANBK. Terdapat dua moda pelaksanaan yaitu online dan semi online.
X_{11}	Sumber Listrik	Sumber aliran listrik yang digunakan oleh suatu sekolah yaitu PLN serta PLN dan diesel.
X_{12}	Akreditasi Sekolah	Peringkat kelayakan program dan/atau satuan pendidikan berdasarkan kriteria yang telah ditetapkan yaitu akreditasi A, B, C dan tidak/belum terakreditasi.
X_{13}	Status Pelaksanaan TIK	Kesediaan dan kualitas suatu sekolah untuk melaksanakan ANBK. Status pelaksanaan ini terbagi dalam 3 kategori yaitu mandiri, mandiri dan ditumpangi serta menumpang.

4.4. Alat dan Cara Organisir Data

Metode analisis yang digunakan dalam penelitian ini adalah metode pengelompokan non-hirarki yakni *Simple K-Medoids Partitioning* (SKM) dengan menggunakan alat bantu *software R Studio*, *Microsoft Excel* serta *Tableau* dan *Geoda* untuk proses visualisasi data. Proses pengumpulan data pada penelitian ini, yaitu berupa studi pustaka. Peneliti melakukan proses pengumpulan data dengan mencari informasi melalui dokumen yang mendukung dalam proses penulisan. Dokumen tersebut berupa buku, jurnal ilmiah, penelitian sebelumnya, berita, website terpercaya, skripsi dan disertasi. Selain itu, peneliti mengumpulkan data sekunder SMA di Provinsi Daerah Istimewa Yogyakarta yang ada dalam website Sekolah Kita milik Kementerian Pendidikan dan Kebudayaan dengan proses input satu persatu dari setiap objek ke dalam *Microsoft Excel*.

4.5. Tahapan Penelitian

Berikut tahapan dalam penelitian ini untuk mengelompokkan sekolah SMA menggunakan metode *Simple K-Medoids Partitioning* (SKM) yang digambarkan dalam *flowchart* pada **Gambar 4.1**.



Gambar 4.1 Diagram Alir Penelitian

Berdasarkan pada **Gambar 4.1** mengenai diagram alir penelitian ini dapat dijabarkan tahapan – tahapan sebagai berikut:

1. Penentuan Topik

Tahap awal penelitian ini dilakukan untuk menentukan topik yang akan diteliti. Dalam hal ini, peneliti menggunakan topik mengenai komponen kesiapan TIK untuk ANBK pada jenjang SMA di Provinsi Yogyakarta.

2. Studi Literatur

Studi literatur adalah suatu teknik untuk mengumpulkan referensi dan tinjauan pustaka yang dapat bersumber dari buku, jurnal, disertasi, skripsi dan artikel yang relevan dan sesuai dengan penelitian yang akan dilakukan.

Pada tahapan ini, penulis mengumpulkan data komponen kesiapan TIK dari website Kementerian Pendidikan dan Kebudayaan yaitu Sekolah Kita.

3. Identifikasi dan Perumusan Masalah

Tahap ini merupakan suatu permasalahan yang lebih spesifik mengenai ruang lingkup masalah yang akan diteliti. Rumusan masalah yang diangkat dalam penelitian ini berupa analisis deskriptif data, hasil pengelompokan dengan

algoritma *Simple K-Medoids Partitioning* (SKM), hasil validasi klaster, dan profilisasi hasil pengelompokan pada data komponen kesiapan TIK untuk ANBK pada jenjang SMA di Provinsi Yogyakarta tahun 2021.

4. Tujuan dan Manfaat Penelitian

Tujuan penelitian adalah rumusan kalimat yang memberikan adanya hasil berupa suatu hal yang didapatkan sesudah penelitian dilakukan serta hal yang ingin dicapai pada suatu penelitian. Adapun tujuan penelitian ini adalah menjawab masalah yang sudah dirumuskan sebelumnya.

Manfaat penelitian adalah hasil yang akan diperoleh dari suatu penelitian. Yang mana manfaat dari penelitian ini telah dituliskan pada Bab I.

5. Pengumpulan Data

Pada tahap ini, peneliti akan mengumpulkan data yang akan digunakan dalam penelitian. Data diambil dari website sekolah.data.kemdikbud.go.id. Peneliti akan menginputkan data sekolah satu – persatu ke dalam *Microsoft Excel*. Kemudian akan diolah menggunakan *R Studio* dan *Microsoft Excel*.

6. Analisis Deskriptif

Analisis deskriptif digunakan untuk menggambarkan karakteristik data yang akan digunakan dalam penelitian. Pada tahapan ini, peneliti akan menyajikan dalam grafik dan tabel statistik yang memuat ukuran statistik seperti nilai minimum, maksimum dan *mean* dari masing – masing variabel numerik serta menyajikan grafik berupa *pie chart* untuk variabel kategorik.

7. Perhitungan Ukuran Jarak *Gower* Data Campuran

Jarak antar objek untuk data campuran akan dihitung menggunakan jarak *Gower*. Perhitungan data numerik menggunakan rentang terbobot *Manhattan* dan untuk data kategorik menggunakan *simple matching*.

8. Analisis *Simple K-Medoids Partitioning* (SKM)

Algoritma *Simple K-Medoids Partitioning* (SKM) merupakan metode pengembangan dari algoritma *Simple and Fast K-Medoids* yang rentan pada permasalahan pada klaster kosong. Pada penelitian ini, algoritma SKM akan bekerja dengan data campuran yaitu numerik dan kategorik.

9. Validasi Klaster

Pada tahapan ini, hasil dari pengelompokan kemudian akan divalidasi dengan menggunakan *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV). Tujuannya untuk mengetahui jumlah k klaster yang optimal pada data komponen kesiapan TIK untuk ANBK jenjang SMA di Provinsi Yogyakarta.

10. Penentuan Karakteristik Hasil Pengelompokan (Profilisasi)

Setelahnya, akan dilakukan proses profilisasi yang bertujuan untuk mengetahui karakteristik variabel masing-masing klaster yang terbentuk. Sehingga peneliti dapat menentukan mana klaster yang tinggi dan rendah.

11. Pengujian Rata – Rata Dua Populasi Independen

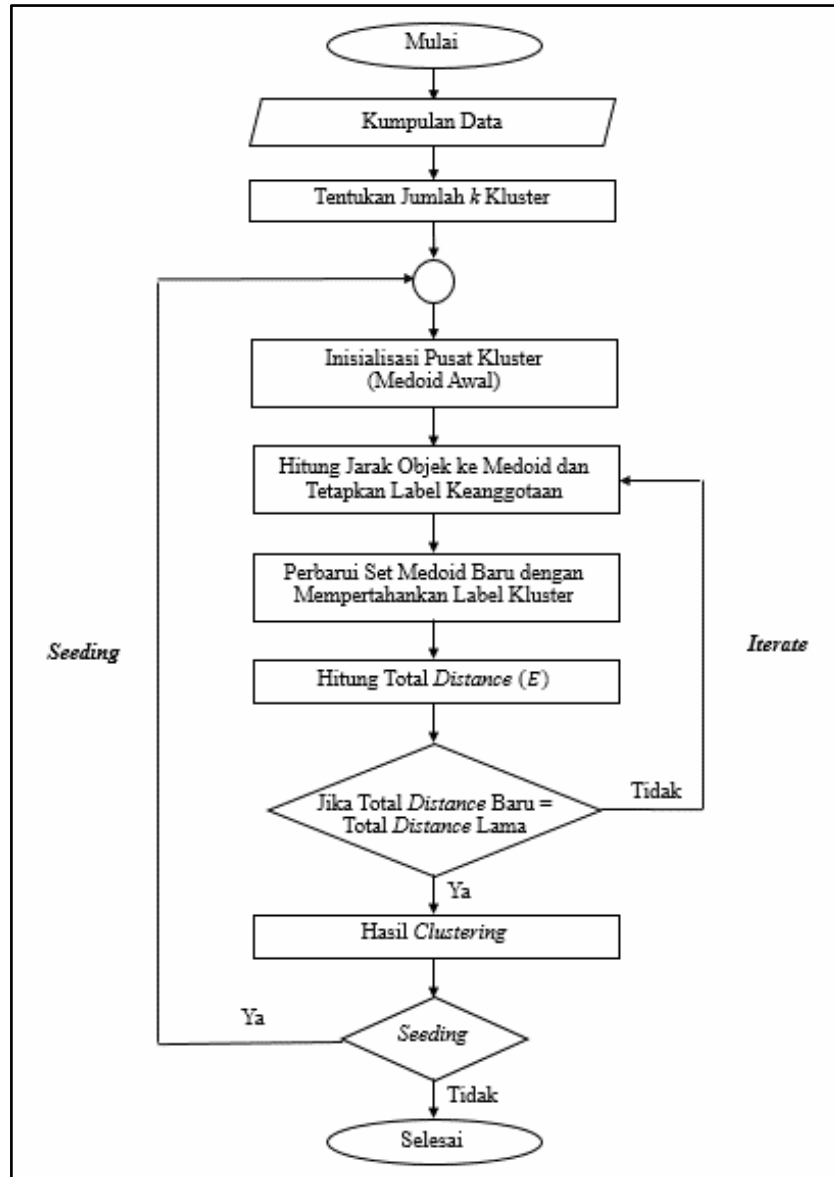
Hasil pengelompokan yang telah didapatkan selanjutnya akan dilakukan pengujian untuk membandingkan ada atau tidaknya perbedaan rata – rata secara signifikan dari variabel numerik pada masing – masing klaster yang terbentuk.

12. Kesimpulan

Tahapan terakhir pada penelitian ini adalah kesimpulan. Tahap ini akan memberikan ringkasan hasil yang dapat menjawab rumusan masalah yang diangkat pada penelitian ini.

4.5.1 Tahapan Analisis *Simple K-Medoids Partitioning*

Berikut tahapan dari algoritma *Simple K-Medoids Partitioning* yang digambarkan dalam *flowchart* pada **Gambar 4.2**.



Gambar 4.2 Diagram Alir Algoritma *Simple K-Medoids Partitioning*

Berdasarkan pada **Gambar 4.2** mengenai diagram alir algoritma SKM ini yang dapat dijabarkan tahapan – tahapannya sebagai berikut:

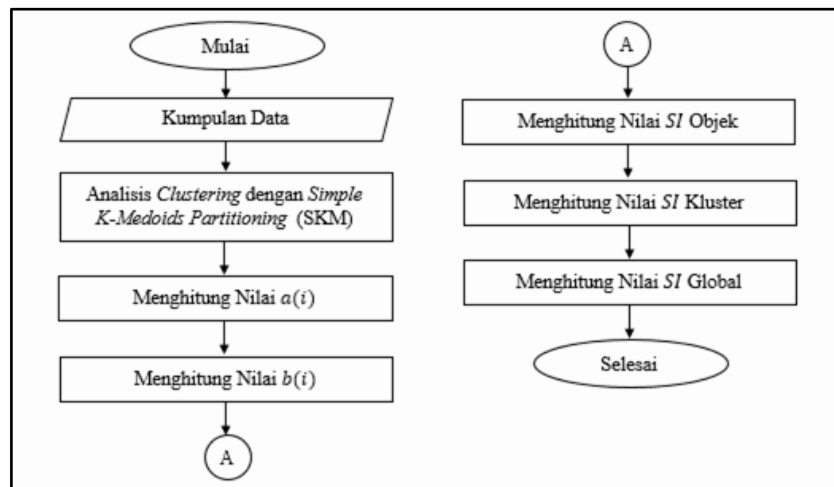
1. Inisialisasi pusat klaster atau medoid sejumlah k klaster. Proses ini dilakukan berdasarkan pada total jarak paling minimum antara seluruh objek.
2. Hitung jarak objek menuju masing – masing medoid dengan menggunakan rumus jarak *Gower* pada **Persamaan 3.10**. Setelah didapatkan jaraknya, kemudian menetapkan label keanggotaan dari masing – masing objek.
3. Perbarui medoid dengan tetap mempertahankan label klaster menggunakan rumus pada **Persamaan 3.11**. Medoid baru didefinisikan sebagai objek yang

meminimalkan jumlah jarak objek satu dengan lainnya yang ada dalam satu klaster.

4. Hitung total jarak (E) yang merupakan jarak antara objek dengan medoid terdekat menggunakan **Persamaan 3.12**.
5. Ulangi langkah 2 – 4, hingga total jarak baru sama dengan total jarak sebelumnya atau himpunan medoid tidak berubah atau jumlah iterasi yang ditentukan telah tercapai.
6. Ulangi langkah 1 – 5, untuk mendapatkan beberapa *seeding* secara acak dari medoid awal. Objek yang ada dalam pusat klaster akan selalu menjadi salah satu medoid awal. Penentuan medoid final dipilih berdasarkan nilai E yang paling minimum.
7. Tetapkan keanggotaan dari setiap objek ke medoid akhir terdekat untuk memperoleh hasil pengelompokan.

4.5.2 Tahapan Menghitung Nilai *Silhouette Index*

Berikut tahapan dalam evaluasi hasil pengelompokan menggunakan nilai *Silhouette Index* (SI) yang digambarkan dalam *flowchart* pada **Gambar 4.3**.



Gambar 4.3 Diagram Alir Perhitungan Nilai *Silhouette Index*

Berdasarkan pada **Gambar 4.3** mengenai diagram alir perhitungan nilai SI dapat dijabarkan tahapan – tahapan sebagai berikut:

1. Analisis Pengelompokan dengan *Simple K-Medoids Partitioning* (SKM)
Data yang digunakan dalam penelitian ini, akan dikelompokkan dengan menggunakan algoritma SKM. Pengelompokan data akan dibagi dalam dua klaster sesuai kedekatan di antara objek yang berada dalam satu klaster.

2. Menghitung Nilai $a(i)$

Nilai $a(i)$ merupakan rata – rata nilai jarak objek terhadap seluruh anggota lainnya yang berada dalam klaster yang sama dengan menggunakan rumus pada **Persamaan 3.13**.

3. Menghitung Nilai $b(i)$

Nilai $b(i)$ merupakan rata – rata nilai jarak objek terhadap seluruh anggota lainnya yang berada pada klaster yang berbeda dengan menggunakan rumus pada **Persamaan 3.14**.

4. Menghitung Nilai SI Objek

Nilai SI objek merupakan nilai *Silhouette Index* untuk objek ke- i yang berada dalam suatu klaster dengan menggunakan rumus pada **Persamaan 3.15**.

5. Menghitung Nilai SI Klaster

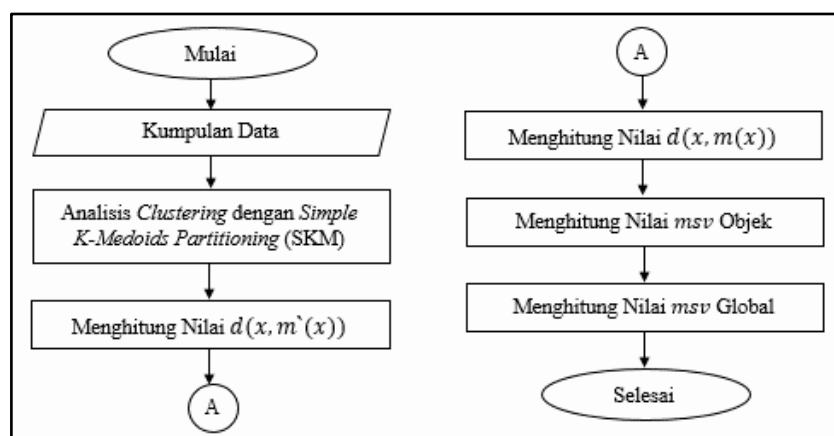
Nilai SI klaster merupakan rata – rata nilai *Silhouette Index* dari seluruh objek ke- i yang berada dalam satu klaster yang sama dengan menggunakan rumus pada **Persamaan 3.16**.

6. Menghitung Nilai SI Global

Nilai SI global merupakan rata – rata nilai *Silhouette Index* dari seluruh SI klaster dengan menggunakan rumus pada **Persamaan 3.17**.

4.5.3 Tahapan Menghitung Nilai *Medoid Shadow Value* (MSV)

Berikut tahapan evaluasi hasil pengelompokan menggunakan nilai *Medoid Shadow Value* (MSV) yang digambarkan dalam *flowchart* pada **Gambar 4.4**.



Gambar 4.4 Diagram Alir Perhitungan Nilai *Medoid Shadow Value*

Berdasarkan pada **Gambar 4.4** mengenai diagram alir perhitungan nilai MSV dapat dijabarkan tahapan – tahapan sebagai berikut:

1. Analisis Pengelompokan dengan *Simple K-Medoids Partitioning* (SKM)

Data yang digunakan dalam penelitian ini dikelompokkan menggunakan algoritma SKM. Pengelompokan data akan dibagi dalam dua kelompok sesuai dengan kedekatan di antara objek yang berada dalam satu klaster.

2. Menghitung Nilai $d(x, m^{\prime}(x))$

Nilai $d(x, m^{\prime}(x))$ adalah jarak antara objek x ke medoid terdekat kedua.

3. Menghitung Nilai $d(x, m(x))$

Nilai $d(x, m(x))$ adalah jarak antara objek x ke medoid terdekat pertama.

4. Menghitung Nilai MSV

Nilai msv merupakan nilai bayangan berbasis medoid yang digunakan dalam proses validasi hasil pengelompokan menggunakan **Persamaan 3.18**.

5. Menghitung Nilai MSV Global

Nilai msv global merupakan rata – rata nilai MSV dari suatu k klaster tertentu.

BAB V

HASIL DAN PEMBAHASAN

5.1. Gambaran Data Penelitian

Gambaran data penelitian yang diambil dari website Sekolah Kita milik Kemendikbud yang akan digunakan dalam penelitian ini disajikan pada **Tabel 5.1**.

Tabel 5.1 Gambaran Data Penelitian

No	Kode SMA	Guru	Siswa	Laboratorium Komputer	Komputer Utama	Komputer Sesuai Spesifikasi (Milik)
1	S1	42	608	3	6	80
2	S2	47	215	3	3	64
3	S3	45	609	2	4	85
⋮	⋮	⋮	⋮	⋮	⋮	⋮
165	S165	11	57	1	0	9

No	Kode SMA	Rata-Rata <i>Bandwidth</i>	<i>Switch Hub</i>	WIFI	Status Sekolah
1	S1	100	3	6	Negeri
2	S2	393	3	7	Negeri
3	S3	20	5	25	Negeri
⋮	⋮	⋮	⋮	⋮	⋮
165	S165	50	24	2	Swasta

No	Kode SMA	Moda Pelaksanaan TIK	Sumber Listrik	Akreditasi	Status Pelaksanaan TIK
1	S1	Semi Online	PLN	A	Mandiri dan Ditumpangi
2	S1	Semi Online	PLN	A	Mandiri dan Ditumpangi
3	S3	Semi Online	PLN	A	Mandiri dan Ditumpangi
⋮	⋮	⋮	⋮	⋮	⋮
165	S165	Semi Online	PLN	B	Mandiri

Berdasarkan **Tabel 5.1** didapatkan informasi mengenai struktur data yang akan dianalisis menggunakan metode *Simple K-Medoids Partitioning* (SKM).

5.2. Statistika Deskriptif

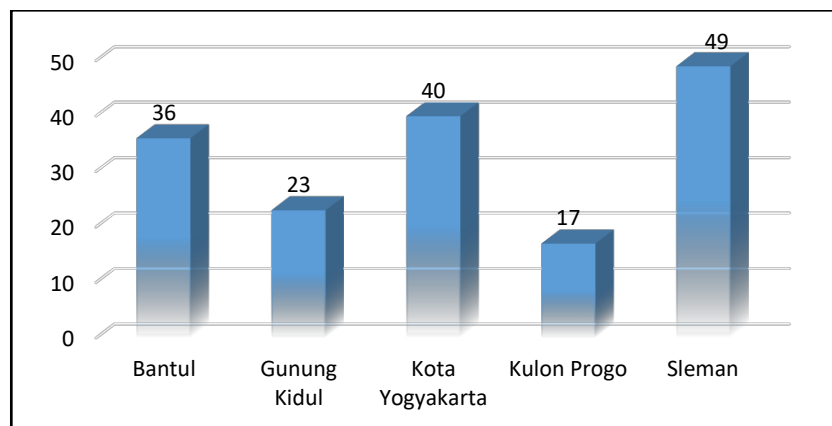
Analisis deskriptif sangat penting dalam penelitian guna mengetahui gambaran umum data yang dianalisis. Pada penelitian ini, peneliti menggunakan dua tipe data yaitu data numerik dan data kategorik. Jumlah variabel pada data numerik sebanyak 8 variabel dan data kategorik ada 5 variabel.

Data numerik pada data komponen kesiapan TIK sekolah dalam penyelenggaraan ANBK terdiri dari 8 variabel. Statistika deskriptif untuk data numerik disajikan dalam bentuk ukuran statistik yang berupa nilai *Mean* (atau nilai rata-rata), nilai *Min.* (atau nilai terendah) dan nilai *Max.* (atau nilai tertinggi) yang disajikan dalam **Tabel 5.2**.

Tabel 5.2 Statistika Deskriptif

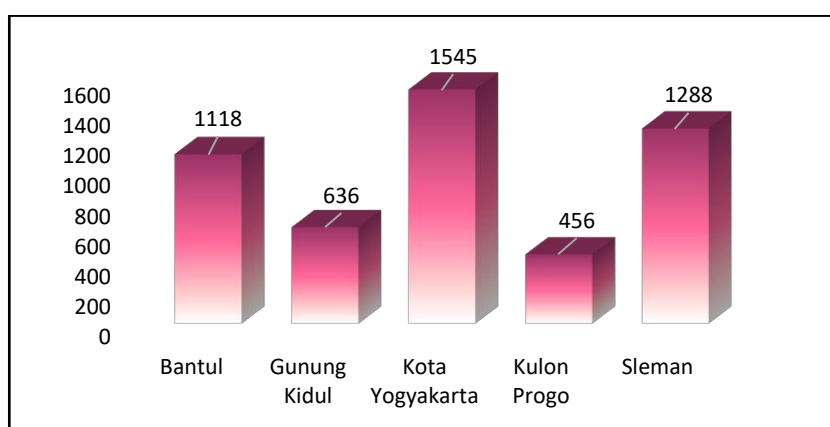
Variabel	Keterangan	Mean	Min.	Max.
X_1	Guru (Orang)	31	3	199
X_2	Siswa (Orang)	325	0	1157
X_3	Laboratorium Komputer (Ruang)	3	0	60
X_4	Komputer Utama / Proktor (Unit)	3	0	50
X_5	Komputer Sesuai Spesifikasi (Milik) (Unit)	51	0	171
X_6	Rata-rata <i>Bandwidth</i> (Mbps)	46.38	0	566
X_7	<i>Switch Hub</i> (Buah)	6	0	85
X_8	WIFI (Buah)	9	0	126

Rata-rata jumlah guru untuk masing-masing sekolah SMA di Provinsi Yogyakarta sebesar 31 orang. Rata-rata jumlah siswa yang menempuh pendidikan pada sekolah SMA di Provinsi Yogyakarta sebanyak 325 orang. Fasilitas penunjang pembelajaran seperti laboratorium komputer rata-rata pada masing-masing sekolah SMA ada sebanyak 3 ruang yang dilengkapi dengan komputer utama sebanyak 3 ruang dan komputer sesuai spesifikasi milik sebanyak 51 unit. Besarnya rata-rata *bandwidth* yang digunakan dalam proses penyelenggaraan ANBK pada masing-masing sekolah SMA sebesar 46.38 Mbps. Setiap sekolah SMA memiliki fasilitas *switch hub* sebanyak 6 buah dan WIFI sebanyak 9 buah.



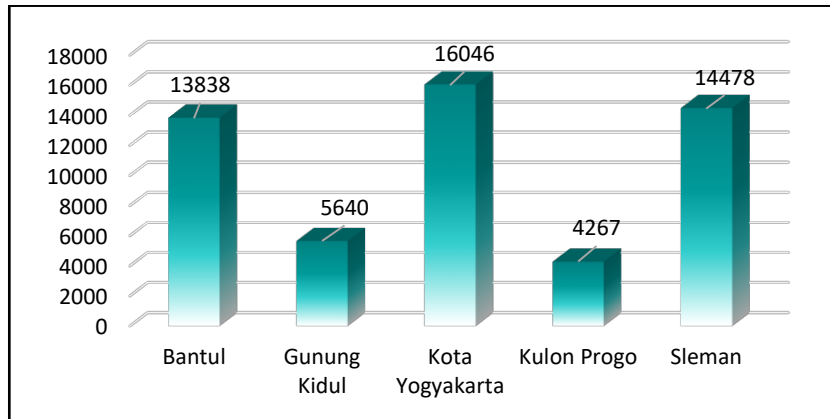
Gambar 5.1 Diagram Batang Jumlah Infrastruktur Sekolah SMA

Gambar 5.1 menggambarkan jumlah sekolah SMA yang ada di Provinsi Yogyakarta pada setiap kabupaten/kota tahun 2021. Kabupaten Sleman memiliki fasilitas infrastruktur sekolah SMA paling banyak di antara kabupaten/kota lainnya yaitu sebanyak 49 infrastruktur. Berbanding terbalik dengan Kabupaten Kulon Progo yang merupakan kabupaten dengan fasilitas infrastruktur SMA paling sedikit yaitu 17 infrastruktur. Rata-rata infrastruktur sekolah SMA di Provinsi Yogyakarta adalah 33 infrastruktur. Hal tersebut mengartikan bahwa terdapat 2 kabupaten/kota yang memiliki jumlah infrastruktur sekolah SMA yang berada dibawah rata-rata Provinsi.



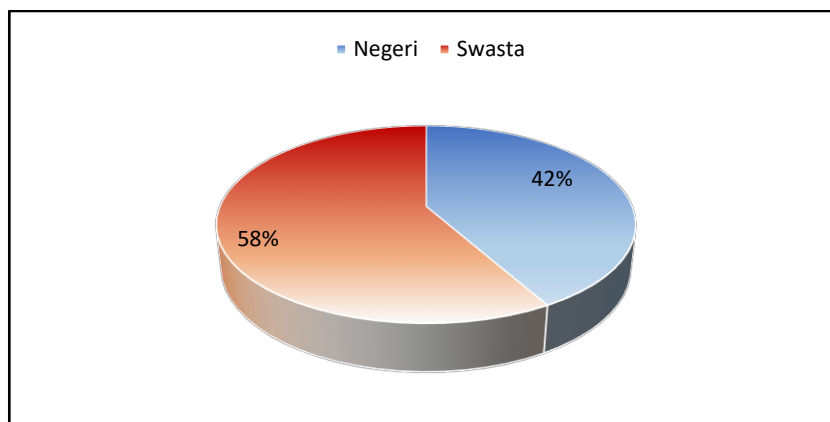
Gambar 5.2 Diagram Batang Jumlah Guru Sekolah SMA

Gambar 5.2 merupakan diagram batang untuk persebaran jumlah guru sekolah SMA berdasarkan kabupaten/kota di Provinsi Yogyakarta. Total keseluruhan guru sekolah SMA di Provinsi Yogyakarta ada sebanyak 5043 guru. Kota Yogyakarta menjadi kabupaten/kota dengan jumlah guru SMA terbanyak di Provinsi Yogyakarta yakni 1545 guru. Berkebalikan dengan Kabupaten Kulon Progo yang hanya memiliki jumlah guru SMA sebanyak 456 guru, nilai ini sangat berbeda jauh jika dibandingkan dengan Kota Yogyakarta. Sedangkan untuk jumlah guru sekolah SMA di Kabupaten Bantul, Kabupaten Gunung Kidul dan Kabupaten Sleman secara berturut-turut adalah 1118 guru, 636 guru dan 1288 guru.



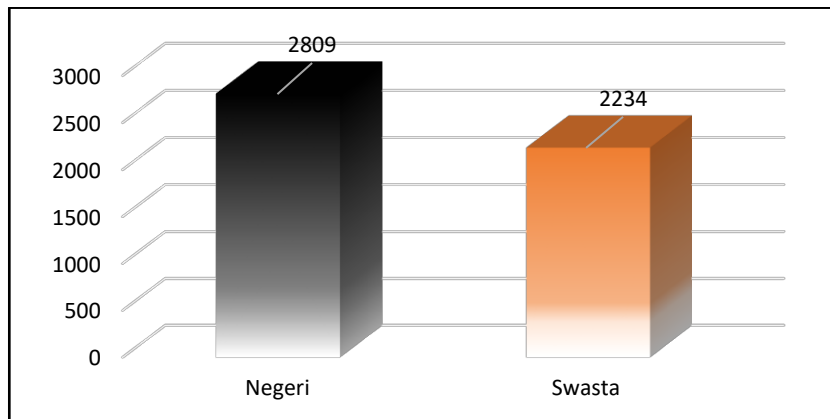
Gambar 5.3 Diagram Batang Jumlah Siswa Sekolah SMA

Gambar 5.3 menggambarkan jumlah siswa sekolah SMA berdasarkan kabupaten/kota di Provinsi Yogyakarta. Total keseluruhan siswa yang menempuh jenjang pendidikan SMA sebanyak 54269 siswa. Terdapat tiga wilayah kabupaten/kota yang memiliki jumlah siswa mencapai lebih dari 10000 siswa. Ketiga wilayah tersebut adalah Kabupaten Bantul, Kota Yogyakarta dan Kabupaten Sleman. Sedangkan untuk Kabupaten Gunung Kidul dan Kulon Progo masing-masing memiliki jumlah siswa SMA sebanyak 5640 siswa dan 4267 siswa.



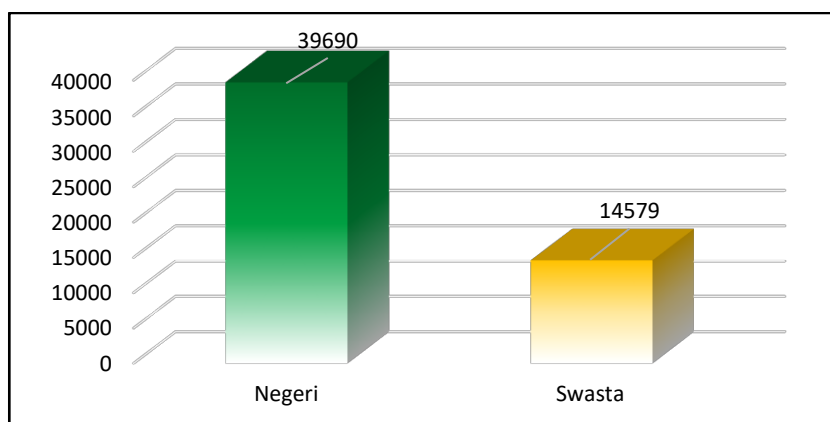
Gambar 5.4 Persentase Sekolah SMA Menurut Status Sekolah

Gambar 5.4 menyajikan diagram lingkaran persentase jumlah sekolah SMA di Provinsi Yogyakarta berdasarkan status sekolahnya yaitu negeri dan swasta. Persentase sekolah SMA swasta memiliki proporsi yang lebih besar yaitu sebesar 58% atau sebanyak 96 sekolah. Jumlah sekolah SMA negeri di Provinsi Yogyakarta mempunyai persentase sebesar 42% atau sebanyak 69 sekolah .



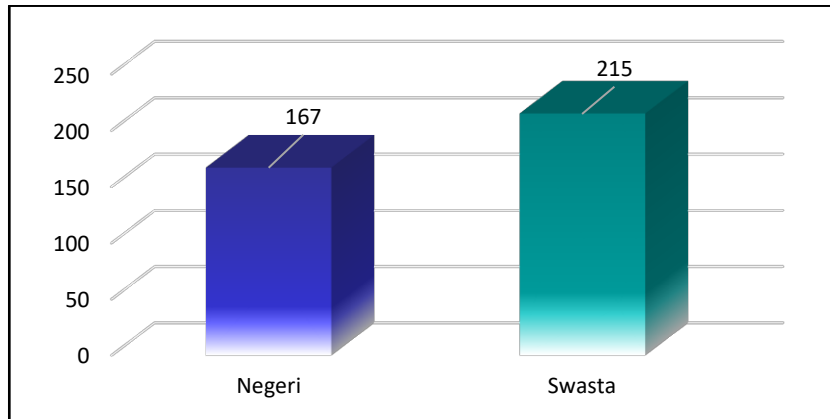
Gambar 5.5 Diagram Batang Guru Sekolah SMA

Gambar 5.5 menggambarkan jumlah guru yang mengajar pada sekolah SMA yang ada di Provinsi Yogyakarta berdasarkan status sekolah tahun 2021. Jumlah guru SMA negeri di Provinsi Yogyakarta sebanyak 2809 guru dan untuk sekolah SMA swasta sebanyak 2234 guru. Jika dilihat pada jumlah infrastruktur sekolah SMA negeri di Provinsi Yogyakarta, rata-rata sekolah SMA negeri mempunyai jumlah guru sebanyak 4115 guru. Pada sekolah SMA swasta, rata-rata setiap sekolah memiliki jumlah guru sebanyak 24 guru.



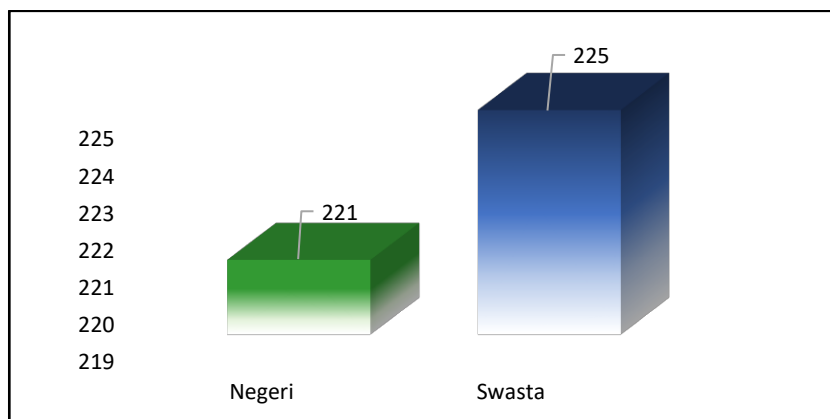
Gambar 5.6 Diagram Batang Siswa Sekolah SMA

Gambar 5.6 menunjukkan jumlah siswa yang menempuh jenjang SMA di Provinsi Yogyakarta menurut status sekolah. Sekolah SMA negeri memiliki jumlah siswa sebanyak 39690 siswa. Angka tersebut memiliki jumlah hampir tiga kali lipat dari jumlah siswa sekolah SMA swasta yang hanya sebanyak 14579 siswa. Sehingga dapat dikatakan bahwa siswa sekolah SMA di Provinsi Yogyakarta lebih didominasi oleh siswa yang menempuh pendidikan pada sekolah negeri. Kemudian, jumlah rata-rata siswa untuk setiap sekolah SMA swasta sebanyak 152 siswa dan untuk sekolah SMA negeri memiliki jumlah rata-rata siswa sebanyak 576 siswa.



Gambar 5.7 Diagram Batang Laboratorium Komputer Sekolah SMA

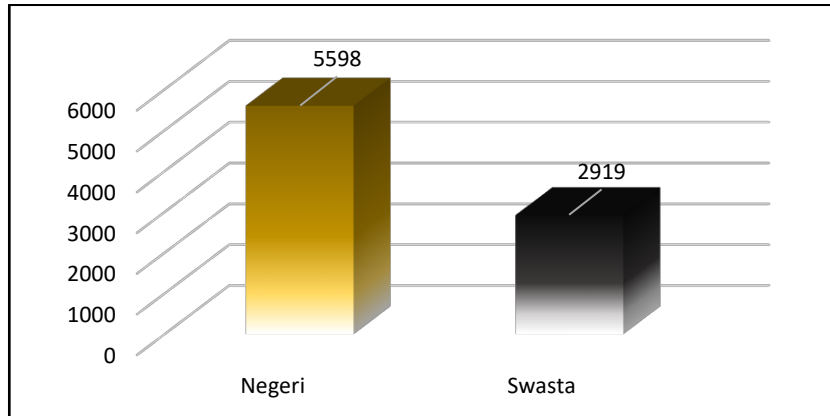
Gambar 5.7 merupakan diagram batang jumlah fasilitas laboratorium komputer yang tersedia pada sekolah SMA di Provinsi Yogyakarta menurut status sekolah. Fasilitas laboratorium komputer pada sekolah SMA swasta jumlahnya lebih banyak dibandingkan pada sekolah SMA negeri. Pada sekolah SMA swasta tersedia sebanyak 215 ruang laboratorium komputer, yang artinya setiap sekolah SMA memiliki jumlah laboratorium komputer sebanyak 2-3 ruang. Sedangkan pada sekolah SMA negeri hanya tersedia 167 ruang laboratorium komputer, di mana rata-rata sekolah SMA negeri mempunyai laboratorium komputer sebanyak 2-3 ruang. Apabila dilihat dari kedua rata-rata ruang laboratorium komputer pada masing-masing sekolah SMA negeri dan swasta, kedua memiliki jumlah yang sama yaitu sebanyak 2-3 ruang.



Gambar 5.8 Diagram Batang Komputer Utama Sekolah SMA

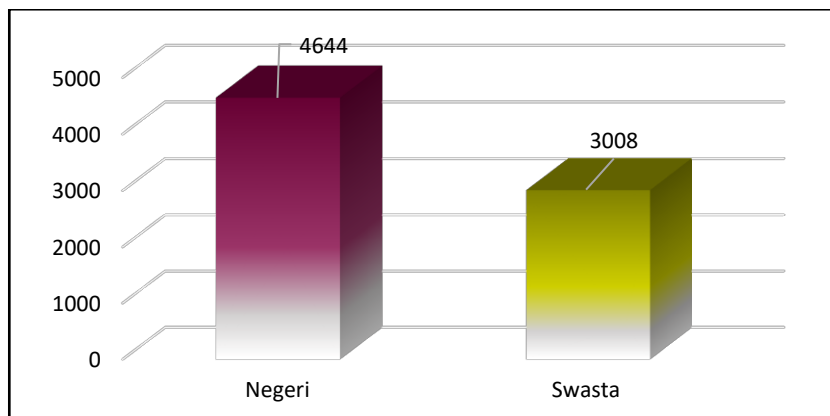
Gambar 5.8 menunjukkan banyaknya fasilitas komputer utama atau proktor yang tersedia dalam sekolah SMA di Provinsi Yogyakarta menurut status sekolah. Pada sekolah SMA swasta memiliki jumlah proktor sebanyak 225 unit sehingga setiap sekolah SMA swasta di Provinsi Yogyakarta rata-rata mempunyai proktor

sebanyak 3 unit. Kemudian untuk sekolah SMA negeri di Provinsi Yogyakarta mempunyai total jumlah proktor sebanyak 221 unit dari 69 sekolah yang ada, maka pada masing-masing sekolah SMA negeri rata-rata memiliki proktor sebanyak 4 unit. Komputer utama ini berfungsi sebagai server yang akan memonitor dalam proses penyelenggaraan ANBK.



Gambar 5.9 Diagram Batang Komputer Spesifikasi (Milik) Sekolah SMA

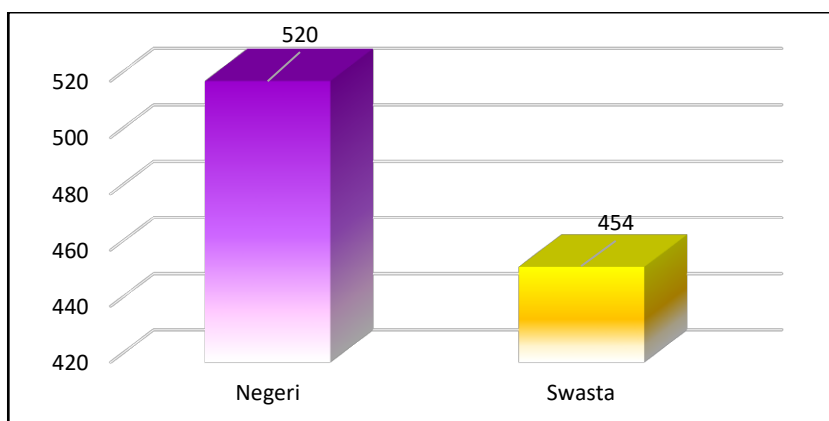
Gambar 5.9 menyajikan diagram batang untuk jumlah komputer sesuai spesifikasi milik yang tersedia pada sekolah SMA di Provinsi Yogyakarta menurut status sekolah. Total komputer sesuai spesifikasi milik pada 69 sekolah SMA negeri di Provinsi Yogyakarta adalah 5598 unit. Setiap sekolah SMA negeri rata-rata tersedia komputer sesuai spesifikasi milik sebanyak 81 unit. Pada sekolah SMA swasta, masing-masing sekolah memiliki rata-rata jumlah komputer sebanyak 31 unit sehingga total seluruh komputer sesuai spesifikasi milik yang tersedia pada sekolah SMA swasta di Provinsi Yogyakarta adalah 2919 unit.



Gambar 5.10 Diagram Batang Rata-Rata *Bandwidth* Sekolah SMA

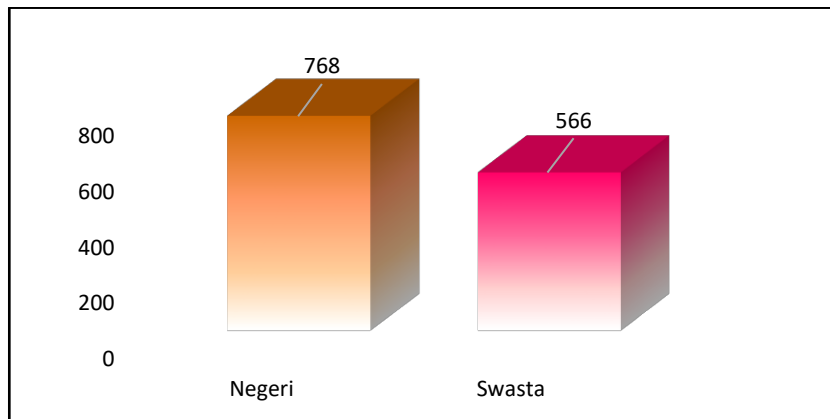
Gambar 5.10 menggambarkan jumlah rata-rata *bandwidth* yang merupakan total keseluruhan penggunaan dari seluruh sekolah SMA di Provinsi Yogyakarta

yaitu 165 sekolah. Sekolah SMA negeri memiliki jumlah rata-rata *bandwidth* sebesar 4644 Mbps yang merupakan gabungan dari 69 sekolah sehingga untuk setiap sekolah memiliki rata-rata nilai *bandwidth* sebesar 67.30 Mbps. Pada 96 sekolah SMA swasta mempunyai total rata-rata *bandwidth* sebesar 3008 Mbps, maka untuk masing-masing sekolah menggunakan rata-rata nilai *bandwidth* sebesar 31.33 Mbps. Nilai rata-rata *bandwidth* ini merupakan hasil dari rata-rata antara nilai *bandwidth upload* dan *bandwidth download*.



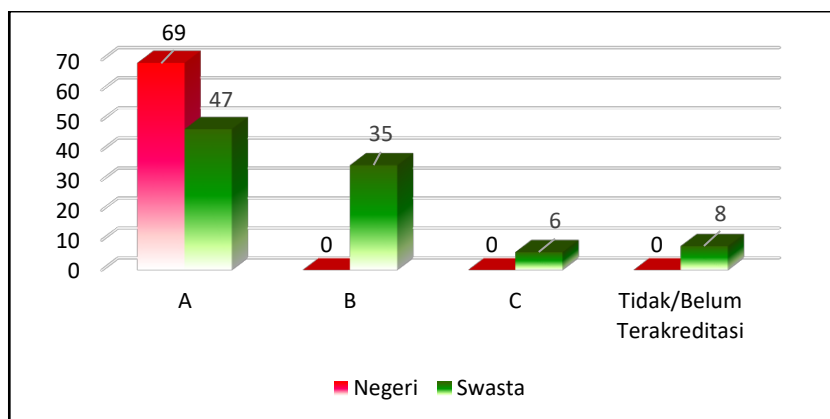
Gambar 5.11 Diagram Batang *Swicth Hub* Sekolah SMA

Gambar 5.11 menunjukkan jumlah *switch hub* yang digunakan oleh sekolah SMA di Provinsi Yogyakarta menurut status sekolah. Total *switch hub* yang digunakan oleh 165 sekolah SMA di Provinsi Yogyakarta yaitu 974 buah, di mana untuk sekolah SMA negeri sebanyak 520 buah dan sekolah SMA swasta sebanyak 454 buah. Setiap sekolah SMA negeri memiliki rata-rata jumlah *switch hub* yaitu 8 buah, sedangkan pada sekolah SMA swasta memiliki rata-rata jumlah *switch hub* yakni 5 buah untuk masing-masing sekolah. Adanya *switch hub* menjadi hal yang penting, karena alat ini berfungsi sebagai komponen jaringan komputer yang digunakan untuk menghubungkan komputer yang mempunyai kebutuhan pada *bandwidth* yang besar.



Gambar 5.12 Diagram Batang WIFI Sekolah SMA

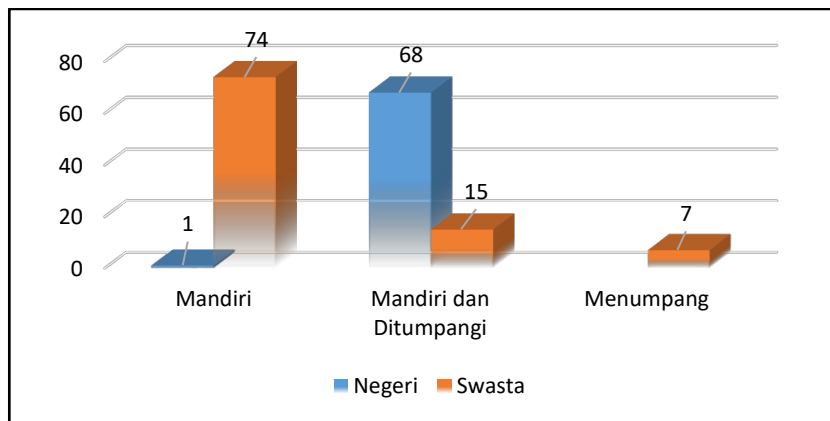
Gambar 5.12 adalah diagram batang dari jumlah keseluruhan WIFI yang dimiliki oleh sekolah SMA di Provinsi Yogyakarta berdasarkan status sekolah. Pada 69 sekolah SMA negeri memiliki total penggunaan WIFI sebanyak 768 buah, sehingga untuk setiap sekolah rata-rata menggunakan WIFI sebanyak 12 buah. Kemudian, pada 96 sekolah SMA swasta mempunyai total penggunaan WIFI sebanyak 566 buah, sehingga untuk setiap sekolah rata-rata menggunakan WIFI sebanyak 6 buah. Adanya WIFI menjadi penghubung bagi komputer untuk mengakses internet ketika pelaksanaan ANBK.



Gambar 5.13 Diagram Batang Akreditasi Sekolah SMA

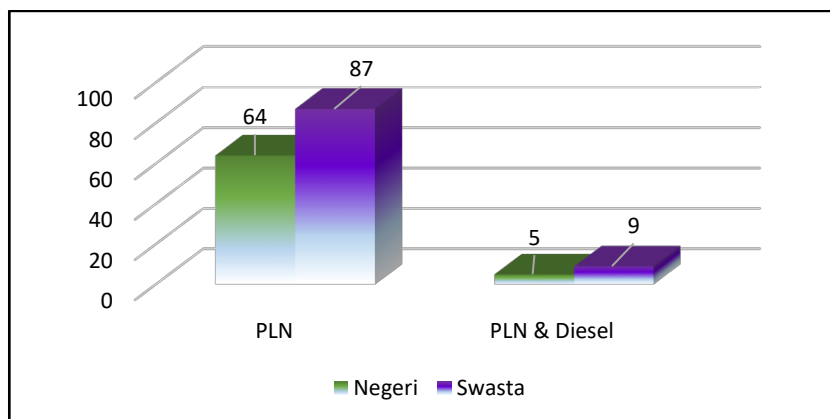
Gambar 5.13 menunjukkan jumlah kategori akreditasi pada sekolah SMA di Provinsi Yogyakarta menurut status sekolah yaitu negeri dan swasta. Seluruh sekolah SMA negeri di Provinsi Yogyakarta telah terakreditasi yaitu sebanyak 69 sekolah. Sedangkan untuk sekolah SMA swasta sudah banyak yang telah terakreditasi A dan B masing-masing sebanyak 47 sekolah dan 35 sekolah. Sekolah SMA swasta di Provinsi Yogyakarta yang masih memiliki akreditasi rendah yaitu C atau bahkan tidak/belum terakreditasi jumlahnya ada 6 sekolah dan 8 sekolah.

Persebaran sekolah SMA swasta yang tidak/belum terakreditasi masih merata untuk seluruh kabupaten/kota di Provinsi Yogyakarta. Kabupaten Bantul, Kabupaten Kulon Progo dan Kota Yogyakarta masing-masing terdapat 1 sekolah SMA swasta yang belum/tidak terakreditasi. Sisanya 5 sekolah SMA yang tidak/belum terakreditasi berada di Kabupaten Sleman dengan 3 sekolah dan Kabupaten Gunung Kidul dengan 2 sekolah.



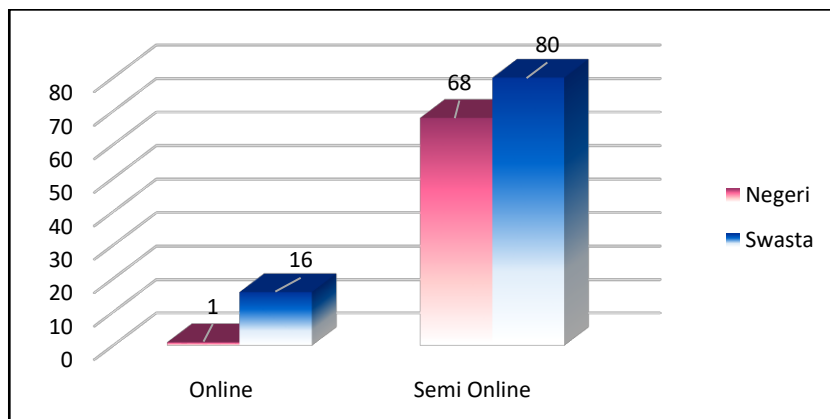
Gambar 5.14 Diagram Batang Status Pelaksanaan TIK Sekolah SMA

Gambar 5.14 menyajikan jumlah kategori status pelaksanaan TIK yang diikuti oleh 165 sekolah SMA di Provinsi Yogyakarta menurut status sekolah. Status pelaksanaan TIK yang paling mendominasi dari sekolah SMA negeri adalah mandiri dan ditumpangi dengan total sebanyak 68 sekolah. Pada sekolah SMA swasta lebih didominasi oleh sekolah dengan status pelaksanaan TIK berupa mandiri yakni sebanyak 74 sekolah. Kemudian, status pelaksanaan TIK berupa menumpang hanya ada pada sekolah SMA swasta dengan total 7 sekolah. Hanya terdapat 1 sekolah SMA negeri yang memiliki status pelaksanaan TIK berupa mandiri, sekolah tersebut adalah SMA Negeri 2 Wates yang terletak di Kabupaten Kulon Progo.



Gambar 5.15 Diagram Batang Sumber Listrik Sekolah SMA

Gambar 5.15 merupakan diagram batang jumlah kategori sumber listrik yang digunakan oleh 165 sekolah SMA di Provinsi Yogyakarta menurut status sekolah. Seluruh sekolah SMA baik negeri maupun swasta lebih mendominasi penggunaan listrik yang berasal dari PLN yaitu sebanyak 151 sekolah masing-masing untuk sekolah negeri ada 64 sekolah dan swasta ada 87 sekolah. Sisanya yaitu 14 sekolah menggunakan listrik yang berasal dari dua sumber yaitu PLN dan diesel di mana terdapat 5 sekolah SMA negeri dan 9 sekolah SMA swasta.



Gambar 5.16 Diagram Batang Moda Pelaksanaan TIK Sekolah SMA

Gambar 5.16 adalah diagram batang tentang jumlah kategori moda pelaksanaan TIK yang diikuti oleh sekolah SMA di Provinsi Yogyakarta berdasarkan status sekolah. Terdapat 1 sekolah SMA negeri yang mengikuti pelaksanaan TIK dengan moda online, sekolah tersebut adalah SMA Negeri 1 Imogiri. Kemudian untuk sisanya sebanyak 68 sekolah SMA negeri mengikuti moda pelaksanaan TIK berupa semi online. Pada sekolah SMA swasta sebesar 83.3% mengikuti moda pelaksanaan TIK berupa semi online dan 16 sekolah lainnya mengikuti moda pelaksanaan TIK berupa online.

5.3. Pengukuran Kedekatan Antar Objek

Dalam melakukan analisis pengelompokan, mengukur kedekatan antar objek menjadi hal yang terpenting. Karena pada dasarnya, dalam analisis pengelompokan merupakan metode untuk mengklasifikasikan objek di mana dalam satu kluster terdiri dari objek yang memiliki kedekatan yang tinggi dan dengan kluster yang lain mempunyai kedekatan yang rendah.

Pada proses pengukuran kedekatan antar objek akan menggunakan jarak *Gower* yang ada pada **Persamaan 3.7**. Pengukuran jarak digunakan secara bersama untuk variabel numerik, biner dan kategorik. Dimana dalam mengukur kedekatan variabel numerik menggunakan rentang terbobot *Manhattan* sedangkan untuk variabel biner dan kategorik menggunakan *simple matching*.

1. Pengukuran Kedekatan Objek Variabel Numerik

Proses pengukuran kedekatan objek dengan variabel numerik maka akan menggunakan **Persamaan 3.1**. Perhitungan dilakukan untuk seluruh data dengan jenis numerik yaitu jumlah guru, siswa, laboratorium komputer, komputer utama, komputer sesuai spesifikasi (milik), rata-rata *bandwidth*, *switch hub* dan WIFI. Misalnya ukuran kedekatan numerik dengan rentang terbobot *manhattan* pada objek S1 dengan S2 dan objek S1 dengan S3 adalah sebagai berikut.

$$s_{(S_1, S_2)} = \left(1 - \frac{|42-47|}{(199-3)}\right) + \left(1 - \frac{|608-215|}{(1157-0)}\right) + \left(1 - \frac{|3-3|}{(60-0)}\right) + \dots + \left(1 - \frac{|6-7|}{(126-0)}\right) \\ = 6.96$$

$$s_{(S_1, S_3)} = \left(1 - \frac{|42-45|}{(199-3)}\right) + \left(1 - \frac{|608-609|}{(1157-0)}\right) + \left(1 - \frac{|3-2|}{(60-0)}\right) + \dots + \left(1 - \frac{|6-25|}{(126-0)}\right) \\ = 7.58$$

Perhitungan dilakukan untuk seluruh objek data. Hasil perhitungan total jarak variabel numerik untuk 165 objek disajikan pada matriks $d_{numerik}$ ordo 165×165 pada **Tabel 5.3**.

Tabel 5.3 Matriks Kedekatan Objek Data Numerik

	S1	S2	S3	S4	S5	S6	S7	S8	...	S165
S1	8	6.96	7.58	7.42	7.49	7.25	7.31	7.01	...	6.43
S2	6.96	8	6.67	6.50	7.04	6.36	6.69	6.82	...	6.37
S3	7.58	6.67	8	7.26	7.36	7.32	7.45	6.89	...	6.35
S4	7.42	6.50	7.26	8	7.14	7.60	6.90	7.08	...	6.08
S5	7.49	7.04	7.36	7.14	8	6.99	7.44	6.60	...	6.94
S6	7.25	6.36	7.32	7.60	6.99	8	6.88	7.27	...	6.06
S7	7.31	6.69	7.45	6.90	7.44	6.88	8	6.53	...	6.90
S8	7.01	6.82	6.89	7.08	6.60	7.27	6.53	8	...	5.62
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
S165	6.43	6.37	6.35	6.08	6.94	6.06	6.90	5.62	...	8

Berdasarkan **Tabel 5.3** diketahui seberapa besar kedekatan antara objek satu dengan yang lain. Misalkan kedekatan objek data numerik untuk S1 (SMA Negeri

1 Pundong) dengan S4 (SMA Negeri 1 Sedayu) adalah sebesar 7.42. Nilai maksimum kedekatan numerik adalah 8 sesuai jumlah variabel numerik.

2. Pengukuran Kedekatan Objek Variabel Biner dan Kategorik

Proses pengukuran kedekatan objek dengan variabel biner dan kategorik maka akan menggunakan **Persamaan 3.2**. Perhitungan dilakukan untuk seluruh data dengan jenis biner dan kategorik yaitu status sekolah, moda pelaksanaan TIK, sumber listrik, akreditasi sekolah dan status pelaksanaan TIK. Misalnya untuk ukuran kedekatan biner dan kategorik dengan *simple matching* pada objek S1 dengan S2 serta objek S1 dengan S3 adalah

$$s_{(S_1, S_2)} = (1 + 1 + 1 + 1 + 1) = 5$$

$$s_{(S_1, S_2)} = (1 + 1 + 1 + 1 + 1) = 5$$

Bernilai 1 jika antar objek memiliki kategori variabel yang sama dan 0 jika kategori variabel antar objek berbeda. Pada objek S1 dan S2, keduanya memiliki nilai pada seluruh variabel biner dan kategori yang sama maka total kedekatannya sebesar 5. Perhitungan dilakukan untuk 165 objek. Hasil perhitungan total jarak variabel biner dan kategorik disajikan matriks $d_{binary/categorical}$ ordo 165×165 dalam **Tabel 5.4**.

Tabel 5.4 Matriks Kedekatan Objek Data Biner dan Kategorik

	S1	S2	S3	S4	S5	S6	S7	S8	...	S165
S1	5	5	5	5	5	5	5	5	...	2
S2	5	5	5	5	5	5	5	5	...	2
S3	5	5	5	5	5	5	5	5	...	2
S4	5	5	5	5	5	5	5	5	...	2
S5	5	5	5	5	5	5	5	5	...	2
S6	5	5	5	5	5	5	5	5	...	2
S7	5	5	5	5	5	5	5	5	...	2
S8	5	5	5	5	5	5	5	5	...	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
S165	2	2	2	2	2	2	2	2	...	5

Berdasarkan **Tabel 5.4** dapat diketahui seberapa besar kedekatan antara objek satu dengan yang lainnya. Misalkan kedekatan objek data numerik untuk S3 (SMA Negeri 1 Bambanglipuro) dengan S6 (SMA Negeri 1 Jetis) adalah sebesar 5. Nilai maksimum kedekatan data biner dan kategorik adalah 5 sesuai dengan jumlah banyaknya variabel biner dan kategorik yaitu 5.

3. Pengukuran Kedekatan Objek Variabel Campuran

Proses pengukuran kedekatan objek dengan variabel numerik, biner dan kategorik maka akan menggunakan **Persamaan 3.3**. Perhitungan ini dilakukan untuk seluruh data dan jenis variabel yaitu 8 variabel numerik dan 5 variabel biner serta kategorik. Misalnya perhitungan ukuran kedekatan variabel campuran dengan *Gower* objek S1 dengan S2 serta objek S1 dengan S3 adalah

$$s_{(S_1, S_2)} = \left(\frac{6.96+5}{13} \right) = 0.92$$

$$s_{(S_1, S_3)} = \left(\frac{7.58+5}{13} \right) = 0.97$$

Tabel 5.5 Matriks Kedekatan Objek Data Campuran

	S1	S2	S3	S4	S5	S6	S7	S8	...	S165
S1	1	0.92	0.97	0.96	0.96	0.94	0.95	0.92	...	0.65
S2	0.92	1	0.90	0.88	0.93	0.87	0.90	0.91	...	0.64
S3	0.97	0.90	1	0.94	0.95	0.95	0.96	0.91	...	0.64
S4	0.96	0.88	0.94	1	0.93	0.97	0.92	0.93	...	0.62
S5	0.96	0.93	0.95	0.93	1	0.92	0.96	0.89	...	0.69
S6	0.94	0.87	0.95	0.97	0.92	1	0.91	0.94	...	0.62
S7	0.95	0.90	0.96	0.92	0.96	0.91	1	0.89	...	0.68
S8	0.92	0.91	0.91	0.93	0.89	0.94	0.89	1	...	0.59
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
S165	0.65	0.64	0.64	0.62	0.69	0.62	0.68	0.59	...	1.00

Berdasarkan **Tabel 5.5** dapat diketahui seberapa besar kedekatan antara objek satu dengan yang lainnya. Misalkan kedekatan objek data campuran S3 (SMA Negeri 1 Bambanglipuro) dengan S6 (SMA Negeri 1 Jetis) sebesar 0.95. Nilai kedekatan data campuran memiliki nilai maksimum sebesar 1 karena hasil dari penjumlahan data numerik dengan biner dan kategorik dibagi dengan jumlah variabel yang digunakan. Pada data numerik ada sebanyak 8 variabel serta data biner dan kategorik ada 5 variabel sehingga jumlah nilai pembaginya adalah 13.

4. Pengukuran Jarak Objek Variabel Campuran

Proses pengukuran jarak objek dengan variabel campuran yaitu numerik, biner dan kategorik akan menggunakan **Persamaan 3.4**. Perhitungan dilakukan untuk seluruh data dan jenis variabel yaitu 165 data dengan 13 variabel. Misalnya ukuran kedekatan variabel campuran dengan jarak *Gower* pada objek S1 dengan S2 serta objek S1 dengan S3 adalah

$$s_{(S_1, S_2)} = 1 - 0.96 = 0.04$$

$$s_{(S_1, S_3)} = 1 - 0.98 = 0.02$$

Hasil perhitungan jarak data dengan variabel campuran disajikan pada **Tabel 5.6**.

Tabel 5.6 Matriks Jarak Data Campuran

	S1	S2	S3	S4	S5	S6	S7	S8	...	S165
S1	0	0.08	0.03	0.04	0.04	0.06	0.05	0.08	...	0.35
S2	0.08	0	0.10	0.12	0.07	0.13	0.10	0.09	...	0.36
S3	0.03	0.10	0	0.06	0.05	0.05	0.04	0.09	...	0.36
S4	0.04	0.12	0.06	0	0.07	0.03	0.08	0.07	...	0.38
S5	0.04	0.07	0.05	0.07	0	0.08	0.04	0.11	...	0.31
S6	0.06	0.13	0.05	0.03	0.08	0	0.09	0.06	...	0.38
S7	0.05	0.10	0.04	0.08	0.04	0.09	0	0.11	...	0.32
S8	0.08	0.09	0.09	0.07	0.11	0.06	0.11	0	...	0.41
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
S165	0.35	0.36	0.36	0.38	0.31	0.38	0.32	0.41	...	0

Tabel 5.6 diperoleh besarnya jarak antara satu objek dengan yang lainnya. Hasil perhitungan jarak data campuran pada **Tabel 5.6** yang digunakan dalam pengelompokan metode SKM. Besar nilai jarak selalu berada pada rentang 0 sampai 1. Jika jarak bernilai 1 menandakan kedekatan antara dua objek lemah dan jika jarak bernilai 0 maka kedekatan antara dua objek tersebut sangat kuat.

Misalkan jarak antara objek S8 (SMA Negeri 1 Banguntapan) dengan S5 (SMA Negeri 1 Kretek) sebesar 0.11 yang artinya kedekatan antara kedua objek tersebut sangat kuat karena nilainya mendekati 0. Sehingga besar kemungkinannya jika antara objek S8 dengan S5 berada dalam satu klaster yang sama.

5.4. Pengelompokan SMA di Provinsi Yogyakarta

Pengelompokan sekolah SMA di Provinsi Yogyakarta dilakukan dengan data campuran yaitu data numerik dan kategorik. Proses pengelompokan menggunakan data campuran yaitu 8 variabel numerik dan 5 kategorik. Profilisasi dilakukan untuk mengetahui karakteristik variabel dari setiap klaster.

5.4.1 Pengelompokan dengan *Simple K-Medoids Partitioning*

Pengelompokan sekolah SMA di Provinsi Yogyakarta akan dilakukan dengan menggunakan metode SKM. Data yang digunakan merupakan data dengan variabel campuran yang terdiri dari numerik, biner dan kategorik.

Pada penelitian ini, percobaan pengelompokan data akan dibagi dalam 2, 3, 4 dan 5 kelompok. Percobaan pengelompokan yang pertama di mana data akan dibagi kedalam 2 kelompok. Sesuai tahapan algoritma SKM, langkah awal adalah memilih medoid awal yang merupakan objek yang meminimalkan jumlah jarak menuju

objek lain dengan menggunakan nilai a_i yang akan dihitung berdasarkan pada **Persamaan 3.10**. Berdasarkan nilai jarak pada **Tabel 5.6** didapatkan nilai a_i yang disajikan pada **Tabel 5.7**.

Misalkan untuk menghitung nilai a_{S1} dan a_{S2} diperoleh sebagai berikut:

$$\begin{aligned} a_{S1} &= 0 + 0.08 + 0.03 + 0.04 + 0.04 + 0.06 + 0.05 + 0.08 + \dots + 0.35 \\ &= 32.6138 \end{aligned}$$

$$\begin{aligned} a_{S2} &= 0.08 + 0 + 0.10 + 0.12 + 0.07 + 0.1 + 0.10 + 0.09 + \dots + 0.36 \\ &= 36.74313 \end{aligned}$$

Tabel 5.7 Nilai a_i

Objek	a_i
S84	26.98996
S103	27.06907
S161	27.45945
S139	27.47312
S163	27.49228
S65	27.51961
S23	27.52065
S63	27.53716
$\vdots$	$\vdots$
S54	61.13408

Medoid awal yang terbentuk dari proses *seeding* yang pertama yaitu objek ke-84 dan objek ke-33. Objek ke-84 merupakan SMAS Kolombo Depok sebagai medoid kluster 1 (atau M1) dan objek ke-33 merupakan SMA PGRI Kasihan sebagai medoid kluster 2 (atau M2). Objek ke-84 merupakan objek dengan nilai a_i paling rendah dan objek ke-33 terpilih secara random untuk menjadi medoid awal kluster 2.

Langkah kedua adalah menghitung jarak setiap objek menuju ke medoidnya yang dapat dilihat pada matriks **Tabel 5.8**. Proses pengelompokan dilakukan dengan batas iterasi sebanyak 10 iterasi dan *seeding* sebanyak 20 kali. Maksud dari `set.seed(1)` adalah untuk mereproduksi hasil agar hasilnya sama ketika dijalankan pada mesin yang lain. Pemilihan *seeding* = 20 berarti akan dilakukan proses pembaruan medoid awal sebanyak 20 kali dengan *iterate* maksimum pada setiap *seeding* adalah 10 kali. Sehingga diperoleh medoid final berdasarkan nilai E terkecil yang disajikan pada **Tabel 5.8**.

Tabel 5.8 Medoid Akhir Hasil Pengelompokan

Objek ke- <i>i</i>	Nama Sekolah	Keterangan
S19	SMA Negeri 1 Sanden	Data ke-19 sebagai medoid akhir klaster 1
S71	SMAS Muhammadiyah 1 Sleman	Data ke-71 sebagai medoid akhir klaster 2

Berdasarkan **Tabel 5.8** didapatkan medoid akhir dari masing-masing klaster. Medoid akhir klaster 1 adalah objek S19 atau SMA Negeri 1 Sanden dan klaster 2 adalah objek S71 atau SMAS Muhammadiyah 1 Sleman. Adapun jarak antar objek dengan medoid akhir disajikan pada **Tabel 5.9**.

Tabel 5.9 Jarak Objek dengan Medoid Akhir

Objek ke- <i>i</i>	Jarak ke Medoids		Jarak Terdekat	Klaster
	S19 (C1)	S71 (C2)		
S1	0.0293	0.3283	0.0293	1
S2	0.0857	0.3328	0.0857	1
S3	0.0286	0.3299	0.0286	1
S4	0.0564	0.3567	0.0564	1
S5	0.0340	0.2906	0.0340	1
S6	0.0507	0.3607	0.0507	1
S7	0.0415	0.3957	0.0415	1
S8	0.0880	0.3981	0.0880	1
⋮	⋮	⋮	⋮	⋮
S165	0.3360	0.0371	0.0371	2
Total Jarak (<i>Minimum Distance</i>)			13.9869	

Berdasarkan **Tabel 5.9** dapat diketahui jarak masing-masing objek menuju pusat klaster akhir atau medoid akhir. Karena data dikelompokkan menjadi 2 maka setiap objek akan mempunyai dua nilai jarak menuju masing – masing pusat klaster (medoid). Akan tetapi yang menjadi jarak terdekat adalah jarak terpendek objek menuju ke salah satu medoid klaster.

Misalkan pada objek S1, jarak menuju medoid klaster 1 (S19) sebesar 0.0293 dan klaster 2 (S71) sebesar 0.3283. Nilai jarak S1 menuju S19 adalah jarak terdekat objek menuju pusat klaster sehingga dapat dikatakan bahwa objek S1 masuk dalam keanggotaan klaster 1. Nilai jarak terdekat seluruh objek dijumlahkan dan didapatkan nilai *minimum distance* yang paling kecil yaitu hasil iterasi akhir dari *seeding* dengan medoid akhirnya adalah S19 dan S71 yaitu sebesar 13.9869. Adapun ringkasan dari hasil pengelompokan menggunakan 20 *seeding* dan iterasi maksimal pada setiap *seeding* sebanyak 10 kali disajikan pada **Tabel 5.10**.

Tabel 5.10 Ringkasan Hasil Pengelompokan dengan *seeding* = 20

<i>Seeding</i>	Iterasi	Medoid		Nilai <i>E</i>	<i>Seeding</i>	Iterasi	Medoid		Nilai <i>E</i>
		<i>M1</i>	<i>M2</i>				<i>M1</i>	<i>M2</i>	
1	Awal	S84	S33	21.5626	11	Awal	S84	S10	20.3830
	Akhir	S19	S25	14.0042		Akhir	S25	S19	14.0042
2	Awal	S84	S91	20.5610	12	Awal	S84	S45	20.5610
	Akhir	S25	S19	14.0042		Akhir	S25	S19	14.0042
3	Awal	S84	S38	21.0737	13	Awal	S84	S165	21.2654
	Akhir	S25	S136	14.0034		Akhir	S19	S71	13.9869
4	Awal	S84	S26	22.2622	14	Awal	S84	S116	20.0711
	Akhir	S19	S71	13.9869		Akhir	S25	S136	14.0034
5	Awal	S84	S10	20.3840	15	Awal	S84	S85	26.8921
	Akhir	S25	S19	14.0042		Akhir	S84	S85	26.8921
6	Awal	S84	S141	22.3068	16	Awal	S84	S49	20.5483
	Akhir	S19	S71	13.9869		Akhir	S25	S136	14.0034
7	Awal	S84	S31	24.4914	17	Awal	S84	S153	22.4264
	Akhir	S19	S71	13.9869		Akhir	S19	S71	13.9869
8	Awal	S84	S81	21.1925	18	Awal	S84	S162	26.5587
	Akhir	S19	S71	13.9869		Akhir	S84	S54	26.4388
9	Awal	S84	S80	36.4640	19	Awal	S84	S116	23.3607
	Akhir	S19	S71	13.9869		Akhir	S25	S136	14.0034
10	Awal	S84	S82	21.5147	20	Awal	S84	S147	25.6213
	Akhir	S19	S25	14.0042		Akhir	S19	S25	14.0042

Proses *seeding* merupakan jumlah inisialisasi untuk melakukan pembaruan medoid awal. Berdasarkan 20 kali *seeding* yang dilakukan terlihat pada **Tabel 5.10** bahwa objek S84 selalu menjadi *initial* medoid pada setiap awal proses iterasi dilakukan. Karena objek S84 merupakan objek yang paling pusat dan meminimalkan jumlah jarak antar objek yang berada dalam satu kluster. Pada setiap proses *seeding* selain *initial* medoid, medoid yang lainnya terpilih secara acak. Penentuan medoid final dipilih berdasarkan nilai *E* yang paling minimum. Beberapa nilai *E* yang dihasilkan, diperoleh nilai *E* terkecil yaitu sebesar 13.9869 yang didapatkan dari medoid *M1* yaitu objek S19 dan medoid *M2* yaitu objek S71.

Hasil pengelompokan SMA di Provinsi Yogyakarta dengan menggunakan algoritma SKM yang dibagi dalam 2 kluster di mana pada kluster 1 terdiri dari 83 sekolah dan anggota kluster 2 ada sebanyak 82 sekolah. Pada hasil proses *seeding* untuk pengelompokan 3, 4 dan 5 kluster di tuliskan pada **Lampiran 8**.

5.4.2 Validasi Pengelompokan *Simple K-Medoids Partitioning*

Hasil pengelompokan menggunakan metode SKM divalidasi menggunakan dua metode kriteria internal yaitu *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV). Berikut penjelasan hasil validasi menggunakan kedua metode tersebut.

a. *Silhouette Index* (SI)

Perhitungan validasi hasil klaster akan menggunakan nilai *Silhouette Index* (SI). Jika nilai SI mendekati angka 1, maka dapat dikatakan bahwa klaster yang dihasilkan semakin baik. Begitu sebaliknya apabila nilai SI mendekati angka -1 maka dikatakan bahwa klaster yang dihasilkan adalah semakin buruk.

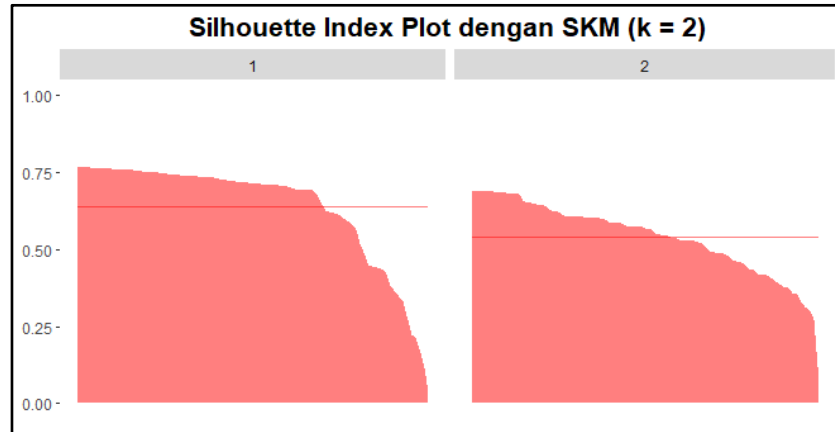
Percobaan dengan jumlah k klaster dilakukan untuk mengetahui nilai SI dari masing – masing klaster. Adapun jumlah k klaster yang akan digunakan adalah 2, 3, 4 dan 5. Nilai SI dari setiap percobaan akan divisualisasikan dalam bentuk grafik. Daerah arsiran berwarna merah merupakan luasan dari seluruh nilai SI objek yang masuk dalam keanggotaan suatu klaster. Sedangkan garis merah adalah rata – rata SI dari suatu klaster.

a) Percobaan dengan 2 Klaster

Pada percobaan pertama menggunakan jumlah k klaster sebanyak 2 klaster berdasarkan **Persamaan 3.15**. Diperoleh nilai SI setiap objek yang disajikan pada **Tabel 5.11** dan grafik nilai SI pada **Gambar 5.17**.

Tabel 5.11 Nilai SI: Percobaan dengan 2 Klaster

	Nilai SI	Klaster
S1	0.7591	1
S2	0.6266	1
S3	0.7585	1
⋮	⋮	⋮
S165	0.6416	2



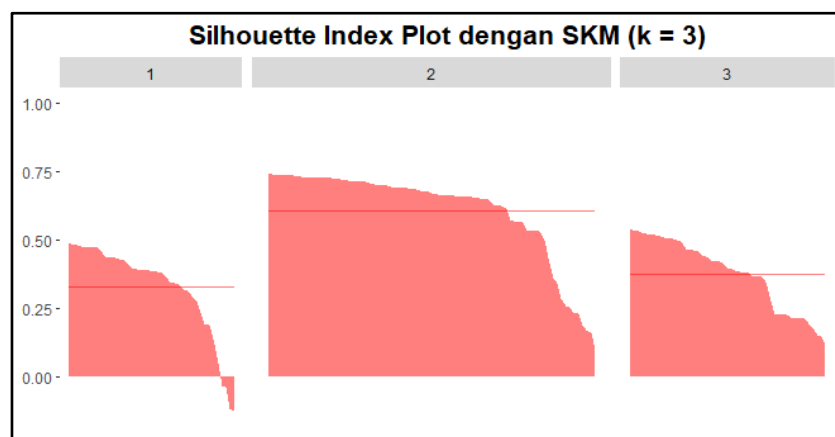
Gambar 5.17 Grafik SI: Percobaan dengan 2 Klaster

b) Percobaan dengan 3 Klaster

Pada percobaan kedua menggunakan jumlah k klaster sebanyak 3 klaster berdasarkan **Persamaan 3.15**. Diperoleh nilai SI setiap objek yang disajikan pada **Tabel 5.12** dan grafik nilai SI pada **Gambar 5.18**.

Tabel 5.12 Nilai SI: Percobaan dengan 3 Klaster

	Nilai SI	Klaster
S1	0.7268	2
S2	0.5652	2
S3	0.7278	2
⋮	⋮	⋮
S165	0.4600	3



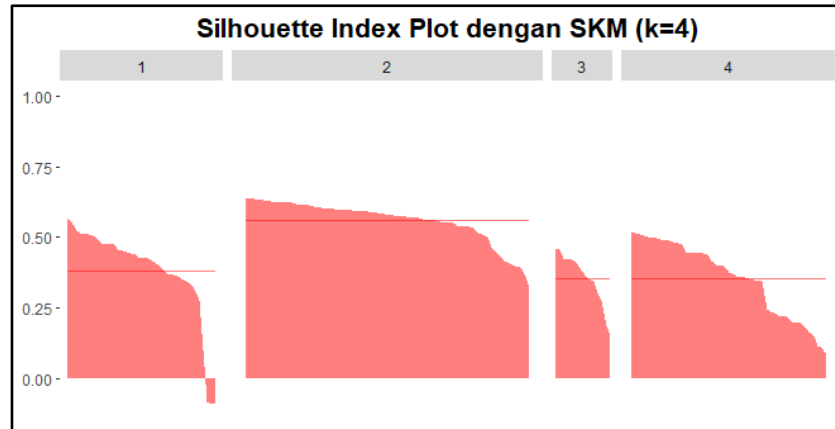
Gambar 5.18 Grafik SI: Percobaan dengan 3 Klaster

c) Percobaan dengan 4 Klaster

Pada percobaan ketiga menggunakan jumlah k klaster sebanyak 4 klaster berdasarkan **Persamaan 3.15**. Diperoleh nilai SI setiap objek yang disajikan pada **Tabel 5.13** dan grafik nilai SI pada **Gambar 5.19**.

Tabel 5.13 Nilai SI: Percobaan dengan 4 Klaster

	Nilai SI	Klaster
S1	0.6231	2
S2	0.4507	2
S3	0.6221	2
⋮	⋮	⋮
S165	0.4397	4



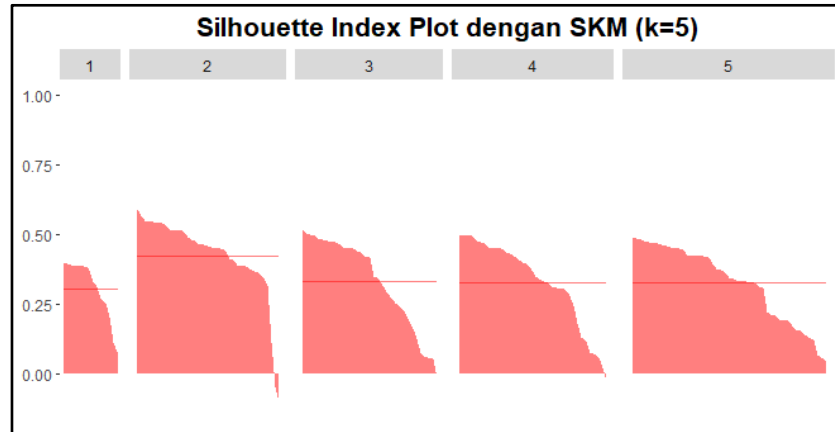
Gambar 5.19 Grafik SI: Percobaan dengan 4 Klaster

d) Percobaan dengan 5 Klaster

Pada percobaan keempat menggunakan jumlah k klaster sebanyak 5 klaster berdasarkan **Persamaan 3.15**. Diperoleh nilai SI setiap objek yang disajikan pada **Tabel 5.14** dan grafik nilai SI pada **Gambar 5.20**.

Tabel 5.14 Nilai SI: Percobaan dengan 5 Klaster

	Nilai SI	Klaster
S1	0.0743	4
S2	0.2310	3
S3	0.1812	4
⋮	⋮	⋮
S165	0.4183	5



Gambar 5.20 Grafik SI: Percobaan dengan 5 Klaster

Nilai SI global merupakan nilai *silhouette index* untuk jumlah k klaster tertentu. Berdasarkan dari percobaan dengan jumlah k klaster, didapatkan nilai SI global dari masing – masing percobaan berdasarkan **Persamaan 3.17** yang disajikan pada **Tabel 5.15**.

Tabel 5.15 Nilai *Silhouette Index*

Percobaan k Klaster	Nilai <i>Silhouette Index</i> (SI) Global
2 Klaster	0.5886
3 Klaster	0.4737
4 Klaster	0.4448
5 Klaster	0.3470

Penentuan nilai SI terbaik apabila nilainya mendekati 1. Jika dilihat pada **Tabel 5.15** nilai tertinggi di antara empat percobaan yakni ketika jumlah k klaster sebanyak 2 dengan nilai SI sebesar 0.5886. Sehingga dengan kata lain jumlah k klaster optimal untuk pengelompokan SMA di Provinsi Yogyakarta menggunakan algoritma SKM adalah 2 klaster.

b. Medoid Shadow Value (MSV)

Validasi kedua dilakukan menggunakan *Medoid Shadow Value* (MSV). Nilai MSV berada pada rentang 0 sampai 1. Jika nilai MSV mendekati angka 1 artinya *well-separated*, dapat dikatakan objek tepat jika berada dalam klaster tersebut sebaliknya jika nilai mendekati angka 0 artinya *poor-separated* maka objek kurang tepat apabila berada dalam klaster tersebut (Budiaji, Weksi, 2019).

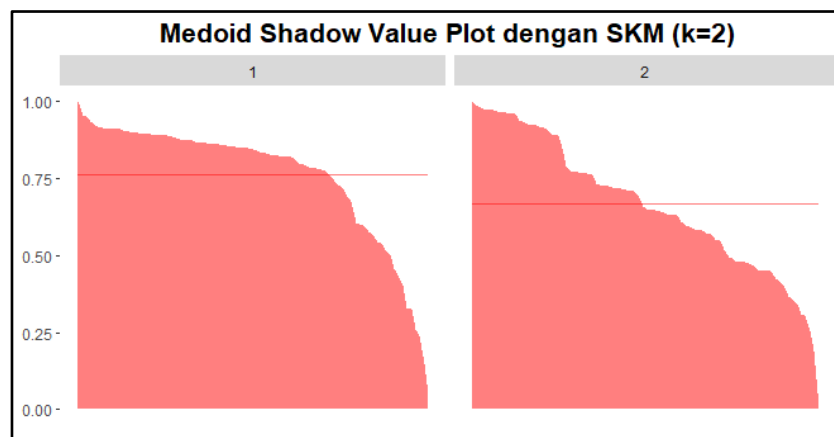
Percobaan dengan jumlah k klaster dilakukan untuk mengetahui nilai MSV dari masing – masing klaster. Jumlah k klaster yang akan digunakan adalah 2, 3, 4 dan 5. Nilai MSV juga divisualisasikan dengan grafik yang sama seperti nilai SI.

a) Percobaan dengan 2 Klaster

Pada percobaan pertama menggunakan jumlah k klaster sebanyak 2 klaster berdasarkan **Persamaan 3.18**. Diperoleh nilai MSV setiap objek yang disajikan pada **Tabel 5.16** dan grafik nilai MSV pada **Gambar 5.21**.

Tabel 5.16 Nilai MSV: Percobaan dengan 2 Klaster

	Nilai MSV	Klaster
S1	0.9107	1
S2	0.7424	1
S3	0.9134	1
⋮	⋮	⋮
S165	0.8895	2



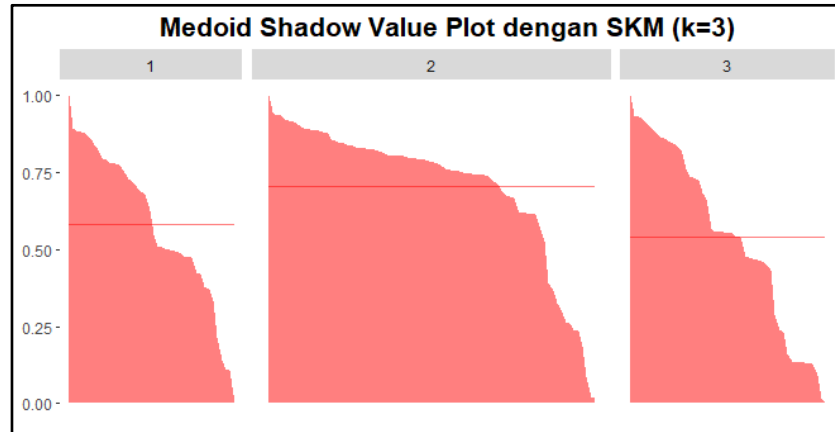
Gambar 5.21 Grafik MSV: Percobaan dengan 2 Klaster

b) Percobaan dengan 3 Klaster

Pada percobaan kedua menggunakan jumlah k klaster sebanyak 3 klaster berdasarkan **Persamaan 3.18**. Diperoleh nilai MSV setiap objek yang disajikan pada **Tabel 5.17** dan grafik nilai MSV pada **Gambar 5.22**.

Tabel 5.17 Nilai MSV: Percobaan dengan 3 Klaster

	Nilai MSV	Klaster
S1	0.8954	2
S2	0.6119	2
S3	0.9217	2
⋮	⋮	⋮
S165	0.7364	3



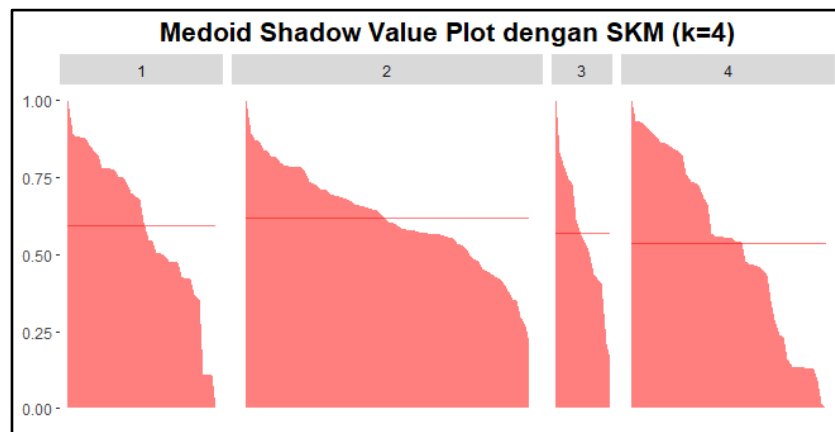
Gambar 5.22 Grafik MSV: Percobaan dengan 3 Klaster

c) Percobaan dengan 4 Klaster

Pada percobaan ketiga menggunakan jumlah k klaster sebanyak 4 klaster berdasarkan **Persamaan 3.18**. Diperoleh nilai MSV setiap objek yang disajikan pada **Tabel 5.18** dan grafik nilai MSV pada **Gambar 5.23**.

Tabel 5.18 Nilai MSV: Percobaan dengan 4 Klaster

	Nilai MSV	Klaster
S1	0.7976	2
S2	0.3551	2
S3	0.8409	2
⋮	⋮	⋮
S165	0.7364	4



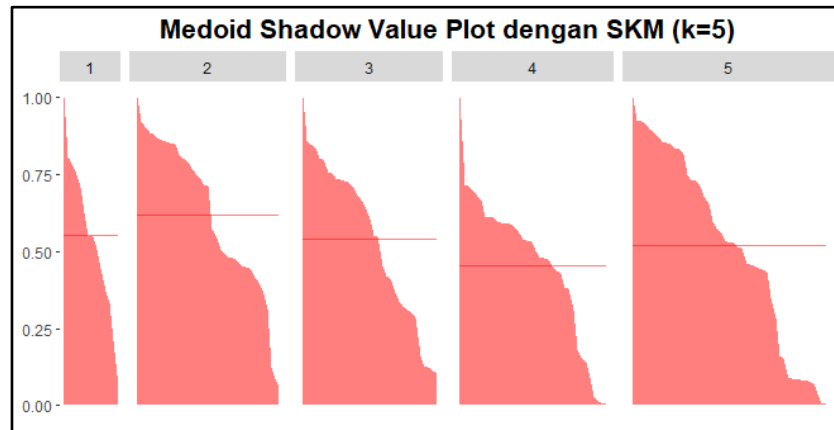
Gambar 5.23 Grafik MSV: Percobaan dengan 4 Klaster

d) Percobaan dengan 5 Klaster

Pada percobaan keempat menggunakan jumlah k klaster sebanyak 5 klaster berdasarkan **Persamaan 3.18**. Diperoleh nilai MSV setiap objek yang disajikan pada **Tabel 5.19** dan grafik nilai MSV pada **Gambar 5.24**.

Tabel 5.19 Nilai MSV: Percobaan dengan 5 Klaster

	Nilai MSV	Klaster
S1	0.0038	4
S2	0.3011	3
S3	0.4289	4
⋮	⋮	⋮
S165	0.7293	5



Gambar 5.24 Grafik MSV: Percobaan dengan 5 Klaster

Nilai MSV global merupakan nilai *medoid shadow value* untuk jumlah k klaster tertentu. Berdasarkan dari percobaan dengan jumlah k klaster, didapatkan nilai MSV global dari masing – masing percobaan yang disajikan pada **Tabel 5.20**.

Tabel 5.20 Nilai *Medoid Shadow Value* (MSV)

Percobaan k Klaster	Nilai <i>Medoid Shadow Value</i> (MSV) Global
2 Klaster	0.7129
3 Klaster	0.6272
4 Klaster	0.5853
5 Klaster	0.5327

Penentuan nilai MSV terbaik apabila nilainya mendekati 1. Jika dilihat pada **Tabel 5.20** nilai tertinggi di antara empat percobaan yang dilakukan adalah percobaan dengan jumlah k klaster sebanyak 2 klaster. Sehingga dapat dikatakan bahwa jumlah k klaster yang optimal untuk pengelompokan SMA di Provinsi Yogyakarta dengan menggunakan algoritma SKM adalah 2 klaster.

5.4.3 Profilisasi Pengelompokan *Simple K-Medoids Partitioning*

Hasil pengelompokan menggunakan algoritma SKM untuk SMA di Provinsi Yogyakarta terbagi dalam 2 klaster. Hasil keanggotaan masing – masing objek yang terbentuk dari hasil pengelompokan tersebut disajikan pada **Tabel 5.21**.

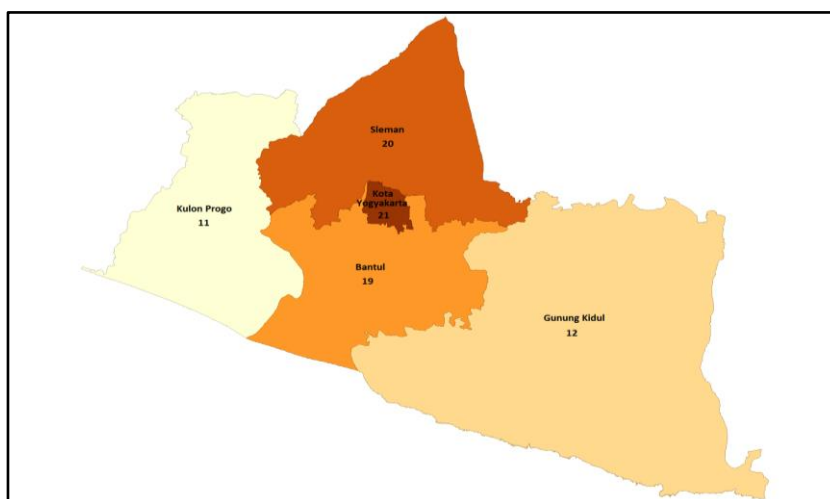
Tabel 5.21 Hasil Pengelompokan dengan SKM

Klaster	Anggota Klaster	Jumlah
1	S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14, S15, S16, S17, S18, S19, S37, S38, S39, S40, S41, S42, S43, S44, S45, S46, S47, S48, S49, S50, S51, S52, S53, S68, S84, S85, S86, S87, S88, S89, S90, S91, S92, S93, S94, S95, S96, S103, S109, S110, S111, S112, S113, S114, S115, S116, S117, S118, S119, S126, S127, S128, S129, S130, S131, S132, S133, S134, S135, S136, S137, S139, S140, S143, S150, S157, S159, S160, S162, S163	83
2	S20, S21, S22, S23, S24, S25, S26, S27, S28, S29, S30, S31, S32, S33, S34, S35, S36, S54, S55, S56, S57, S58, S59, S60, S61, S62, S63, S64, S65, S66, S67, S69, S70, S71, S72, S73, S74, S75, S76, S77, S78, S79, S80, S81, S82, S83, S97, S98, S99, S100, S101, S102, S104, S105, S106, S107, S108, S120, S121, S122, S123, S124, S125, S138, S141, S142, S144, S145, S146, S147, S148, S149, S151, S152, S153, S154, S155, S156, S158, S161, S164, S165	82

Adapun karakteristik data dari masing – masing keanggotaan klaster dijelaskan sebagai berikut:

a. Karakteristik Klaster 1

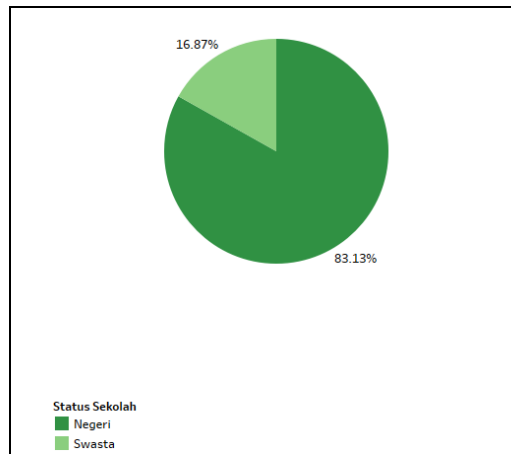
Pada klaster 1 hasil pengelompokan menggunakan SKM diperoleh anggota klaster sebanyak 83 sekolah. Berikut adalah karakteristik variabel dari anggota pada klaster 1. Jumlah sekolah SMA di Provinsi Yogyakarta yang masuk dalam klaster 1 berdasarkan kabupaten dan akreditasi sekolah ditunjukkan pada **Gambar 5.25**.



Gambar 5.25 Profil Klaster 1: Akreditasi Berdasarkan Wilayah

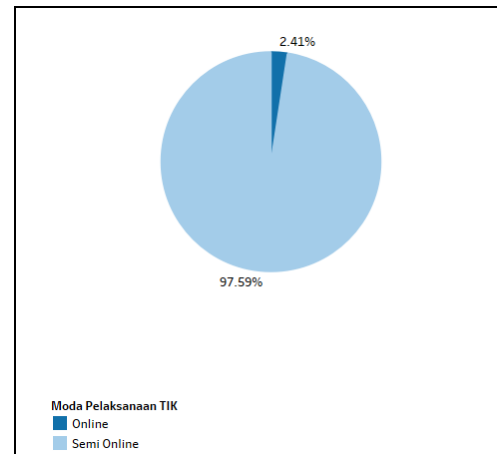
Pada **Gambar 5.25** menyajikan peta sebaran banyaknya sekolah berdasarkan akreditasi dan wilayah sekolah. Seluruh sekolah yang masuk dalam anggota klaster

1 terakreditasi A. Terlihat bahwa Kota Yogyakarta menjadi wilayah terbanyak sekolah yang masuk dalam kluaster 1 yaitu 21 sekolah, disusul Kabupaten Sleman sebanyak 20 sekolah. Kulon Progo menjadi wilayah dengan jumlah sekolah paling rendah yang masuk dalam klaster 1 yaitu 11 sekolah. Persentase status sekolah untuk klaster 1 disajikan pada **Gambar 5.26** sedangkan persentase moda pelaksanaan TIK disajikan pada **Gambar 5.27**.



Gambar 5.26 Profil Klaster 1:

Persentase Status Sekolah

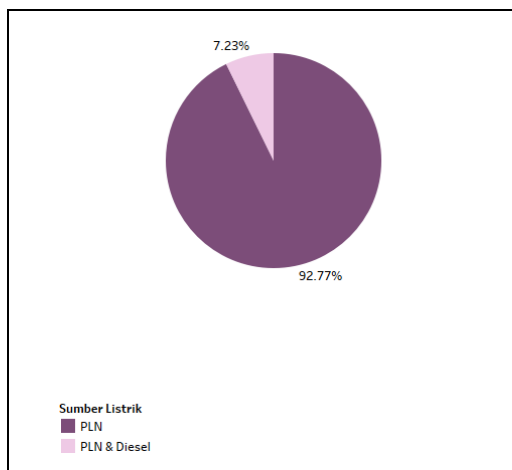


Gambar 5.27 Profil Klaster 1:

Persentase Moda Pelaksanaan TIK

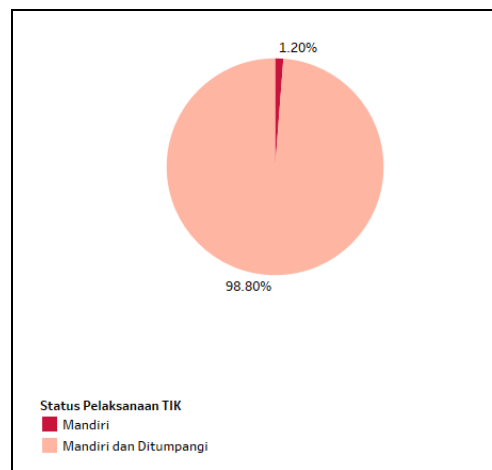
Berdasarkan *pie chart* pada **Gambar 5.26** menunjukkan bahwa pada klaster 1 lebih didominasi oleh sekolah negeri yaitu sebesar 83.13% atau 69 sekolah. Sisanya adalah sekolah swasta dengan persentase sebesar 16.87% atau 14 sekolah.

Berdasarkan *pie chart* pada **Gambar 5.27** menunjukkan bahwa pada klaster 1, sekolah dengan moda pelaksanaan TIK secara semi online lebih mendominasi dengan persentase sebesar 97.59% atau 81 sekolah dan sisanya sebesar 2.41 % atau 2 sekolah merupakan sekolah dengan menggunakan moda pelaksanaan TIK secara online. Persentase sumber listrik untuk klaster 1 disajikan pada **Gambar 5.28** sedangkan persentase status pelaksanaan TIK disajikan pada **Gambar 5.29**.



Gambar 5.28 Profil Klaster 1:

Persentase Sumber Listrik

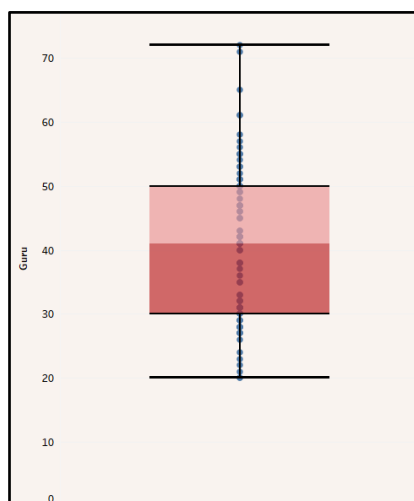


Gambar 5.29 Profil Klaster 1:

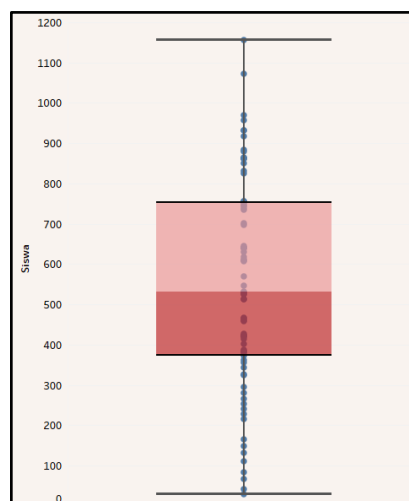
Persentase Status Pelaksanaan TIK

Berdasarkan *pie chart* pada **Gambar 5.28** menunjukkan bahwa pada klaster 1 berisikan sekolah yang memiliki sumber listrik berasal dari PLN sebesar 92.77% atau 77 sekolah serta sisanya sebesar 7.23% atau 6 sekolah menggunakan sumber listrik yang berasal dari PLN dan diesel.

Berdasarkan *pie chart* pada **Gambar 5.29** menunjukkan bahwa klaster 1, sebesar 98.80% atau 82 sekolah memiliki status pelaksanaan TIK berupa mandiri dan ditumpangi. Sisanya 1.27% atau 1 sekolah, merupakan sekolah dengan status pelaksanaan TIK berupa mandiri. Karakteristik jumlah guru untuk klaster 1 disajikan pada **Gambar 5.30** dan untuk jumlah siswa disajikan pada **Gambar 5.31**.



Gambar 5.30 Profil Klaster 1: Guru

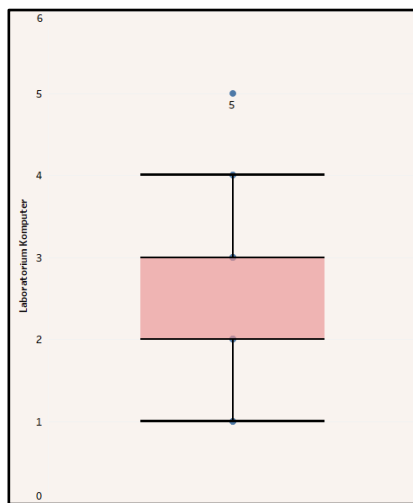


Gambar 5.31 Profil Klaster 1: Siswa

Berdasarkan *boxplot* pada **Gambar 5.30** didapatkan bahwa nilai rata – rata jumlah guru pada klaster 1 sebesar 41 guru. Nilai minimum jumlah guru klaster 1 sebesar 20 guru dan nilai maksimum sebesar 72 guru. Nilai tengah (median) jumlah

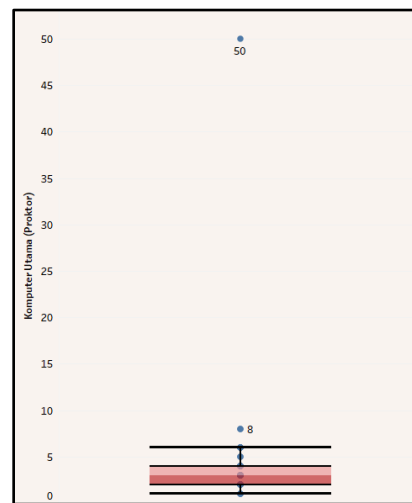
guru pada klaster 1 sebesar 41 guru, nilai kuartil pertama (Q1) sebesar 30 guru dan nilai kuartil ketiga (Q3) sebesar 50 guru.

Berdasarkan *boxplot* pada **Gambar 5.31** didapatkan bahwa nilai rata – rata jumlah siswa klaster 1 sebesar 554 siswa dengan nilai minimum sebesar 30 siswa dan nilai maksimum sebesar 1157 siswa. Nilai tengah jumlah siswa klaster 1 sebesar 532 siswa, kuartil pertama sebesar 375 siswa dan kuartil ketiga sebesar 753 siswa. Karakteristik jumlah laboratorium komputer untuk klaster 1 disajikan pada **Gambar 5.32** dan jumlah komputer utama/ proktor disajikan pada **Gambar 5.33**.



Gambar 5.32 Profil Klaster 1:

Laboratorium Komputer

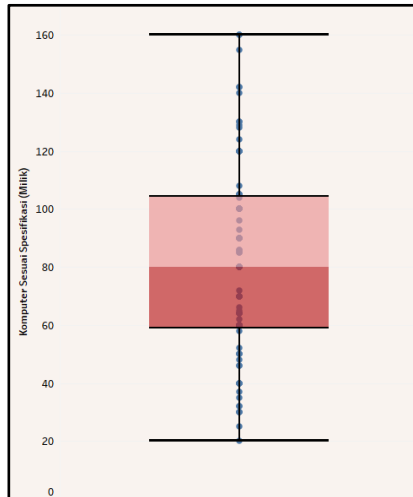


Gambar 5.33 Profil Klaster 1:

Komputer Utama

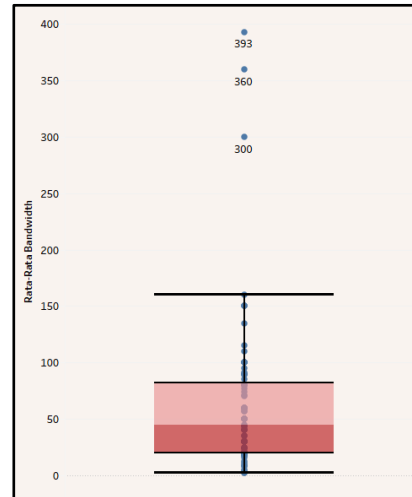
Berdasarkan *boxplot* pada **Gambar 5.32** didapatkan bahwa nilai rata – rata jumlah laboratorium komputer pada klaster 1 adalah 3 ruang. Nilai minimum jumlah laboratorium komputer pada klaster 1 adalah 1 ruang dan nilai maksimum adalah 5 ruang. Kemudian untuk nilai tengah (atau median) jumlah laboratorium komputer pada klaster 1 adalah 2 ruang, nilai kuartil pertama (atau Q1) adalah 2 ruang dan nilai kuartil ketiga (atau Q3) adalah 3 ruang.

Berdasarkan *boxplot* pada **Gambar 5.33** didapatkan bahwa nilai rata – rata jumlah komputer utama pada klaster 1 adalah 4 unit. Nilai minimum jumlah komputer utama pada klaster 1 adalah 1 unit dan nilai maksimum adalah 50 unit. Untuk nilai tengah (median) jumlah komputer utama klaster 1 adalah 3 unit, nilai kuartil pertama (Q1) adalah 2 unit dan nilai kuartil ketiga (Q3) adalah 4 unit. Karakteristik jumlah komputer sesuai spesifikasi (milik) klaster 1 disajikan pada **Gambar 5.34** sedangkan untuk jumlah *switch hub* disajikan pada **Gambar 5.35**.



Gambar 5.34 Profil Klaster 1:

Komputer Sesuai Spesifikasi (Milik)

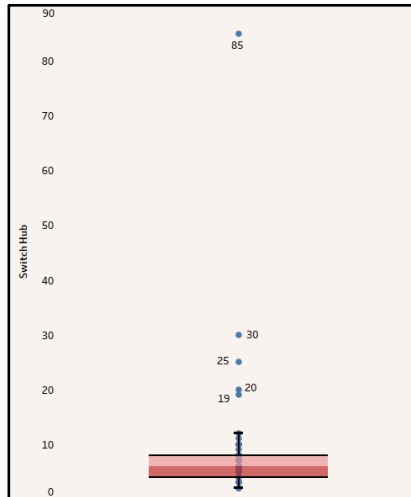


Gambar 5.35 Profil Klaster 1: Rata-

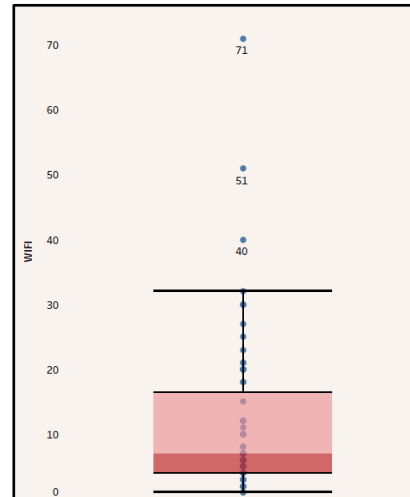
Rata *Bandwidth*

Berdasarkan *boxplot* pada **Gambar 5.34** didapatkan bahwa nilai rata – rata jumlah komputer sesuai spesifikasi (milik) pada klaster 1 adalah 82 unit. Nilai minimum jumlah komputer sesuai spesifikasi (milik) pada klaster 1 adalah 20 unit dan nilai maksimum adalah 160 unit. Untuk nilai tengah (atau median) jumlah komputer sesuai spesifikasi (milik) pada klaster 1 adalah 80 unit, nilai kuartil pertama (atau Q1) adalah 59 unit dan nilai kuartil ketiga (atau Q3) adalah 105 unit.

Berdasarkan *boxplot* pada **Gambar 5.35** didapatkan bahwa nilai *mean* rata-rata *bandwidth* pada klaster 1 adalah 62.98 Mbps. Nilai minimum rata-rata *bandwidth* pada klaster 1 adalah 2 Mbps dan nilai maksimum adalah 393 Mbps. Nilai tengah (atau median) rata-rata *bandwidth* klaster 1 adalah 45 Mbps, nilai kuartil pertama (atau Q1) adalah 20 Mbps dan nilai kuartil ketiga (atau Q3) adalah 82.50 Mbps. Karakteristik jumlah *switch hub* untuk klaster 1 disajikan pada **Gambar 5.36** sedangkan untuk jumlah WIFI disajikan pada **Gambar 5.37**.



Gambar 5.36 Profil Klaster 1: *Switch Hub*



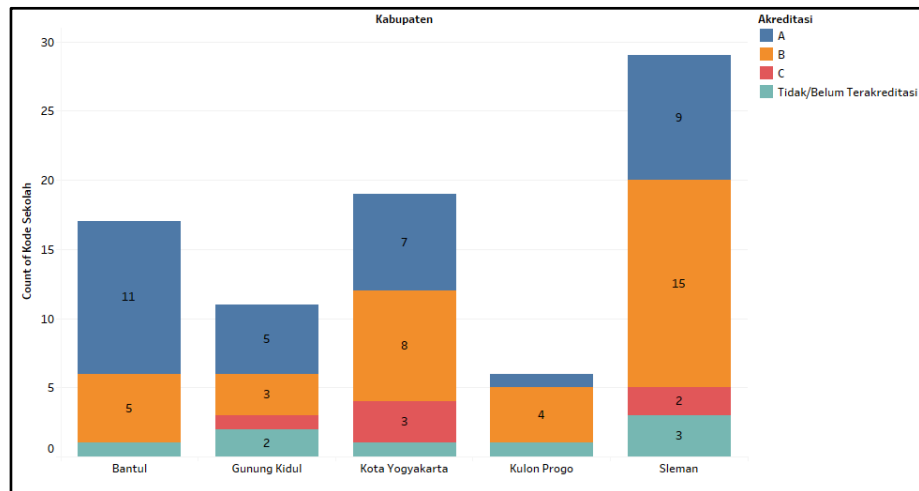
Gambar 5.37 Profil Klaster 1: WIFI

Berdasarkan *boxplot* pada **Gambar 5.36** didapatkan bahwa nilai rata – rata jumlah *switch hub* pada klaster 1 adalah 8 buah. Nilai minimum jumlah *switch hub* pada klaster 1 adalah 2 buah dan nilai maksimum adalah 85 buah. Nilai tengah (atau median) jumlah *switch hub* klaster 1 adalah 6 buah, nilai kuartil pertama (atau Q1) adalah 4 buah dan nilai kuartil ketiga (atau Q3) adalah 8 buah.

Berdasarkan *boxplot* pada **Gambar 5.37** didapatkan bahwa nilai rata – rata jumlah WIFI pada klaster 1 adalah 12 buah. Nilai minimum jumlah WIFI pada klaster 1 adalah 1 buah dan nilai maksimum adalah 71 buah. Kemudian untuk nilai tengah (atau median) jumlah WIFI pada klaster 1 adalah 7 buah, nilai kuartil pertama (atau Q1) adalah 4 buah dan nilai kuartil ketiga (atau Q3) adalah 17 buah.

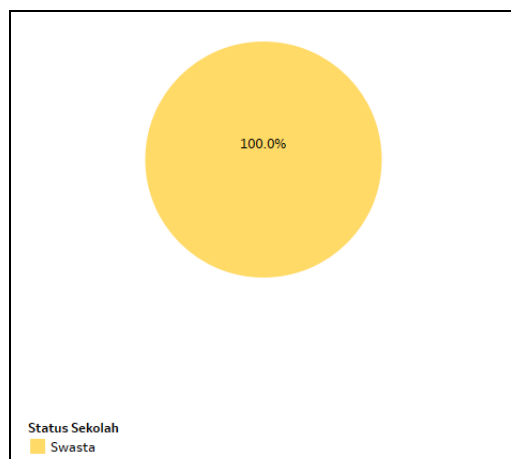
b. Karakteristik Klaster 2

Pada klaster 2 hasil pengelompokan menggunakan SKM diperoleh anggota klaster sebanyak 82 sekolah. Berikut adalah karakteristik variabel dari anggota pada klaster 2. Jumlah sekolah SMA di Provinsi Yogyakarta yang masuk dalam klaster 2 berdasarkan kabupaten dan akreditasi sekolah ditunjukkan pada **Gambar 5.38**.

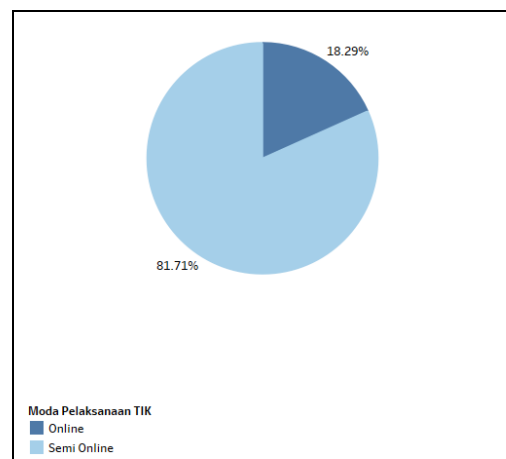


Gambar 5.38 Profil Klaster 2: Akreditasi Berdasarkan Wilayah

Pada **Gambar 5.38** menyajikan *stacked bar* banyaknya sekolah pada klaster 2 berdasarkan akreditasi dan wilayah sekolah. Sekolah yang masuk dalam anggota klaster 2 telah terakreditasi A, B, C dan tidak/ belum terakreditasi. Namun jika dilihat pada *stacked bar* sekolah dengan akreditasi A dan B lebih mendominasi pada klaster 2. Terlihat bahwa Kabupaten Bantul dan Kabupaten Sleman menjadi wilayah terbanyak sekolah yang terakreditasi A yaitu sebanyak 11 sekolah. Persebaran anggota klaster 2 meliputi seluruh kabupaten/kota di Provinsi Yogyakarta. Persentase status sekolah untuk klaster 2 disajikan pada **Gambar 5.39** sedangkan persentase moda pelaksanaan TIK disajikan pada **Gambar 5.40**.



Gambar 5.39 Profil Klaster 2:
Persentase Status Sekolah

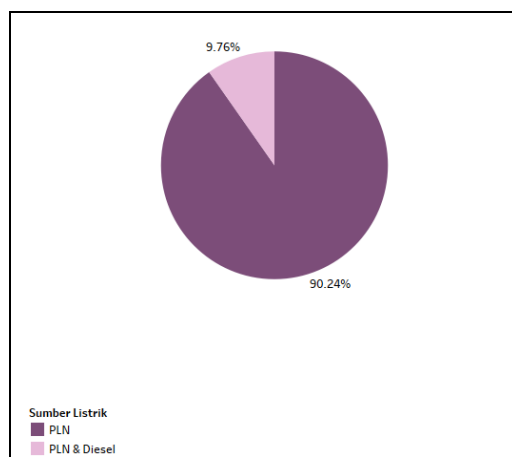


Gambar 5.40 Profil Klaster 2:

Persentase Moda Pelaksanaan TIK

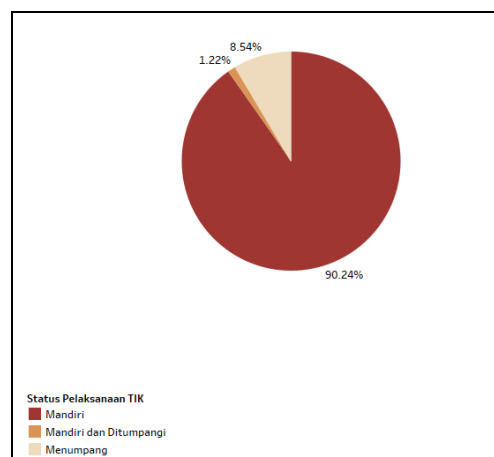
Berdasarkan *pie chart* pada **Gambar 5.39** menunjukkan bahwa pada klaster 2 seluruh anggota klaster merupakan sekolah swasta. Persentase untuk sekolah swasta pada klaster 2 sebesar 100% atau 82 sekolah.

Berdasarkan *pie chart* pada **Gambar 5.40** menunjukkan bahwa pada klaster 2, sekolah dengan moda pelaksanaan TIK secara semi online lebih mendominasi dengan persentase sebesar 81.71% atau 67 sekolah dan sisanya sebesar 18.29% atau 15 sekolah merupakan sekolah dengan menggunakan moda pelaksanaan TIK secara online. Persentase sumber listrik untuk klaster 2 disajikan pada **Gambar 5.41** sedangkan persentase sumber listrik disajikan pada **Gambar 5.42**.



Gambar 5.41 Profil Klaster 2:

Persentase Sumber Listrik

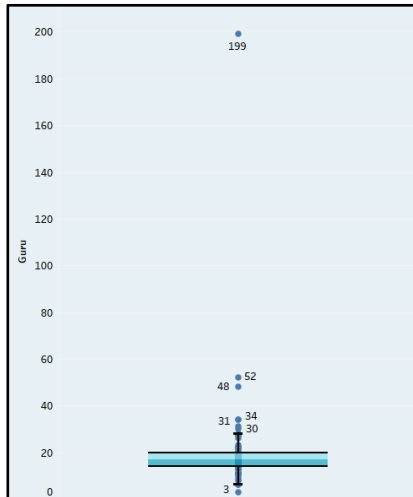


Gambar 5.42 Profil Klaster 2:

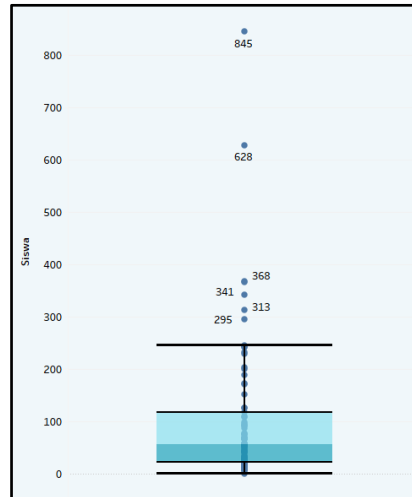
Persentase Status Pelaksanaan TIK

Berdasarkan *pie chart* pada **Gambar 5.41** menunjukkan bahwa pada klaster 2 berisikan sekolah yang memiliki sumber listrik berasal dari PLN sebesar 90.24% atau 74 sekolah serta sisanya sebesar 9.76% atau 8 sekolah merupakan sekolah yang menggunakan sumber listrik yang berasal dari PLN dan diesel.

Berdasarkan *pie chart* pada **Gambar 5.42** menunjukkan bahwa pada klaster 2, sekolah dengan status pelaksanaan TIK adalah mandiri lebih mendominasi dengan persentase sebesar 90.24% atau 74 sekolah. Selain itu, terdapat sekolah dengan status pelaksanaan TIK berupa mandiri dan ditumpangi serta menumpang. Masing – masing memiliki persentase sebesar 1.22% atau 1 sekolah dan 8.54% atau 7 sekolah. Karakteristik jumlah guru untuk klaster 2 disajikan pada **Gambar 5.43** sedangkan untuk jumlah siswa disajikan pada **Gambar 5.44**.



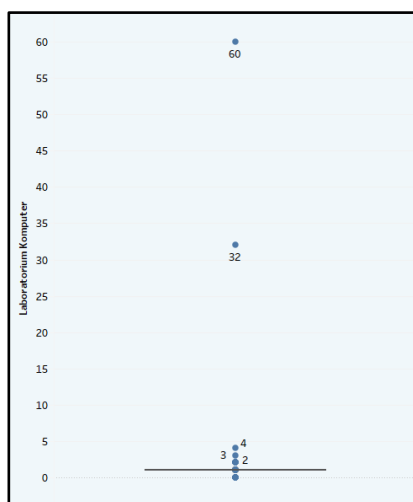
Gambar 5.43 Profil Klaster 2: Guru



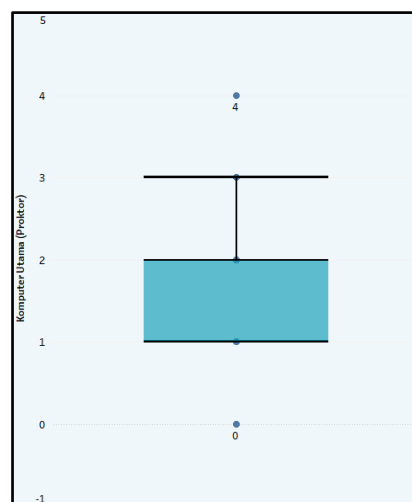
Gambar 5.44 Profil Klaster 2: Siswa

Berdasarkan *boxplot* pada **Gambar 5.43** didapatkan bahwa nilai rata – rata jumlah guru pada klaster 2 sebesar 21 guru. Nilai minimum jumlah guru pada klaster 2 sebesar 3 guru dan nilai maksimum sebesar 199 guru. Nilai tengah (atau median) jumlah guru klaster 2 sebesar 18 guru, nilai kuartil pertama (atau Q1) sebesar 14 guru dan nilai kuartil ketiga (atau Q3) sebesar 22 guru.

Berdasarkan *boxplot* pada **Gambar 5.44** didapatkan bahwa nilai rata – rata jumlah siswa klaster 2 sebesar 102 siswa dengan nilai minimum sebesar 0 dan nilai maksimum sebesar 845 siswa. Nilai tengah jumlah siswa klaster 2 sebesar 57 siswa, nilai Q1 sebesar 25 siswa dan nilai Q3 sebesar 116 siswa. Karakteristik jumlah laboratorium komputer klaster 2 disajikan pada **Gambar 5.45** sedangkan untuk jumlah komputer utama / proktor disajikan pada **Gambar 5.46**.



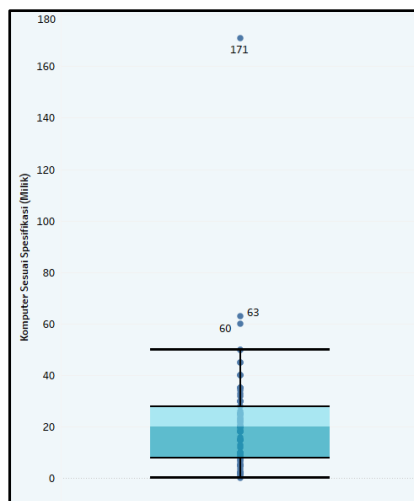
Gambar 5.45 Profil Klaster 2:
Laboratorium Komputer



Gambar 5.46 Profil Klaster 2:
Komputer Utama

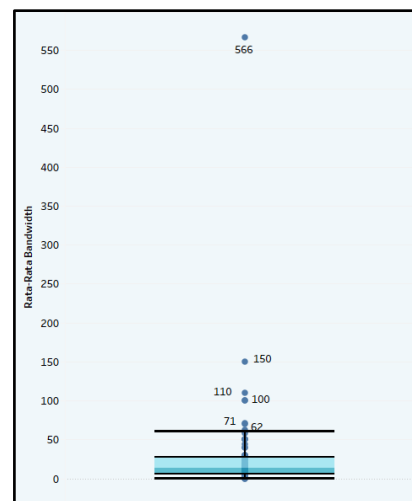
Berdasarkan *boxplot* pada **Gambar 5.45** didapatkan bahwa nilai rata – rata jumlah laboratorium komputer pada klaster 2 adalah 3 ruang dengan nilai minimum adalah 0 ruang dan nilai maksimum adalah 60 ruang. Nilai tengah (atau median) jumlah laboratorium komputer pada klaster 2 adalah 1 ruang, nilai kuartil pertama (atau Q1) adalah 1 ruang dan nilai kuartil ketiga (atau Q3) adalah 1 ruang.

Berdasarkan *boxplot* pada **Gambar 5.46** didapatkan bahwa nilai rata – rata jumlah komputer utama pada klaster 2 adalah 3 unit dengan nilai minimum adalah 0 unit dan nilai maksimum adalah 50 unit. Nilai tengah jumlah komputer utama klaster 2 adalah 2 unit, nilai Q1 adalah 1 unit dan nilai Q3 adalah 2 unit. Karakteristik jumlah komputer sesuai spesifikasi (milik) klaster 2 disajikan pada **Gambar 5.47** sedangkan jumlah *switch hub* disajikan pada **Gambar 5.48**.



Gambar 5.47 Profil Klaster 2:

Komputer Sesuai Spesifikasi (Milik)



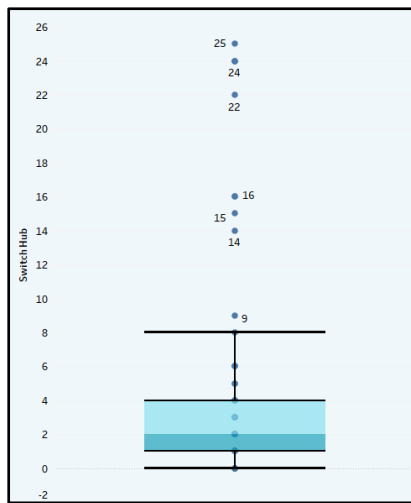
Gambar 5.48 Profil Klaster 2: Rata-

Rata *Bandwidth*

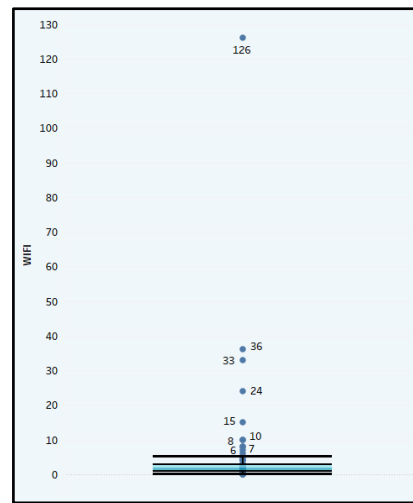
Berdasarkan *boxplot* pada **Gambar 5.47** didapatkan nilai rata – rata jumlah komputer sesuai spesifikasi (milik) klaster 2 adalah 23 unit dengan nilai minimum adalah 0 unit dan nilai maksimum 171 unit. Nilai tengah jumlah komputer sesuai spesifikasi (milik) klaster 2 adalah 20 unit, nilai Q1 adalah 9 unit dan nilai Q3 adalah 30 unit.

Berdasarkan *boxplot* pada **Gambar 5.48** didapatkan bahwa nilai *mean* rata-rata *bandwidth* pada klaster 2 sebesar 30 Mbps. Nilai minimum rata-rata *bandwidth* pada kaster 2 sebesar 0 Mbps dan nilai maksimum sebesar 566 Mbps. Nilai tengah rata-rata *bandwidth* sebesar 13 Mbps, nilai Q1 sebesar 7 Mbps dan nilai Q3 sebesar

28 Mbps. Karakteristik jumlah *switch hub* klaster 2 disajikan pada **Gambar 5.49** dan jumlah WIFI disajikan pada **Gambar 5.50**.



Gambar 5.49 Profil Klaster 2: *Switch Hub*



Gambar 5.50 Profil Klaster 2: WIFI

Berdasarkan *boxplot* pada **Gambar 5.49** didapatkan bahwa nilai rata – rata jumlah *switch hub* pada klaster 2 adalah 5 buah. Nilai minimum jumlah *switch hub* klaster 2 adalah 0 buah dan nilai maksimum adalah 25 buah. Nilai tengah (atau median) jumlah *switch hub* pada klaster 2 adalah 2 buah, nilai kuartil pertama (atau Q1) adalah 1 buah dan nilai kuartil ketiga (atau Q3) adalah 4 buah.

Berdasarkan *boxplot* pada **Gambar 5.50** didapatkan bahwa nilai rata – rata jumlah WIFI pada klaster 2 adalah 5 buah dengan nilai minimum 0 buah dan nilai maksimum adalah 126 buah. Nilai median jumlah WIFI klaster 2 adalah 2 buah, nilai Q1 adalah 1 buah dan nilai Q3 adalah 3 buah.

Untuk mengetahui perbedaan antar klaster hasil dari pengelompokan dengan metode SKM maka akan dilakukan menggunakan perhitungan rata – rata nilai dari setiap variabel numerik. Hasil rata – rata nilai tersebut disajikan pada **Tabel 5.22**.

Tabel 5.22 Profilisasi Anggota Klaster Berdasarkan Variabel Numerik

Klaster	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	40.5	553	2.45	3.82	81.2	63	7.53	11.1
2	20.5	102	2.18	1.57	21.7	29.6	4.26	4.99

Apabila dilihat pada **Tabel 5.22** baris yang diberi warna merah tua merupakan kelompok yang mempunyai nilai variabel numerik kategori tinggi. Sedangkan

untuk warna merah muda merupakan kelompok dengan variabel numerik kategori rendah. Berdasarkan penjelasan tersebut maka dapat dikatakan bahwa:

- Klaster 1 merupakan kelompok kategori tinggi, di mana pada jumlah SDM dan infrastruktur komponen kesiapan TIK untuk ANBK memiliki nilai yang tinggi apabila dibandingkan dengan klaster 2.
- Klaster 2 merupakan kelompok kategori rendah, di mana pada jumlah SDM dan infrastruktur komponen kesiapan TIK untuk ANBK memiliki nilai yang tinggi apabila dibandingkan dengan klaster 1.

Penjelasan mengenai karakteristik keanggotaan dari masing – masing klaster dapat dilihat pada **Tabel 5.23**.

Tabel 5.23 Karakteristik Klaster Hasil Pengelompokan

Klaster	Anggota Klaster	Karakteristik
1 (83 Anggota)	S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14, S15, S16, S17, S18, S19, S37, S38, S39, S40, S41, S42, S43, S44, S45, S46, S47, S48, S49, S50, S51, S52, S53, S68, S84, S85, S86, S87, S88, S89, S90, S91, S92, S93, S94, S95, S96, S103, S109, S110, S111, S112, S113, S114, S115, S116, S117, S118, S119, S126, S127, S128, S129, S130, S131, S132, S133, S134, S135, S136, S137, S139, S140, S143, S150, S157, S159, S160, S162, S163	Pada klaster 1, lebih didominasi oleh sekolah SMA dengan status negeri yang telah terakreditasi A dengan moda pelaksanaan TIK adalah semi online, status pelaksanaan TIK adalah mandiri dan ditumpangi, serta sumber listrik berasal dari PLN. Jika dilihat pada nilai rata – rata, klaster 1 memiliki nilai yang lebih tinggi pada jumlah SDM dan infrastruktur komponen kesiapan TIK untuk ANBK dibandingkan dengan klaster 2.
2 (82 Anggota)	S20, S21, S22, S23, S24, S25, S26, S27, S28, S29, S30, S31, S32, S33, S34, S35, S36, S54, S55, S56, S57, S58, S59, S60, S61, S62, S63, S64, S65, S66, S67, S69, S70, S71, S72, S73, S74, S75, S76, S77, S78, S79, S80, S81, S82, S83, S97, S98, S99, S100, S101, S102, S104, S105, S106, S107, S108, S120, S121, S122, S123, S124, S125, S138, S141, S142, S144, S145, S146, S147, S148, S149, S151, S152, S153, S154, S155, S156, S158, S161, S164, S165	Pada klaster 2, lebih didominasi oleh sekolah SMA dengan status swasta yang telah terakreditasi A dan B dengan moda pelaksanaan TIK adalah semi online, status pelaksanaan TIK adalah mandiri, serta sumber listrik berasal dari PLN. Jika dilihat pada nilai rata – rata, klaster 2 memiliki nilai yang lebih rendah pada jumlah SDM dan infrastruktur komponen kesiapan TIK untuk ANBK dibandingkan dengan klaster 1.

5.5. Pengujian Rata – Rata Dua Populasi Independen

Pengujian rata-rata dua populasi independen digunakan untuk mengetahui ada atau tidak adanya perbedaan yang signifikan pada nilai rata – rata variabel numerik antara klaster 1 dengan klaster 2. Pengujian ini dilakukan dengan menggunakan statistik uji Z karena jumlah sampel yang besar dan diasumsikan normal dengan menggunakan pendekatan pada Teori Limit Pusat (Long, 2020). Adapun hasil pengujian rata – rata dua populasi independen tersebut disajikan pada **Tabel 5.24**.

Tabel 5.24 Statistik Uji: Pengujian Rata – Rata Dua Populasi Independen
Variabel Numerik Penelitian

Variabel	Nilai Z_{hitung}
X_1 (Guru)	7.4087
X_2 (Siswa)	13.811
X_3 (Laboratorium Komputer)	0.32395
X_4 (Komputer Utama / Proktor)	3.37781
X_5 (Komputer Sesuai Spesifikasi (Milik))	13.182
X_6 (Rata – Rata <i>Bandwidth</i>)	3.1875
X_7 (<i>Switch</i> Hub)	2.5919
X_8 (WIFI)	2.9504

Nilai $Z_{\alpha/2}$ sebesar 1.96 dengan keputusan akan tolak H_0 apabila nilai $Z_{hitung} > Z_{\alpha/2}$. Berdasarkan hasil pengujian dua rata – rata populasi independen pada **Tabel 5.24**, didapatkan bahwa hanya terdapat satu variabel numerik yang tidak memiliki perbedaan yang signifikan antara klaster 1 dan klaster 2. Variabel tersebut adalah variabel X_3 atau jumlah laboratorium komputer. Apabila dilihat dari rata – rata variabel numerik pada jumlah laboratorium untuk klaster 1 dan klaster 2 memiliki nilai yang hampir sama yaitu jika dibulatkan jumlah laboratorium pada klaster 1 dan klaster 2 adalah 3 ruang.

Selain melakukan pengujian pada variabel numerik yang digunakan dalam penelitian ini, peneliti ingin mengetahui apakah variabel numerik yang tidak masuk dalam penelitian memiliki perbedaan yang signifikan apabila dimasukkan pada hasil pengelompokan untuk klaster 1 dan klaster 2. Adapun beberapa variabel numerik non penelitian yang akan diuji adalah variabel daya listrik, *bandwidth upload* dan *bandwidth download*. Berikut hasil pengujian rata – rata dua populasi independen untuk variabel numerik non penelitian yang disajikan pada **Tabel 5.25**.

Tabel 5.25 Statistik Uji: Pengujian Rata – Rata Dua Populasi Independen
Variabel Numerik Non Penelitian

Variabel	Nilai Z_{hitung}
Daya Listrik	2.2328
<i>Bandwidth Upload</i>	2.9302
<i>Bandwidth Download</i>	3.3635

Pada hasil pengujian hipotesis rata – rata dua populasi independen untuk variabel numerik non penelitian diperoleh bahwa ketiga variabel numerik yaitu daya listrik, *bandwidth upload* dan *bandwidth download* memiliki perbedaan yang signifikan antara kluster 1 dan kluster 2. Karena nilai Z_{hitung} ketiga variabel tersebut memiliki nilai yang lebih besar dari $Z_{\alpha/2}$ yaitu sebesar 1.96.

5.6. Validasi Metode

Validasi metode dilakukan dalam menentukan metode pengelompokan terbaik untuk data sekolah SMA di Provinsi Yogyakarta berdasarkan komponen kesiapan TIK untuk ANBK. Proses validasi metode ini, akan dilakukan dengan menggunakan tiga metode pengelompokan non-hirarki berbasis medoid untuk data dengan variabel campuran. Ketiga metode tersebut meliputi algoritma *K-Medoids* (KM), *Simple and Fast K-Medoids* (SFKM) dan *Simple K-Medoids* (SKM).

Proses pengelompokan menggunakan metode KM dilakukan dengan jumlah iterasi maksimum sebanyak 10, dimana *initial* objek yang diperoleh secara random adalah objek ke-12 untuk kluster 1 dan objek ke-135 untuk kluster 2. Selanjutnya, pada pengelompokan dengan metode SFKM menggunakan jumlah iterasi maksimum sebanyak 10. Pada pengelompokan metode SKM dilakukan dengan menggunakan jumlah *seeding* sebanyak 20 kali dimana setiap 1 kali *seeding* memiliki jumlah iterasi maksimum sebanyak 10 kali. Ketiga metode pengelompokan yaitu KM, SFKM dan SKM akan menggunakan perhitungan jarak untuk data campuran adalah jarak *Gower* dimana data numerik menggunakan rentang terbobot *manhattan* serta data biner dan kategorik menggunakan *simple matching*. Adapun hasil pengelompokan dengan menggunakan ketiga metode tersebut disajikan pada **Tabel 5.26**.

Tabel 5.26 Validasi Metode: KM, SFKM dan SKM

Metode	Nilai <i>Minimum Distance (E)</i>	Medoid Akhir		Jumlah Anggota Klaster	
		Klaster 1	Klaster 2	Klaster 1	Klaster 2
<i>K-Medoids</i> (KM)	14.00341	25	136	87	78
<i>Simple and Fast K-Medoids</i> (SFKM)	14.00341	136	25	78	87
<i>Simple K-Medoids</i> (SKM)	13.98689	19	71	83	82

Penentuan metode terbaik didasarkan ketika jumlah k klaster yang terbentuk sama maka anggota yang terbentuk sama atau relatif sama. Pada **Tabel 5.26** menunjukkan bahwa metode terbaik untuk data sekolah SMA di Provinsi Yogyakarta berdasarkan komponen kesiapan TIK untuk ANBK adalah metode *Simple K-Medoids* (SKM) dengan nilai *minimum distance* (E) paling kecil. Pengelompokan dengan menggunakan algoritma SKM memberikan hasil bahwa nilai *minimum distance* (E) sebesar 13.98689 dimana medoid akhir klaster 1 adalah objek ke-19 dan klaster 2 adalah objek ke-71. Jumlah anggota klaster yang terbentuk untuk klaster 1 sebanyak 83 anggota sedangkan klaster 2 sebanyak 82 anggota. Selain itu, adanya metode SKM dikembangkan untuk mengatasi kekurangan dari metode KM dan SFKM yang rentan terjadi permasalahan pada *local optima* dan klaster kosong.

BAB VI

PENUTUP

6.1. Kesimpulan

Berdasarkan rumusan masalah pada bab sebelumnya, didapatkan kesimpulan untuk menjawab rumusan masalah yang ada dengan penjelasan sebagai berikut:

1. Gambaran umum komponen kesiapan TIK sekolah SMA di Provinsi Yogyakarta dalam penyelenggaraan ANBK 2021 yaitu rata – rata sekolah SMA memiliki jumlah guru sebanyak 31 orang, siswa sebanyak 329 orang, laboratorium komputer ada 3 ruang, komputer utama ada 4 unit, komputer sesuai spesifikasi (milik) ada 52 unit, rata-rata *bandwidth* sebesar 46.38 Mbps, *switch hub* ada 6 buah dan WIFI ada 9 buah serta didominasi oleh sekolah SMA swasta yang terakreditasi A dengan status pelaksanaan TIK berupa mandiri serta mandiri dan ditumpangi, moda pelaksanaan TIK berupa semi online dan sumber listrik berasal dari PLN.
2. Pengelompokan komponen kesiapan TIK sekolah SMA di Provinsi Yogyakarta dalam penyelenggaraan ANBK 2021 menggunakan algoritma *Simple K-Medoids Partitioning* (SKM) dengan *set.seed(1)*, *seeding* = 20 dan *iterate* = 10 dikelompokkan dalam 2 klaster. Klaster 1 terdiri dari 83 sekolah sedangkan klaster 2 terdiri dari 82 sekolah. Sedangkan untuk hasil validasi menggunakan *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV) menunjukkan bahwa jumlah *k* klaster optimal pada pengelompokan SMA di Provinsi Yogyakarta berdasarkan komponen kesiapan TIK untuk ANBK adalah 2 klaster.
3. Karakteristik yang terbentuk dari hasil pengelompokan SMA menunjukkan berdasarkan data numerik yaitu klaster 1 merupakan kelompok sekolah dengan jumlah SDM dan infrastruktur pada komponen kesiapan TIK untuk ANBK yang tinggi sedangkan untuk klaster 2 menjadi kelompok sekolah dengan jumlah SDM dan infrastruktur pada komponen kesiapan TIK untuk ANBK yang rendah. Sedangkan karakteristik klaster berdasarkan data kategorik yaitu klaster 1 didominasi oleh sekolah negeri yang terakreditasi A, dengan moda pelaksanaan TIK semi online, sumber listrik berasal dari PLN,

serta status pelaksanaan TIK berupa mandiri dan ditumpangi. Karakteristik klaster 2 didominasi oleh sekolah swasta yang terakreditasi A dan B, dengan moda pelaksanaan TIK semi online, sumber listrik berasal dari PLN, dan status pelaksanaan TIK berupa mandiri.

6.2. Saran

Berdasarkan hasil dari analisis yang telah dilakukan dalam penelitian ini, saran yang dapat diambil adalah Pemerintah melalui Kementerian Pendidikan dan Kebudayaan (Kemendikbud) agar lebih memperhatikan sekolah-sekolah yang masih memiliki fasilitas yang rendah terutama yang berkaitan dengan sarana dan prasarana Teknologi Informasi dan Komunikasi (TIK) serta kualitas dan kuantitas dari segi SDM.

Saran selanjutnya apabila penelitian ini akan dikembangkan, peneliti menyarankan yaitu dengan membandingkan metode pengelompokan yang lain untuk membandingkan hasil antara metode yang digunakan, tidak membatasi hanya pada jenjang pendidikan tertentu, menambah variabel pendukung lainnya yang masuk dalam profil kesiapan TIK, serta pada proses validasi dapat menggunakan metode pengukuran yang lain untuk mendukung keakuratan hasil analisis yang telah dilakukan.

DAFTAR PUSTAKA

- Asmani, J. M. (2011). *Tips Praktis Membangun dan Mengolah Administrasi Sekolah*. Yogyakarta: Diva Press.
- Ayang Sari, d. (2020). Penghapusan Ujian Nasional Tahun 2021 dalam Perspektif Guru SMA di Kota Tebing Tinggi. *Prodising Seminar Nasional PBSI-III* (hal. 213 - 220). Medan: Universitas Negeri Medan.
- Budiaji, W. (2019). *Partitioning Clustering Algorithms in Food Security Mapping*. Vienna: University of Natural Resources and Life Sciences Vienna .
- Budiaji, Weksi. (2019, Juli). Medoid-Based Shadow Value Validation And Visualization. *International Journal of Advances in Intelligent Informatics*, 5(2), 76 - 88. Diambil kembali dari <http://dx.doi.org/10.26555/ijain.v5i2.326>
- Dani Hermawan, M. (2020). *Diktat Statistika Pendidikan dengan SPSS*. Jember: IAIN Jember.
- Hae-Sang Park, C.-H. J. (2009). A Simple and Fast Algorithm for K-Medoids Clustering. *ELSEVIER*, 3336 - 3341.
- Halkidi, M. Y. (2001). On Clustering Validation Techniques. *Journal of Intelligent Information Systems*, 107 - 145.
- Hariani, N. P. (2021). *Pengelompokan Perguruan Tinggi Negeri Di Indonesia Menggunakan Metode Hirarki Dan K-Medoids Dengan Ukuran Jarak Modifikasi Data Campuran (Studi Kasus : Perguruan Tinggi Negeri Di Indonesia Tahun 2020)*. Yogyakarta: Skripsi Program Sarjana Statistika, Universitas Islam Indonesia.
- Hermawan, D. (2020). *Diktat Statistika Pendidikan dengan SPSS*. Jember: IAIN Jember.
- Hidayat, A. (2016). *Penjelasan Tentang Analisis Multivariat dan Jenisnya*. Dipetik Januari 8, 2022, dari Statistikian: <https://www.statistikian.com/2016/11/analisis-multivariat.html>
- Huang, Z. (1997). Clustering Large Data Sets With Mixed Numeric And Categorical Values. *The First Pacific-Asia Conference on Knowledge Discovery and Data Mining*, (hal. 21 - 34).

- Ikasakti, N. A. (2017). *Efektivitas Algoritma K-Means, K-Medoids, Dan Optimal Jumlah Cluster Untuk Modifikasi Algoritma K-Medoids Pada Kumpulan Data Besar*. Yogyakarta: Universitas Islam Indonesia.
- Insani, R. W. (2020). *Memahami Teorema Limit Pusat*. Dipetik Januari 15, 2022, dari rachmatwahid: <https://www.rachmatwahid.com/2020/04/teorema-limit-pusat.html>
- Ivana Indriani Putri Damanik, d. (2019). Algoritma K-Medoids untuk Mengelompokkan Desa yang Memiliki Fasilitas Sekolah di Indonesia. *Prosiding Seminar Nasional Riset Information Science (SENARIS)* (hal. 520 - 527). Pematangsiantar: STIKOM Tunas Bangsa.
- J. Han, M. K. (2006). *Data Mining: Concept and Techniques*. Waltham: Morgan Kauffman Publisher.
- Johnson, R. &. (2007). *Applied Multivariate Statistical Analysis (6th)*. New Jersey: Pearson Education.
- Kaufman, R. d. (1990). *Finding Groups in Data*. New York, USA: John Wiley & Sons.
- Kemendikbud. (2021). *Penyelenggaraan Asesmen Nasional*. Balitbang dan Perbukuan.
- Kemendikbud Ristek. (2020). *Kesiapan TIK*. Dipetik Januari 10, 2022, dari Vervaltik Data Kemdikbud: <https://vervaltik.data.kemdikbud.go.id/index.php/publik/keterangan>
- Kemendikbudristek, P. (2021). *Menilik Kesiapan Asesmen Nasional dari Kacamata Kecakapan Digital*. Dipetik Januari 28, 2022, dari Pusat Data dan Teknologi Informasi Kementerian Pendidikan dan Kebudayaan: <https://pusdatin.kemdikbud.go.id/menilik-kesiapan-asesmen-nasional-dari-kacamata-kecakapan-digital/>
- L, I. (2019). Evaluasi dalam Proses Pembelajaran. *Adaara (Jurnal Manajemen Pendidikan Islam)*, 920-935.
- Long, T. J. (2020). *Teorema Limit Pusat*. Dipetik Maret 1, 2022, dari Jagostat.com: <https://jagostat.com/statistika-matematika-1/teorema-limit-pusat>

- Long, Tju Ji. (2020). *Uji Hipotesis Rata - Rata Dua Populasi*. Dipetik Januari 20, 2022, dari jagostat.com: <https://jagostat.com/metode-statistika-2/uji-hipotesis-rata-rata-dua-populasi>
- Maharani, T. (2019). *Klarifikasi Nadiem Makarim: UN Tidak Dihapus, Hanya Diganti*. Dipetik Januari 12, 2022, dari KOMPAS.com: <https://nasional.kompas.com/read/2019/12/13/08224851/klarifikasi-nadiem-makarim-un-tidak-dihapus-hanya-diganti?page=all>
- Muhyidin. (2020). *Data & Variabel*. Dipetik Maret 5, 2022, dari Muhyidin SKM: <https://muhyidin.id/data-variabel/>
- Nooraeni, R. (2016). Metode Clustering Menggunakan Kombinasi Algoritma Cluster K-Prototype dan Algoritma Genetika untuk Data Bertipe Campuran. *Jurnal Aplikasi Statistika & Komputasi Statistik*, 81 - 97.
- P. IndiraPriya, D. D. (2013). A Survey on Different Clustering Algorithms in Data Mining Technique. *International Journal of Modern Engineering Research (IJMER)*, 267 - 274.
- Pendidikan, D. (2021). *Verifikasi dan Validasi TIK Satuan Pendidikan*. Jakarta: Kementerian Pendidikan dan Kebudayaan.
- Puspitasari, V. (2016). *Pengelompokan Sekolah Menengah Atas Di Provinsi Daerah Istimewa Yogyakarta Berdasarkan Nilai Ujian Nasional Menggunakan Algoritma Agglomerative Hierarchical Clustering*. Yogyakarta: Universitas Sanata Dharma.
- Putri, I. P. (2017). *Pengelompokan Perguruan Tinggi Negeri di Indonesia Menggunakan Metode Ensemble Robust Clustering Using Links*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Rakyat, D. S. (2021). *Statistik Pendidikan*. Jakarta: Badan Pusat Statistik.
- Riyanto, B. (2019). Penerapan Algoritma K-Medoids Clustering untuk Pengelompokan Penyebaran Diare di Kota Medan (Studi Kasus: Kantor Dinas Kesehatan Kota Medan). *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 3(1), 562 - 568. doi:10.30865/komik.v3i1.1659
- Rousseeuw, P. (1984). Least Median of Squares Regression. *Journal of the American Statistical Association*, 871 - 880.

- Sandhya Harikumar, S. P. (2015). K-Medoid Clustering for Heterogeneous DataSets. *4th International Conference on Eco-Friendly Computing and Communication Systems (ICECCS)* (hal. 226 - 237). Elsevier. doi:10.1016/j.procs.2015.10.077
- Sharma, S. (1996). *Applied Multivariate Techniques*. New York: John Wiley and Sons, Inc.
- SMP, A. (2021). *Ketahui Perbedaan ANBK Online dan Semi-Online*. Dipetik Februari 2, 2022, dari Direktorat Sekolah Menengah Pertama: <https://ditsmp.kemdikbud.go.id/ketahui-perbedaan-anbk-online-dan-semi-online/>
- Syafril. (2010). *Statistika*. Padang: Sukabina Press.
- Umi Wahyuningsih Muhadi, d. (2017). *Sekolah Menengah Atas Dari Masa ke Masa*. Jakarta Selatan: Direktorat Pembinaan SMA, Kementerian Pendidikan dan Kebudayaan.
- Weksi Budiaji, F. L. (2019). Simple K-Medoids Partitioning Algorithm for Mixed Variable Data. *Algorithms*.
- Wikipedia. (2018). *Sekolah Swasta*. Dipetik Februari 15, 2022, dari Wikipedia: https://id.wikipedia.org/wiki/Sekolah_swasta
- Wikipedia. (2021). *Sekolah Negeri (Pemerintah)*. Dipetik Januari 28, 2022, dari Wikipedia: [https://id.wikipedia.org/wiki/Sekolah_negeri_\(pemerintah\)](https://id.wikipedia.org/wiki/Sekolah_negeri_(pemerintah))
- Yuvalianda. (2019). *Statistika Deskriptif : Pembahasan Lengkap dan Contoh*. Dipetik Februari 20, 2022, dari Blog Yuva: <https://yuvalianda.com/statistik-deskriptif/>

RINGKASAN TA

IMPLEMENTASI *SIMPLE K-MEDOIDS PARTITIONING* PADA PENGELOMPOKAN SMA BERDASARKAN KOMPONEN KESIAPAN TEKNOLOGI INFORMASI (Studi Kasus : Sekolah Menengah Atas (SMA) di Provinsi Yogyakarta)

SYAFIRA IRSALINA^{1*}, KARIYAM²

^{1,2} Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia

*18611105@students.uii.ac.id

ABSTRAK

Evaluasi pendidikan menjadi hal yang tidak dapat dipisahkan dengan komponen pendidikan. Secara nasional, evaluasi dilakukan dengan sistem Ujian Nasional Berbasis (UNBK). Namun, pada tahun 2021, UNBK diganti menjadi Asesmen Nasional Berbasis Komputer (ANBK). Provinsi Yogyakarta memiliki tingkat penyelesaian sekolah jenjang SM/ sederajat tertinggi dan tingkat kesenjangan dalam kecakapan digital paling rendah. Persiapan Teknologi Informasi dan Komunikasi (TIK) menjadi hal penting pada pelaksanaan ANBK. Tujuan penelitian yakni mengelompokkan sekolah SMA di Provinsi Yogyakarta berdasarkan data komponen kesiapan TIK untuk ANBK menggunakan algoritma *Simple K-Medoids Partitioning* (SKM) dengan jarak *Gower*, validasi kluster menggunakan *Silhouette Index* dan *Medoid Shadow Value* serta karakteristik kelompok yang terbentuk dari hasil pengelompokan. Data berasal dari website Sekolah Kita milik Kemendikbud yang merupakan data campuran yakni terdiri dari 8 variabel numerik dan 5 variabel kategorik. Data dikelompokkan menjadi dua kelompok. Hasil pengelompokan dengan *seeding* = 20 dan *iterate* = 10 menunjukkan bahwa kluster 1 terdiri dari 83 sekolah yang merupakan kelompok dengan jumlah SDM dan infrastruktur TIK yang tinggi dan kluster 2 terdiri dari 82 sekolah yang merupakan kelompok dengan jumlah SDM dan infrastruktur TIK yang rendah serta validasi kluster diperoleh jumlah k kluster yang optimal adalah ketika $k = 2$.

Kata Kunci: Komponen Kesiapan TIK untuk ANBK, SMA, *Simple K-Medoids Partitioning*

ABSTRACT

Educational evaluation is something that cannot be separated from the education component. Nationally, the evaluation is carried out using the National Based Examination (UNBK) system. However, in 2021, UNBK was changed to the Computer-Based National Assessment (ANBK). Yogyakarta province has the highest secondary school completion rate and the lowest disparity in digital skills. Preparation of Information and Communication Technology (ICT) is an important thing in the implementation of ANBK. The research objectives are to group high school schools in Yogyakarta Province based on the component data of ICT readiness for ANBK using the Simple K-Medoids Partitioning (SKM) algorithm with Gower distance, cluster validation using Silhouette Index and Medoid Shadow Value and the characteristics of the groups formed from the grouping results. The data comes from the Sekolah Kita website belonging to the Ministry of Education and Culture which is mixed data consisting of 8 numerical variables and 5 categorical variables. The data are grouped into two groups. The results of grouping with seeding = 20 and iterate = 10 indicate that cluster 1 consists of 83 schools which are a group with a high number of human resources and ICT infrastructure and cluster 2 consists of 82 schools which are a group with a low number of human resources and ICT infrastructure. In cluster validation, the optimal number of k clusters is when $k = 2$.

Keywords: ICT Readiness Components for ANBK, Senior High School, Simple K-Medoids Partitioning

1 Pendahuluan

Evaluasi pendidikan menjadi hal yang tidak terpisahkan dari komponen pendidikan. Proses evaluasi secara nasional dilakukan melalui Ujian Nasional (UN). Terkhusus sistem UN yang diterapkan pemerintah adalah Ujian Nasional Berbasis Komputer (UNBK). Tahun 2021, Menteri Pendidikan dan Kebudayaan, Nadiem Makarim resmi mengganti UNBK menjadi Asesmen Nasional Berbasis Komputer (ANBK). Asesmen Nasional terdiri dari tiga instrumen salah satunya yaitu Asesmen Kompetensi Minimum (AKM) Literasi-Numerasi (Kemendikbud, 2021). ANBK adalah bentuk penerapan ujian berbasiskan komputer dari Asesmen Nasional. Pelaksanaan ANBK hanya ditujukan untuk sampel siswa setiap jenjang. Jenjang SMA/Sederajat dilakukan siswa kelas 11 (Ayang Sari, 2020). Dengan penyelenggaraan ANBK yang berbasiskan komputer maka komponen kesiapan TIK menjadi suatu hal yang penting untuk menunjang terlaksananya ANBK. Pelaksanaan per jenjang secara bertahap memungkinkan sekolah untuk berbagi sumber daya. Sekolah yang belum mempunyai infrastruktur memadai dapat mengikuti ANBK di satuan pendidikan lain. Kemendikbud Ristek melaksanakan AKM dengan moda asesmen berbasis komputer secara online maupun semi online (Maharani, 2019).

Provinsi Yogyakarta menjadi provinsi pertama dengan tingkat penyelesaian pendidikan pada jenjang SM/Sederajat yang tertinggi yaitu dengan persentase sebesar 90.12% (Rakyat, 2021). Selain itu, Provinsi Yogyakarta memiliki tingkat kesenjangan dalam kecakapan digital paling rendah dibandingkan dengan provinsi lain karena di wilayah Yogyakarta mempunyai infrastruktur *wireless* terbesar setelah provinsi DKI Jakarta. Sehingga kemampuan untuk mengakses TIK juga besar (Kemendikbudristek, 2021). Kesiapan sekolah dalam bidang penggunaan teknologi untuk melaksanakan ANBK menjadi tolak ukur keberhasilan ANBK. Berdasarkan data Dapodik tahun 2021 mencatat ada sebanyak 169 sekolah SMA

dengan 4.281 guru sebagai tenaga pendidikan dan 50.032 siswa sebagai peserta didik di Provinsi Yogyakarta.

Analisis kluster adalah alat eksplorasi penting dalam suatu penyelidikan struktur data. Setiap objek yang berada dalam suatu kelompok bersifat homogen satu dengan lainnya, dan objek yang berada dalam kelompok lain bersifat heterogen satu sama lain (Kaufman, 1990). Salah satu algoritma non-hirarki *k-medoids* yang dikembangkan untuk data variabel campuran yaitu numerik, ordinal, biner dan kategorik adalah *Simple K-Medoids Partitioning* (SKM) (Weksi Budiaji, 2019). Algoritma ini menggunakan *medoid* sebagai pusat klasternya.

Tujuan penelitian ini untuk mengetahui gambaran umum data komponen kesiapan TIK untuk ANBK pada sekolah SMA di Provinsi Yogyakarta, hasil pengelompokan menggunakan metode *Simple K-Medoids Partitioning* (SKM) dengan perhitungan jarak *Gower*, validasi kluster menggunakan *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV) serta karakteristik kelompok yang terbentuk dari hasil pengelompokan. Penelitian ini diharapkan memiliki manfaat Kemendikbud dalam membantu dan memberikan bantuan perbaikan pada infrastruktur sekolah serta bermanfaat bagi sekolah SMA di Provinsi Yogyakarta agar meningkatkan kualitas sumber daya maupun fasilitas yang menunjang proses pembelajaran.

2 Tinjauan Pustaka

2.1 Analisis Kluster

Analisis kluster yakni alat eksplorasi penting dalam penyelidikan struktur data. Setiap objek yang berada dalam suatu kelompok bersifat homogen satu dengan lainnya, dan objek yang berada dalam kelompok lain bersifat heterogen satu sama lain (Kaufman, 1990). Pada pengelompokan berbasis jarak, kuantifikasi kesamaan antar objek didasarkan pada jaraknya, di mana benda yang lebih mirip mempunyai jarak yang lebih dekat dan sebaliknya. Setelah jarak didefinisikan, baik algoritma hirarki maupun non hirarki dapat diterapkan (Weksi Budiaji, 2019).

2.2 Ukuran Kedekatan Antar Objek

Secara umum, jarak digunakan dalam perhitungan ukuran kemiripan dan ketidakmiripan. Hal paling berpengaruh pada analisis pengelompokan adalah jarak antar objek (Sharma, 1996). Maka, perlu alat ukur untuk menghitung jarak antar objek. Terdapat tiga metode untuk mengukur jarak berdasarkan jenis variabel yaitu numerik, biner dan kategorik serta campuran.

a. Pengukuran Kedekatan Variabel Numerik

Perhitungan kedekatan antar objek akan menggunakan rentang *manhattan* yang terbobot apabila variabel yang diteliti merupakan variabel numerik. Ukuran kedekatan variabel numerik merupakan kemiripan antar objek yang dituliskan pada persamaan (1).

$$s_{ijl_{numerical}} = 1 - \frac{|x_{il} - x_{jl}|}{R_l} \quad (1)$$

dimana:

- $s_{ijl_{numerical}}$ = kemiripan antara objek ke- i dan j pada variabel numerik ke- l
- x_{il} = nilai objek ke- i pada variabel numerik ke- l
- x_{jl} = nilai objek ke- j pada variabel numerik ke- l
- R_l = *range* variabel numerik ke- l

b. Pengukuran Kedekatan Variabel Biner dan Kategorik

Pada variabel dengan dua kategori atau biner serta variabel dengan lebih dari dua kategori, perhitungan kedekatan antar objek menggunakan *simple matching*. Ukuran kedekatan variabel biner dan kategorik merupakan kemiripan antar objek ke-*i* dan ke-*j* yang dituliskan pada persamaan (2).

$$s_{ijl_{binary/categorical}} \in \{0,1\} \quad (2)$$

dimana:

$s_{ijl_{binary/categorical}} = 1$; jika objek ke-*i* dan ke-*j* memiliki nilai sama pada variabel biner/kategorik ke-*l*

$s_{ijl_{binary/categorical}} = 0$; jika objek ke-*i* dan ke-*j* memiliki nilai yang beda pada variabel biner/kategorik ke-*l*

c. Pengukuran Kedekatan Gower Variabel Campuran

Ukuran kedekatan atau kemiripan *Gower* pada variabel campuran antara objek ke-*i* dan objek ke-*j* dituliskan pada persamaan (3).

$$s_{ij} = \frac{s_{ij}^{numerical} + s_{ij}^{binary} + s_{ij}^{categorical}}{p_n + p_b + p_c} \quad (3)$$

dimana:

s_{ij} = kesamaan *Gower* objek ke-*i* dan ke-*j*

$s_{ij}^{numerical}$ = kesamaan numerik objek ke-*i* dan ke-*j*

s_{ij}^{binary} = kesamaan biner objek ke-*i* dan ke-*j*

$s_{ij}^{categorical}$ = kesamaan kategorik objek ke-*i* dan ke-*j*

p_n, p_b, p_c = jumlah variabel numerik, biner, kategorik

d. Pengukuran Jarak Gower Variabel Campuran

Kemiripan *Gower* (1971) dapat digunakan untuk menghitung data dengan variabel campuran yaitu variabel numerik, biner dan kategorik. Jarak *Gower* antara objek ke-*i* dan ke-*j* didefinisikan sebagai ukuran ketidakmiripan antar objek yang tulis pada persamaan (4).

$$d_{ij} = 1 - s_{ij} \cdot \omega_{ijl} \quad (4)$$

Pada jarak *Gower* mengatur kontribusi setiap variabel. Apabila terdapat data *missing* maka ω_{ijl} bernilai nol. Perhitungan jarak *Gower* mengasumsikan tidak adanya data *missing* dan setiap variabel memiliki kontribusi yang sama maka $\omega_{ijl} = 1$ (Weksi Budiaji, 2019).

2.3 Algoritma Simple K-Medoids Partitioning (SKM)

Algoritma SKM dikembangkan untuk mengelola *local* optima dan kluster kosong yang dapat terjadi dalam algoritma KM dan SFKM. Solusi untuk menghindari *local* optima, rekomendasi dari Steinley and Brusco (2009) menerapkan beberapa inisialisasi. Mempertahankan label saat *medoid* non unik diinisialisasi merupakan alternatif untuk mengatasi kluster yang kosong. Adapun tahapan dalam algoritma SKM adalah sebagai berikut (Weksi Budiaji, 2019):

1. Pilih satu set medoid awal M_k dengan memilih *k* objek dari X_n .

Medoid awal M_k adalah objek yang terletak paling sentral. Dikatakan medoid optimal ketika $k = 1$ merupakan objek yang terletak di pusat yang meminimalkan jumlah jarak menuju objek lain dengan persamaan (5).

$$a_i = \sum_{j=1}^n d_{ij} = \sum_{j=1}^n d_{ji}, \quad i = 1, 2, 3, \dots, n \quad (5)$$

a_i adalah jumlah jarak dari objek ke- i menuju objek yang lain. Apabila vektor a_i diurutkan, M_k pertama dipilih dari nilai yang terkecil.

2. Tetapkan label keanggotaan dari setiap objek $x_n \in X_n, l(x_n)$ menuju ke medoid M_k yang paling dekat.
3. Perbarui set medoid M_k baru dengan mempertahankan label kluster $l(x_n)$ yang tetap dengan persamaan (6).

$$m_p := \underset{p \in K}{\operatorname{argmin}} \sum_{n: l(x_n)=m_p} d(x_n, m_p), \quad p = 1, 2, 3, \dots, k \quad (6)$$

m_p = objek yang meminimalkan jumlah jarak antar objek dalam satu kluster.

2. Hitung jumlah jarak *sum of within cluster* (E) yang merupakan jarak antara objek dengan medoidnya dalam kluster yang sama, yang dituliskan dalam persamaan (7).

$$E = \sum_{p=1}^k \sum_{n: l(x_n)=m_p} d(x_n, m_p) \quad (7)$$

3. Ulangi langkah 2 – 4 hingga nilai E bernilai sama dengan nilai E sebelumnya atau himpunan medoid tidak berubah atau jumlah iterasi yang ditentukan telah terpenuhi.
4. Ulangi langkah 1 – 5, untuk memperoleh beberapa *seeding* secara acak dari medoid awal. Perhatikan bahwa objek yang berada pada pusat kluster akan selalu menjadi salah satu medoid awal. Penentuan medoid akhir dipilih berdasarkan nilai E paling minimum.
5. Tetapkan keanggotaan dari setiap objek menuju ke medoid akhir.

2.4 Validasi Kluster

Validasi kluster merupakan tahapan dalam melakukan evaluasi hasil pengelompokan yang diklasifikasikan menjadi tiga jenis yaitu kriteria eksternal, internal dan relatif (Weksi Budiaji, 2019). Kriteria internalnya yaitu *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV).

a. *Silhouette Index* (SI)

Silhouette Index (SI) yaitu ukuran validasi dalam pengujian keoptimalan suatu kluster yang terbentuk dari pengelompokan. Untuk menghitung nilai SI dari objek i , terdapat dua komponen yakni a_i dan b_i . Nilai a_i adalah jarak rata – rata antara objek ke- i dengan seluruh objek lainnya yang berada dalam satu kluster, sedangkan nilai b_i adalah hasil perhitungan jarak rata-rata objek ke- i terhadap seluruh objek yang berada dalam kluster terdekat yang kemudian diambil nilai terkecilnya. Berikut rumus perhitungan nilai SI (Puspitasari, 2016). Rumus menghitung nilai a_j^i dituliskan pada persamaan (8).

$$a_j^i = \frac{1}{m_j - 1} \sum_{\substack{r=1 \\ r \neq i}}^{m_j} d(x_i^j, x_r^j) \quad i = 1, 2, \dots, m_j \quad (8)$$

dimana:

- i = indeks data
- j = kluster
- m_j = banyaknya data dalam kluster ke- j
- $d(x_i^j, x_r^j)$ = jarak data ke- i dengan data ke- r dalam satu kluster ke- j
- a_j^i = jarak rata – rata antara objek ke- i dengan semua objek dalam satu kluster

Rumus menghitung nilai b_j^i dituliskan pada persamaan (9).

$$b_i^j = \min_{\substack{n=1, \dots, k \\ n \neq j}} \left\{ \frac{1}{m_n} \sum_{\substack{r=1 \\ r \neq i}}^{m_j} d(x_i^j, x_r^n) \right\} \quad i = 1, 2, \dots, m_j \quad (9)$$

dimana:

- n = klaster yang lain
- $d(x_i^j, x_r^n)$ = jarak data ke- i dengan data ke- r dalam satu klaster ke- j terhadap data ke- r dalam klaster n
- b_j^i = nilai terkecil dari jarak rata – rata antara objek ke- i dengan semua objek dalam klaster lain, tidak berada dalam satu klaster dengan data ke- i

Rumus menghitung nilai SI data ke- i dituliskan pada persamaan (10).

$$SI_i^j = \frac{b_i^j - a_i^j}{\max(a_i^j, b_i^j)} \quad (10)$$

SI_k^j merupakan nilai *Silhouette Index* dari data ke- i . Nilai SI berada pada rentang -1 sampai 1. Ketika nilai SI sebesar 1 menunjukkan bahwa klaster terpisah dengan baik dan sebaliknya jika nilai SI sebesar 0 maka klaster terpisah dengan buruk. Nilai Kaufman untuk struktur nilai SI disajikan pada Tabel 27.

Tabel 27 Nilai Nilai Kaufman Karakteristik *Silhouette Index*

<i>Silhouette Index</i>	Struktur
$0.7 < SC \leq 1$	Kuat
$0.5 < SC \leq 0.7$	Baik
$0.25 < SC \leq 0.5$	Lemah
$SC \leq 0.25$	Tidak Mempunyai Struktur

b. Medoid Shadow Value (MSV)

MSV yakni teknik validasi berbasis medoid yang meniru *Silhouette Index* serta karakter nilai bayangan berbasis *centroid* dikembangkan oleh (Budiaji, Weksi, 2019). Teknik ini menghasilkan bentuk yang hampir sama dengan SI. Rumus MSV ditulis pada persamaan (14).

$$msv = \frac{d(x, m'(x)) - d(x, m(x))}{d(x, m'(x))} \quad (11)$$

dimana:

- $d(x, m(x))$ = jarak objek ke- i dengan medoid terdekat pertama
- $d(x, m'(x))$ = jarak objek ke- i dengan medoid terdekat kedua

2.5 Pengujian Rata – Rata Dua Populasi Independen

Hasil pengelompokan akan diuji dengan membandingkan variabel numerik dari setiap klaster untuk mengetahui rata – rata variabel secara independen berbeda signifikan.

a. Teorema Limit Pusat

Teorema limit pusat (*Central Limit Theorem*) mengatakan apabila suatu distribusi sampel dari nilai *mean* akan membentuk distribusi normal jika ukuran sampel semakin besar (Insani, 2020).

b. Uji Rata – Rata Dua Sampel Independen

Uji rata-rata dua sampel bebas adalah salah satu uji statistik parametrik yang digunakan dalam melakukan perbandingan dua kelompok independen untuk mengetahui apakah rata – rata populasi secara statistik signifikan berbeda. Misalkan

dipunyai dua sampel acak independen masing – masing memiliki ukuran n_1 dan n_2 diambil dari populasi dengan rata – rata μ_1 dan μ_2 serta variansi σ_1^2 dan σ_2^2 dengan peubah acak berdistribusi normal standar (Hermawan, 2020).

Hipotesis:

$H_0 : \mu_1 = \mu_2$ (Tidak ada perbedaan yang signifikan pada rata – rata kelompok 1 dan 2)

$H_1 : \mu_1 \neq \mu_2$ (Ada perbedaan yang signifikan pada rata – rata kelompok 1 dan 2)

Statistik Uji:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (12)$$

dimana:

$\bar{X}_1, \bar{X}_2$ = rata – rata sampel 1 dan 2

n_1, n_2 = ukuran sampel 1 dan 2

σ_1^2, σ_2^2 = variansi sampel 1 dan 2

Daerah kritis: tolak H_0 apabila $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}}$.

3 Tinjauan Pustaka

3.1 Variabel Penelitian

Penelitian menggunakan data sekunder dari website Sekolah Kita milik Kemendikbud. Populasi penelitian yaitu seluruh SMA di Provinsi Yogyakarta sebanyak 169 sekolah. Sampel penelitian adalah sekolah SMA yang telah melengkapi dan memperbaiki data komponen kesiapan TIK untuk ANBK serta bukan sekolah dengan jenjang SPK (Satuan Pendidikan Kerjasama) SMA sebanyak 165 sekolah. Penjelasan variabel penelitian disajikan pada Tabel 28.

Tabel 28 Variabel dan Jenis Data Penelitian

Variabel	Keterangan	Jenis	Kategori dan/atau Satuan
X_1	Guru	Numerik	Orang
X_2	Siswa	Numerik	Orang
X_3	Laboratorium	Numerik	Ruang
	Komputer		
X_4	Komputer Utama / Proktor	Numerik	Unit
X_5	Komputer Sesuai Spesifikasi (Milik)	Numerik	Unit
X_6	Rata-Rata <i>Bandwidth</i>	Numerik	Mbps
X_7	<i>Switch Hub</i>	Numerik	Buah
X_8	WIFI	Numerik	Buah
X_9	Status Sekolah	Biner	0: Negeri; 1: Swasta
X_{10}	Moda Pelaksanaan TIK	Biner	0: Online; 1: Semi Online
X_{11}	Sumber Listrik	Biner	0: PLN; 1: PLN & Diesel
			0: Terakreditasi A;
X_{12}	Akreditasi Sekolah	Kategorik (Ordinal)	1: Terakreditasi B; 2: Terakreditasi C; 3: Tidak/ Belum Terakreditasi
X_{13}	Status Pelaksanaan TIK	Kategorik	0: Mandiri dan Ditumpangi; 1: Mandiri; 2: Menumpang

4 Hasil Dan Pembahasan

4.1 Gambaran Data Penelitian

Sebelum analisis pengelompokan diperlukan gambaran data komponen kesiapan TIK untuk ANBK pada SMA di Provinsi Yogyakarta. Data tersebut disajikan pada Tabel 29.

Tabel 29 Gambaran Data Penelitian

Kode	Nama Sekolah	X_1	X_2	X_3	...	X_{13}
S1	SMA Negeri 1 Pundong	42	608	3	...	Mandiri dan Ditumpangi
S2	SMA Negeri 3 Bantul	47	215	3	...	Mandiri dan Ditumpangi
S3	SMA Negeri 1 Bambanglipuro	45	609	2	...	Mandiri dan Ditumpangi
S4	SMA Negeri 1 Sedayu	61	879	4	...	Mandiri dan Ditumpangi
⋮	⋮	⋮	⋮	⋮	...	⋮
S165	SMAS IT-Bina Anak Sholeh	11	57	1	...	Mandiri

4.3 Pengukuran Jarak Gower Objek Variabel Campuran

Pada pengukuran kedekatan antar objek dengan variabel campuran menggunakan jarak *Gower*. Pada data numerik menggunakan rentang terbobot *Manhattan* serta biner dan kategorik menggunakan *simple matching*. Jarak data campuran disajikan pada Tabel 30.

Tabel 30 Matriks Jarak Data Campuran

	S1	S2	S3	S4	S5	...	S165
S1	0	0.08	0.03	0.04	0.04	...	0.35
S2	0.08	0	0.10	0.12	0.07	...	0.36
S3	0.03	0.10	0	0.06	0.05	...	0.36
S4	0.04	0.12	0.06	0	0.07	...	0.38
S5	0.04	0.07	0.05	0.07	0	...	0.31
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
S165	0.37	0.34	0.38	0.41	0.34	...	0

Karena jarak data campuran merupakan ukuran ketidakmiripan objek. Jika jarak bernilai 1 artinya kedekatan dua objek lemah dan jika jarak bernilai 0 maka kedekatan sangat kuat.

4.4 Pengelompokan Menggunakan Simple K-Medoids Partitioning

Pada penelitian ini, pengelompokan SMA di Provinsi Yogyakarta akan dibagi dalam 2 kelompok menggunakan *software R Studio*. Pengelompokan dilakukan dengan menggunakan *set.seed(1)*, *seeding = 20* dan *iterate = 10*. Maksud dari *set.seed(1)* adalah untuk mereproduksi hasil agar hasilnya sama ketika dijalankan pada mesin yang lain. Pemilihan *seeding = 20* berarti akan dilakukan proses pembaruan medoid awal sebanyak 20 kali dengan *iterate* maksimum pada setiap *seeding* adalah 10 kali. Ringkasan proses pengelompokan disajikan pada Tabel 31.

Tabel 31 Ringkasan Hasil Pengelompokan dengan *seeding = 20*

<i>Seeding</i>	Iterasi	Medoid		Nilai <i>E</i>	<i>Seeding</i>	Iterasi	Medoid		Nilai <i>E</i>
		M1	M2				M1	M2	
1	Awal	S84	S33	21.5626	11	Awal	S84	S10	20.3830
	Akhir	S19	S25	14.0042		Akhir	S25	S19	14.0042
2	Awal	S84	S91	20.5610	12	Awal	S84	S45	20.5610
	Akhir	S25	S19	14.0042		Akhir	S25	S19	14.0042
3	Awal	S84	S38	21.0737	13	Awal	S84	S165	21.2654
	Akhir	S25	S136	14.0034		Akhir	S19	S71	13.9869

<i>Seeding</i>	Iterasi	Medoid		Nilai <i>E</i>	<i>Seeding</i>	Iterasi	Medoid		Nilai <i>E</i>
		<i>M1</i>	<i>M2</i>				<i>M1</i>	<i>M2</i>	
4	Awal	S84	S26	22.2622	14	Awal	S84	S116	20.0711
	Akhir	S19	S71	13.9869		Akhir	S25	S136	14.0034
5	Awal	S84	S10	20.3840	15	Awal	S84	S85	26.8921
	Akhir	S25	S19	14.0042		Akhir	S84	S85	26.8921
6	Awal	S84	S141	22.3068	16	Awal	S84	S49	20.5483
	Akhir	S19	S71	13.9869		Akhir	S25	S136	14.0034
7	Awal	S84	S31	24.4914	17	Awal	S84	S153	22.4264
	Akhir	S19	S71	13.9869		Akhir	S19	S71	13.9869
8	Awal	S84	S81	21.1925	18	Awal	S84	S162	26.5587
	Akhir	S19	S71	13.9869		Akhir	S84	S54	26.4388
9	Awal	S84	S80	36.4640	19	Awal	S84	S116	23.3607
	Akhir	S19	S71	13.9869		Akhir	S25	S136	14.0034
10	Awal	S84	S82	21.5147	20	Awal	S84	S147	25.6213
	Akhir	S19	S25	14.0042		Akhir	S19	S25	14.0042

Berdasarkan Tabel 31 dapat dilihat bahwa nilai *E* minimum didapatkan ketika medoid akhirnya adalah objek ke-19 (atau S19) dan ke-71 (atau S71) yaitu sebesar 13.9869. Sehingga medoid final yang diperoleh dari hasil pengelompokan SMA adalah S19 dan S71. Maka didapatkan hasil pengelompokan yang menunjukkan bahwa klaster 1 terdiri dari 83 anggota dan klaster 2 terdiri dari 82 anggota yang disajikan pada Tabel 32.

Tabel 32 Hasil Pengelompokan SMA di Provinsi Yogyakarta

Klaster	Anggota Klaster
1 (83 SMA)	S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14, S15, S16, S17, S18, S19, S37, S38, S39, S40, S41, S42, S43, S44, S45, S46, S47, S48, S49, S50, S51, S52, S53, S68, S84, S85, S86, S87, S88, S89, S90, S91, S92, S93, S94, S95, S96, S103, S109, S110, S111, S112, S113, S114, S115, S116, S117, S118, S119, S126, S127, S128, S129, S130, S131, S132, S133, S134, S135, S136, S137, S139, S140, S143, S150, S157, S159, S160, S162, S163
2 (82 SMA)	S20, S21, S22, S23, S24, S25, S26, S27, S28, S29, S30, S31, S32, S33, S34, S35, S36, S54, S55, S56, S57, S58, S59, S60, S61, S62, S63, S64, S65, S66, S67, S69, S70, S71, S72, S73, S74, S75, S76, S77, S78, S79, S80, S81, S82, S83, S97, S98, S99, S100, S101, S102, S104, S105, S106, S107, S108, S120, S121, S122, S123, S124, S125, S138, S141, S142, S144, S145, S146, S147, S148, S149, S151, S152, S153, S154, S155, S156, S158, S161, S164, S165

4.4 Validasi Pengelompokan *Simple K-Medoids Partitioning* (SKM)

Pada hasil analisis pengelompokan yang telah didapatkan akan membutuhkan perhitungan keakuratan hasil. Percobaan dengan jumlah *k* klaster dilakukan untuk mengetahui nilai SI dan MSV setiap klaster. Jumlah *k* klaster yang digunakan yaitu 2, 3, 4 dan 5. Hasil validasi untuk nilai SI dan MSV yang disajikan pada Tabel 33.

Tabel 33 Nilai Validasi SI dan MSV

Percobaan <i>k</i> Klaster	Nilai <i>SI</i> Global	Nilai <i>MSV</i> Global
2	0.5886	0.7129

Percobaan k Klaster	Nilai SI Global	Nilai MSV Global
3	0.4737	0.6272
4	0.4448	0.5853
5	0.3470	0.5327

Penentuan nilai SI dan MSV terbaik apabila nilainya mendekati 1. Berdasarkan Tabel 33 didapatkan bahwa nilai tertinggi di antara empat percobaan yang dilakukan adalah percobaan dengan jumlah k klaster sebanyak 2. Sehingga jumlah k klaster optimal untuk pengelompokan data SMA di Provinsi Yogyakarta menggunakan algoritma SKM adalah 2 klaster.

4.5 Profilisasi Pengelompokan *Simple K-Medoids Partitioning* (SKM)

Profilisasi klaster dilakukan untuk mengetahui karakteristik dari suatu kelompok yang terbentuk. Pada pengelompokan ini menggunakan data campuran maka terdapat dua karakteristik data yaitu data numerik dan kategorik. Karakteristik klaster berdasarkan data numerik dilihat dari rata – rata, disajikan pada Tabel 34.

Tabel 34 Profilisasi Klaster Berdasarkan Variabel Numerik

Klaster	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	40.5	553	2.45	3.82	81.2	63	7.53	11.1
2	20.5	102	2.18	1.57	1.57	29.6	4.26	4.99

Apabila dilihat pada Tabel 34, baris berwarna merah tua merupakan kelompok dengan jumlah SDM dan infrastruktur komponen kesiapan TIK untuk ANBK kategori tinggi dan untuk baris berwarna merah muda merupakan kelompok dengan jumlah SDM dan infrastruktur komponen kesiapan TIK untuk ANBK kategori rendah. Karakteristik klaster berdasarkan data kategorik dilihat dari besarnya persentase salah satu kategori dalam variabel tertentu yang mendominasi pada suatu klaster, di mana hasilnya disajikan pada Tabel 35.

Tabel 35 Profilisasi Klaster Berdasarkan Variabel Kategorik

Klaster	X_9		X_{10}		X_{11}	
	0	1	0	1	0	1
1	83.13%	16.87%	2.41%	97.59%	92.77%	7.23%
2	-	100%	18.29%	81.71%	90.24%	9.76%

Klaster	X_{12}				X_{13}		
	0	1	2	3	0	1	2
1	100%	-	-	-	98.80%	1.20%	-
2	40.24%	42.68%	7.32%	9.76%	1.22%	90.24%	8.54%

Berdasarkan Tabel 35 diperoleh informasi karakteristik klaster sebagai berikut:

- Klaster 1: merupakan kelompok sekolah SMA yang didominasi oleh sekolah negeri dengan moda pelaksanaan TIK berupa semi online, dengan sumber listrik yang berasal dari PLN, telah terakreditasi A serta memiliki status pelaksanaan TIK secara mandiri dan ditumpangi.
- Klaster 2: merupakan kelompok sekolah SMA yang didominasi oleh sekolah swasta dengan moda pelaksanaan TIK berupa semi online, dengan sumber listrik berasal dari PLN, terakreditasi A dan B dan memiliki status pelaksanaan TIK secara mandiri.

4.6 Pengujian Rata – Rata Dua Populasi Independen

Pengujian rata-rata dua populasi independen digunakan untuk mengetahui ada atau tidaknya perbedaan yang signifikan pada nilai rata – rata variabel numerik

klaster 1 dan 2. Pengujian ini menggunakan statistik uji Z karena jumlah sampel yang besar dan diasumsikan normal (Long, 2020).

H_0 : Tidak ada perbedaan yang signifikan pada rata – rata variabel numerik klaster 1 dan 2

H_1 : Ada perbedaan signifikan pada rata – rata variabel numerik klaster 1 dan 2

Statistik Uji:

Tabel 36 Statistik Uji Rata – Rata Dua Populasi Independen Variabel Numerik

Variabel	Nilai Z_{hitung}	Keputusan
X_1 (Guru)	7.4087	Tolak H_0
X_2 (Siswa)	13.811	Tolak H_0
X_3 (Laboratorium Komputer)	0.32395	Gagal Tolak H_0
X_4 (Komputer Utama / Proktor)	3.37781	Tolak H_0
X_5 (Komputer Sesuai Spesifikasi (Milik))	13.182	Tolak H_0
X_6 (Rata – Rata <i>Bandwidth</i>)	3.1875	Tolak H_0
X_7 (<i>Switch Hub</i>)	2.5919	Tolak H_0
X_8 (WIFI)	2.9504	Tolak H_0

Nilai $Z_{\alpha/2}$ sebesar 1.96. Berdasarkan Tabel 36 diperoleh kesimpulan bahwa terdapat satu variabel numerik yang tidak berbeda signifikan antara klaster 1 dan 2. Variabel tersebut adalah variabel X_3 atau jumlah laboratorium komputer. Selain pengujian pada variabel numerik yang digunakan dalam penelitian ini, peneliti melakukan uji rata – rata dua populasi pada variabel numerik yang tidak masuk dalam penelitian. Variabel numerik tersebut adalah daya listrik, *bandwidth upload* dan *bandwidth download*. Diperoleh hasil pengujian pada Tabel 36.

Tabel 37 Statistik Uji Rata – Rata Dua Populasi Independen Variabel Numerik Non Penelitian

Variabel	Nilai Z_{hitung}	Keputusan
Daya Listrik	2.2328	Tolak H_0
<i>Bandwidth Upload</i>	2.9302	Tolak H_0
<i>Bandwidth Download</i>	3.3635	Tolak H_0

Pada hasil pengujian hipotesis rata – rata dua populasi independen untuk variabel numerik non penelitian diperoleh bahwa ketiga variabel numerik yaitu daya listrik, *bandwidth upload* dan *bandwidth download* memiliki perbedaan yang signifikan antara klaster 1 dan klaster 2.

5 Kesimpulan

Hasil pengelompokan SMA di Provinsi Yogyakarta menggunakan algoritma *Simple K-Medoids Partitioning* (SKM) dengan *set.seed(1)*, *seeding* = 20 dan *iterate* = 10 menunjukkan bahwa klaster 1 yang terdiri dari 83 sekolah merupakan kelompok sekolah dengan jumlah SDM dan infrastruktur pada komponen kesiapan TIK untuk ANBK yang tinggi dan klaster 2 terdiri dari 82 sekolah, merupakan kelompok sekolah dengan jumlah SDM dan infrastruktur pada komponen kesiapan TIK untuk ANBK yang rendah. Validasi klaster menggunakan nilai *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV) memberikan hasil jika jumlah k klaster optimal yang direkomendasikan untuk pengelompokan sekolah SMA di Provinsi Yogyakarta adalah 2 klaster. Pada hasil pengujian rata – rata dua populasi independen menggunakan uji Z memberikan hasil bahwa seluruh variabel numerik antara klaster 1 dan 2 memiliki perbedaan nilai rata – rata yang signifikan kecuali

pada variabel jumlah laboratorium komputer yang tidak mempunyai perbedaan rata – rata antara kedua klaster.

Daftar Pustaka

- [1] Kemendikbud, Penyelenggaraan Asesmen Nasional, Balitbang dan Perbukuan, 2021.
- [2] d. Ayang Sari, "Penghapusan Ujian Nasional Tahun 2021 dalam Perspektif Guru SMA di Kota Tebing Tinggi," in *Prodising Seminar Nasional PBSI-III*, Medan, 2020.
- [3] T. Maharani, "Klarifikasi Nadiem Makarim: UN Tidak Dihapus, Hanya Diganti," 2019. [Online]. Available: <https://nasional.kompas.com/read/2019/12/13/08224851/klarifikasi-nadiem-makarim-un-tidak-dihapus-hanya-diganti?page=all>.
- [4] D. S. K. Rakyat, Statistik Pendidikan, Jakarta: Badan Pusat Statistik, 2021.
- [5] P. Kemendikbudristek, "Menilik Kesiapan Asesmen Nasional dari Kacamata Kecakapan Digital," 2021. [Online]. Available: <https://pusdatin.kemdikbud.go.id/menilik-kesiapan-asesmen-nasional-dari-kacamata-kecakapan-digital/>.
- [6] R. d. Kaufman, *Finding Groups in Data*, New York, USA: John Wiley & Sons, 1990.
- [7] F. L. Weksi Budiaji, "Simple K-Medoids Partitioning Algorithm for Mixed Variable Data," *Algorithms*, 2019.
- [8] S.Sharma, *Applied Multivariate Techniques*, New York: John Wiley and Sons, Inc., 1996.
- [9] V. Puspitasari, *Pengelompokan Sekolah Menengah Atas Di Provinsi Daerah Istimewa Yogyakarta Berdasarkan Nilai Ujian Nasional Menggunakan Algoritma Agglomerative Hierarchical Clustering*, Yogyakarta: Universitas Sanata Dharma, 2016.
- [10] W. Budiaji, "Medoid-Based Shadow Value Validation And Visualization," *International Journal of Advances in Intelligent Informatics*, vol. 5, no. 2, pp. 76 - 88, Juli 2019.
- [11] R. W. S. Insani, "Memahami Teorema Limit Pusat," 2020. [Online]. Available: <https://www.rachmatwahid.com/2020/04/teorema-limit-pusat.html>.
- [12] D. Hermawan, *Diktat Statistika Pendidikan dengan SPSS*, Jember: IAIN Jember, 2020.

LAMPIRAN

Lampiran 1 Data Kode Sekolah Menengah Atas (SMA)

No.	Nama Sekolah	Kode Sekolah
1	SMA Negeri 1 Pundong	S1
2	SMA Negeri 3 Bantul	S2
3	SMA Negeri 1 Bambanglipuro	S3
4	SMA Negeri 1 Sedayu	S4
5	SMA Negeri 1 Kretek	S5
6	SMA Negeri 1 Jetis	S6
7	SMA Negeri 1 Pleret	S7
8	SMA Negeri 1 Banguntapan	S8
9	SMA Negeri 1 Dlingo	S9
10	SMA Negeri 1 Srandakan	S10
11	SMA Negeri 1 Sewon	S11
12	SMA Negeri 1 Pajangan	S12
13	SMA Negeri 1 Kasihan	S13
14	SMA Negeri 2 Bantul	S14
15	SMA Negeri 1 Piyungan	S15
16	SMA Negeri 1 Imogiri	S16
17	SMA Negeri 1 Bantul	S17
18	SMA Negeri 2 Banguntapan	S18
19	SMA Negeri 1 Sanden	S19
20	SMA Muhammadiyah Pleret	S20
21	SMA Ali Maksum	S21
22	SMA BOPKRI Banguntapan	S22
23	SMA UII Banguntapan	S23
24	SMA Pangudi Luhur Sedayu	S24
25	SMA Stella Duce Bambanglipuro	S25
26	SMA Tumbuh	S26
27	SMA Muhammadiyah Bantul	S27
28	SMA Muhammadiyah Piyungan	S28
29	SMA Muhammadiyah Imogiri	S29
30	SMAS Muhammadiyah Sewon	S30
31	SMA Mutiara Persada	S31
32	SMA 17 Bantul	S32
33	SMA PGRI Kasihan	S33
34	SMA Muhammadiyah Kasihan	S34
35	SMA Dharma Amiluhur	S35
36	SMA Patria Bantul	S36

No.	Nama Sekolah	Kode Sekolah
37	SMAN 1 Seyegan	S37
38	SMAN 1 Ngaglik	S38
39	SMAN 1 Prambanan	S39
40	SMAN 1 Turi	S40
41	SMAN 1 Pakem	S41
42	SMAN 1 Sleman	S42
43	SMAN 1 Kalasan	S43
44	SMAN 1 Mlati	S44
45	SMAN 1 Cangkringan	S45
46	SMAN 1 Tempel	S46
47	SMAN 1 Gamping	S47
48	SMAN 2 Sleman	S48
49	SMAN 1 Godean	S49
50	SMAN 2 Ngaglik	S50
51	SMAN 1 Minggir Sleman	S51
52	SMAN 1 Ngemplak	S52
53	SMAN 1 Depok	S53
54	SMAS Kolese De Britto Depok	S54
55	SMAS Institut Indonesia Berbah	S55
56	SMA Islam Al Azhar 9 Yogyakarta	S56
57	SMAS Muhammadiyah Pakem	S57
58	SMAS Islam Prambanan	S58
59	SMA Insan Mulia Boarding School	S59
60	SMA Islam 3 Pakem	S60
61	SMAS Ma'arif 1 Sleman	S61
62	SMAS Immanuel Kalasan	S62
63	SMAS Muhammadiyah 1 Prambanan	S63
64	SMAS Islam 1 Sleman	S64
65	SMAS Terpadu Bina Umat	S65
66	SMAS Santo Mikael Mlati	S66
67	SMAS Muhammadiyah Kalasan	S67
68	SMAS Muhammadiyah Boarding School (MBS) Sleman	S68
69	SMA Islam Terpadu Baitussalam	S69
70	SMAS Sunan Kalijogo	S70
71	SMAS Muhammadiyah 1 Sleman	S71
72	SMAS Gama Depok Yogyakarta	S72
73	SMAS Cahaya Bangsa Utama	S73
74	SMAS Darul Hikmah	S74
75	SMAS Sulaiman Sleman	S75
76	SMA Budi Utama	S76
77	SMAS Angkasa Adisucipto	S77
78	SMAS Mandala Bhakti	S78

No.	Nama Sekolah	Kode Sekolah
79	SMA Olifant	S79
80	SMAS IKIP Veteran	S80
81	SMAS Muhammadiyah Mlati	S81
82	SMA Sains Wahid Hasyim	S82
83	SMAS Dr Wahidin Mlati	S83
84	SMAS Kolombo Depok	S84
85	SMA Budi Mulia Dua	S85
86	SMAN 1 Tanjungsari	S86
87	SMAN 1 Rongkop	S87
88	SMAN 1 Semanu	S88
89	SMAN 1 Patuk	S89
90	SMAN 2 Wonosari	S90
91	SMAN 2 Playen	S91
92	SMAN 1 Playen	S92
93	SMAN 1 Semin	S93
94	SMAN 1 Karangmojo	S94
95	SMAN 1 Panggang	S95
96	SMAN 1 Wonosari	S96
97	SMA Gotong Royong Semin	S97
98	SMA IKIP Veteran Iii Tambakromo	S98
99	SMA PGRI Playen	S99
100	SMA Muhammadiyah Daarul Khoir Gunungkidul	S100
101	SMA Dominikus Wonosari	S101
102	SMA Muhammadiyah Ngawen	S102
103	SMA Pembangunan 2 Karangmojo	S103
104	SMA Muhammadiyah Ponjong	S104
105	SMA Muhammadiyah Al Mujahidin Wonosari	S105
106	SMA Baitul Qur'an Yogyakarta	S106
107	SMA Pembangunan 3 Ponjong	S107
108	SMAS Pembangunan 1 Wonosari	S108
109	SMA Negeri 1 Temon	S109
110	SMA Negeri 1 Galur	S110
111	SMA Negeri 1 Lendah	S111
112	SMA Negeri 1 Girimulyo	S112
113	SMA Negeri 1 Kalibawang	S113
114	SMA Negeri 1 Kokap	S114
115	SMA Negeri 1 Samigaluh	S115
116	SMA Negeri 1 Pengasih	S116
117	SMA Negeri 1 Sentolo	S117
118	SMA Negeri 2 Wates	S118
119	SMA Negeri 1 Wates	S119
120	SMA IT Al Ghifari Kulon Progo	S120

No.	Nama Sekolah	Kode Sekolah
121	SMA Muhammadiyah Al Manar Galur	S121
122	SMA Muhammadiyah Wates	S122
123	SMA IT Abu Bakar Boarding School Kulon Progo	S123
124	SMA Maarif Wates	S124
125	SMA Sanjaya XIV Nanggulan	S125
126	SMAN 1 Yogyakarta	S126
127	SMAN 8 Yogyakarta	S127
128	SMAN 6 Yogyakarta	S128
129	SMAN 2 Yogyakarta	S129
130	SMAN 9 Yogyakarta	S130
131	SMAN 5 Yogyakarta	S131
132	SMAN 11 Yogyakarta	S132
133	SMAN 3 Yogyakarta	S133
134	SMAN 4 Yogyakarta	S134
135	SMAN 7 Yogyakarta	S135
136	SMAN 10 Yogyakarta	S136
137	SMAS Stella Duce 2	S137
138	SMAS Gotong Royong	S138
139	SMAS Muhammadiyah 4 Yogyakarta	S139
140	SMAS Muhammadiyah 7 Yogyakarta	S140
141	SMAS Gadjah Mada	S141
142	SMAS Budaya Wacana	S142
143	SMAS Muhammadiyah 2 Yogyakarta	S143
144	SMAS BOPKRI 1	S144
145	SMA 17 I Yogyakarta	S145
146	SMAS Katolik Sang Timur	S146
147	SMAS Muhammadiyah 5 Yogyakarta	S147
148	SMAS Maarif Yogyakarta	S148
149	SMAS Santo Thomas	S149
150	SMAS Stella Duce 1	S150
151	SMAS Muhammadiyah 6 Yogyakarta	S151
152	SMAS Budi Luhur	S152
153	SMAS Sultan Agung	S153
154	SMAS Piri 1 Yogyakarta	S154
155	SMAS Perak	S155
156	SMAS Santa Maria	S156
157	SMAS Muhammadiyah 3 Yogyakarta	S157
158	SMAS Bhinneka Tunggal Ika	S158
159	SMAS Pangudi Luhur	S159
160	SMAS IT Abu Bakar	S160
161	SMAS BOPKRI 2	S161
162	SMAS Muhammadiyah 1 Yogyakarta	S162

No.	Nama Sekolah	Kode Sekolah
163	SMAS Taman Madya IP	S163
164	SMAS Taman Madya Jetis	S164
165	SMAS IT Bina Anak Sholeh	S165

Lampiran 2 Data Penelitian

Kode Sekolah	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃
S1	42	608	3	6	80	100	3	6	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S2	47	215	3	3	64	393	3	7	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S3	45	609	2	4	85	20	5	25	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S4	61	879	4	4	100	80	4	2	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S5	30	402	2	2	70	57	4	2	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S6	50	862	3	4	124	30	8	7	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S7	32	526	2	4	30	50	5	20	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S8	51	735	3	4	140	300	7	11	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S9	31	414	3	3	60	40	6	2	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S10	29	372	2	4	46	59	7	7	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S11	65	1073	3	3	120	10	6	10	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S12	32	513	2	2	40	3	2	7	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S13	55	862	4	4	130	60	10	5	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S14	55	755	5	6	130	150	10	20	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S15	37	532	3	3	72	19	7	12	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S16	41	630	2	3	64	57	6	4	Negeri	Online	PLN	A	Mandiri dan Ditumpangi
S17	51	883	3	8	120	150	5	30	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S18	42	739	3	3	105	60	7	1	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S19	38	641	3	3	65	20	6	7	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi

Kode Sekolah	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃
S20	13	199	1	1	19	6	1	1	Swasta	Semi Online	PLN	A	Mandiri
S21	26	313	1	2	25	58	6	4	Swasta	Semi Online	PLN & Diesel	A	Mandiri
S22	16	40	1	1	15	16	1	1	Swasta	Semi Online	PLN	A	Mandiri
S23	19	125	1	2	28	20	5	5	Swasta	Semi Online	PLN	A	Mandiri
S24	22	295	0	4	63	30	5	3	Swasta	Online	PLN	A	Mandiri
S25	17	77	1	1	20	18	2	1	Swasta	Semi Online	PLN	A	Mandiri
S26	16	28	1	1	20	71	2	3	Swasta	Online	PLN	B	Mandiri
S27	27	92	1	3	30	9	2	6	Swasta	Semi Online	PLN	A	Mandiri
S28	15	29	1	1	10	5	1	1	Swasta	Semi Online	PLN	A	Mandiri
S29	27	110	1	2	30	16	1	0	Swasta	Semi Online	PLN	A	Mandiri
S30	14	41	1	1	23	40	24	2	Swasta	Semi Online	PLN	A	Mandiri
S31	3	6	1	2	40	23	2	1	Swasta	Online	PLN	Tidak /Belum Terakreditasi	Mandiri
S32	11	12	1	1	6	2	2	2	Swasta	Online	PLN	B	Mandiri
S33	16	67	1	1	22	10	3	33	Swasta	Semi Online	PLN	A	Mandiri
S34	17	50	1	2	5	15	1	15	Swasta	Online	PLN	B	Mandiri
S35	17	86	1	2	15	7	16	3	Swasta	Online	PLN & Diesel	B	Mandiri
S36	8	18	1	1	2	566	1	1	Swasta	Semi Online	PLN	B	Mandiri
S37	47	931	2	2	32	80	4	15	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S38	42	638	3	3	120	35	25	3	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S39	41	702	1	2	40	5	2	3	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S40	27	387	3	3	142	150	19	30	Negeri	Semi Online	PLN & Diesel	A	Mandiri dan Ditumpangi

Kode Sekolah	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃
S41	33	525	1	2	60	20	4	21	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S42	48	746	2	1	90	70	5	6	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S43	52	746	3	3	96	40	8	21	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S44	28	421	2	2	80	20	7	20	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S45	31	422	2	2	32	11	6	6	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S46	27	362	1	2	37	43	4	20	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S47	28	423	2	3	46	100	5	6	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S48	30	427	1	4	50	90	4	4	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S49	40	640	3	6	104	17	11	23	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S50	46	697	3	3	93	2	6	1	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S51	24	357	2	4	40	30	5	5	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S52	29	421	1	3	58	25	4	4	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S53	41	641	3	3	108	43	6	20	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S54	52	845	4	4	171	150	14	3	Swasta	Online	PLN & Diesel	A	Mandiri
S55	19	35	1	1	10	13	1	1	Swasta	Online	PLN	A	Mandiri
S56	34	341	2	3	50	13	4	1	Swasta	Semi Online	PLN	A	Mandiri
S57	18	15	1	1	23	13	4	2	Swasta	Semi Online	PLN	B	Mandiri
S58	22	65	1	2	20	5	1	2	Swasta	Semi Online	PLN	B	Mandiri
S59	9	12	1	1	7	10	2	2	Swasta	Semi Online	PLN	Tidak /Belum Terakreditasi	Menumpang
S60	13	69	1	1	13	1	1	2	Swasta	Semi Online	PLN	B	Mandiri
S61	11	20	1	1	10	9	1	1	Swasta	Semi Online	PLN	B	Mandiri

Kode Sekolah	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃
S62	17	47	1	1	18	1	3	2	Swasta	Semi Online	PLN	B	Mandiri
S63	19	242	1	2	35	5	3	7	Swasta	Semi Online	PLN	A	Mandiri
S64	28	172	2	4	45	23	4	2	Swasta	Semi Online	PLN	B	Mandiri
S65	22	366	2	2	30	28	2	2	Swasta	Semi Online	PLN	A	Mandiri
S66	15	74	1	1	25	2	6	5	Swasta	Online	PLN	A	Mandiri
S67	17	53	1	1	3	12	1	1	Swasta	Semi Online	PLN	B	Mandiri
S68	57	917	2	2	66	20	5	3	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S69	17	229	1	2	40	12	2	1	Swasta	Semi Online	PLN	B	Mandiri
S70	22	51	32	1	32	18	1	3	Swasta	Semi Online	PLN	B	Mandiri
S71	19	57	1	2	20	10	2	1	Swasta	Semi Online	PLN	B	Mandiri
S72	18	45	1	2	20	8	4	3	Swasta	Semi Online	PLN & Diesel	B	Mandiri
S73	20	18	1	1	20	28	1	1	Swasta	Semi Online	PLN	Tidak /Belum Terakreditasi	Mandiri
S74	11	25	1	1	8	100	16	2	Swasta	Online	PLN	B	Mandiri
S75	17	6	1	2	20	2	6	2	Swasta	Semi Online	PLN	B	Mandiri
S76	28	203	1	2	35	0	0	0	Swasta	Semi Online	PLN & Diesel	A	Mandiri
S77	31	245	60	2	45	60	3	4	Swasta	Semi Online	PLN	A	Mandiri
S78	6	12	0	0	2	14	0	0	Swasta	Online	PLN	C	Mandiri
S79	20	98	1	2	25	50	2	1	Swasta	Semi Online	PLN	Tidak /Belum Terakreditasi	Mandiri
S80	14	68	1	2	15	6	24	2	Swasta	Semi Online	PLN	B	Mandiri
S81	13	18	1	2	15	5	2	36	Swasta	Semi Online	PLN	B	Mandiri

Kode Sekolah	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃
S82	21	368	1	2	22	50	3	3	Swasta	Semi Online	PLN	A	Mandiri
S83	12	1	1	1	18	3	1	24	Swasta	Semi Online	PLN	C	Mandiri
S84	21	110	2	2	52	13	4	1	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S85	31	165	2	50	50	60	8	10	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S86	36	465	2	2	85	74	85	6	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S87	26	280	2	4	48	50	4	5	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S88	28	253	4	4	80	50	4	4	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S89	38	378	2	2	60	77	4	6	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S90	42	753	3	3	90	16	6	8	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S91	47	327	3	4	105	115	5	10	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S92	32	296	2	2	60	25	3	5	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S93	41	570	4	4	120	30	6	4	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S94	43	228	3	6	120	50	6	3	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S95	29	459	2	2	60	30	2	2	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S96	49	754	4	5	128	24	20	51	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S97	14	13	1	1	2	100	3	2	Swasta	Semi Online	PLN	B	Mandiri
S98	30	151	1	2	24	14	2	10	Swasta	Online	PLN	A	Mandiri
S99	13	27	1	1	5	10	1	1	Swasta	Semi Online	PLN	C	Mandiri
S100	14	27	0	1	1	10	1	1	Swasta	Semi Online	PLN	Tidak /Belum Terakreditasi	Menumpang
S101	19	21	1	2	26	20	3	2	Swasta	Semi Online	PLN	A	Mandiri
S102	15	188	1	2	35	6	2	4	Swasta	Semi Online	PLN	B	Mandiri

Kode Sekolah	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃
S103	27	84	2	2	35	4	2	2	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S104	20	66	1	2	9	110	1	1	Swasta	Semi Online	PLN	A	Mandiri
S105	15	125	1	2	21	15	2	2	Swasta	Semi Online	PLN	A	Mandiri
S106	16	68	1	1	20	13	4	2	Swasta	Semi Online	PLN	Tidak /Belum Terakreditasi	Mandiri
S107	20	52	1	1	25	9	1	1	Swasta	Semi Online	PLN	A	Mandiri
S108	22	55	1	2	2	11	4	1	Swasta	Semi Online	PLN	B	Mandiri
S109	30	384	1	2	62	18	2	2	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S110	28	68	1	2	50	8	3	3	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S111	43	384	3	3	70	18	8	20	Negeri	Semi Online	PLN & Diesel	A	Mandiri dan Ditumpangi
S112	23	240	1	2	40	45	2	4	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S113	28	343	3	3	80	95	3	12	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S114	22	30	1	2	30	100	4	3	Negeri	Semi Online	PLN & Diesel	A	Mandiri dan Ditumpangi
S115	20	132	1	2	25	360	2	5	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S116	46	612	3	4	105	24	10	4	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S117	20	529	3	4	90	31	8	12	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S118	38	466	1	1	30	40	4	5	Negeri	Semi Online	PLN	A	Mandiri
S119	51	756	3	6	142	100	7	4	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S120	17	22	1	2	18	2	22	126	Swasta	Semi Online	PLN	Tidak /Belum Terakreditasi	Mandiri
S121	14	7	1	0	1	0	0	0	Swasta	Semi Online	PLN	B	Mandiri
S122	23	92	1	2	7	8	4	0	Swasta	Semi Online	PLN	B	Menumpang

Kode Sekolah	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃
S123	22	94	2	2	5	70	2	2	Swasta	Semi Online	PLN & Diesel	A	Mandiri
S124	20	67	1	1	0	9	1	1	Swasta	Semi Online	PLN	B	Mandiri
S125	11	41	1	2	15	10	1	3	Swasta	Semi Online	PLN	B	Mandiri
S126	71	932	1	2	80	30	5	40	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S127	50	863	2	3	90	91	12	18	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S128	55	825	3	3	90	80	9	1	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S129	61	958	3	3	100	89	10	27	Negeri	Semi Online	PLN & Diesel	A	Mandiri dan Ditumpangi
S130	54	753	2	3	120	90	8	30	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S131	50	861	1	3	100	85	6	18	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S132	58	970	2	2	64	160	8	4	Negeri	Semi Online	PLN & Diesel	A	Mandiri dan Ditumpangi
S133	53	753	2	3	86	100	8	32	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S134	53	864	3	3	129	21	10	12	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S135	56	851	3	3	100	71	4	7	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S136	40	617	3	3	86	8	6	8	Negeri	Semi Online	PLN	A	Mandiri dan Ditumpangi
S137	33	460	2	4	80	50	5	8	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S138	6	18	0	0	2	15	0	0	Swasta	Semi Online	PLN	C	Menumpang
S139	41	267	3	3	60	4	6	2	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S140	35	324	2	3	70	19	5	10	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S141	19	0	2	2	20	28	8	2	Swasta	Online	PLN	B	Mandiri
S142	20	107	2	3	50	50	4	4	Swasta	Semi Online	PLN	B	Mandiri
S143	49	646	5	6	155	135	10	5	Swasta	Semi Online	PLN & Diesel	A	Mandiri dan Ditumpangi
S144	48	628	3	3	34	30	9	0	Swasta	Semi Online	PLN	A	Mandiri

Kode Sekolah	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃
S145	14	25	0	0	2	5	0	1	Swasta	Semi Online	PLN	B	Menumpang
S146	199	118	1	2	33	15	2	3	Swasta	Semi Online	PLN	A	Mandiri
S147	20	232	1	1	30	30	15	8	Swasta	Semi Online	PLN & Diesel	A	Mandiri
S148	18	25	1	1	20	5	25	1	Swasta	Semi Online	PLN	B	Mandiri
S149	10	13	1	1	12	7	2	2	Swasta	Online	PLN	B	Mandiri
S150	50	832	3	8	100	60	4	12	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S151	17	31	1	2	16	40	2	1	Swasta	Semi Online	PLN	A	Mandiri
S152	15	31	1	2	5	10	2	2	Swasta	Semi Online	PLN	C	Mandiri
S153	8	16	1	1	8	5	2	1	Swasta	Semi Online	PLN	C	Mandiri
S154	20	76	1	1	25	62	2	8	Swasta	Online	PLN	A	Mandiri
S155	14	9	1	1	5	3	0	0	Swasta	Semi Online	PLN	Tidak /Belum Terakreditasi	Menumpang
S156	20	171	2	3	60	25	4	10	Swasta	Semi Online	PLN & Diesel	A	Mandiri
S157	45	514	2	3	70	2	10	7	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S158	14	8	0	0	4	15	0	1	Swasta	Semi Online	PLN	B	Menumpang
S159	31	148	3	5	160	50	8	20	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S160	35	547	3	3	60	35	6	4	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S161	34	242	2	2	40	44	2	5	Swasta	Semi Online	PLN	A	Mandiri
S162	72	1157	4	4	160	110	30	71	Swasta	Online	PLN	A	Mandiri dan Ditumpangi
S163	27	41	1	1	20	21	2	2	Swasta	Semi Online	PLN	A	Mandiri dan Ditumpangi
S164	19	56	1	1	20	6	3	1	Swasta	Semi Online	PLN	B	Mandiri dan Ditumpangi
S165	11	57	1	0	9	50	24	2	Swasta	Semi Online	PLN	B	Mandiri

Lampiran 3 Data Penelitian untuk *R Studio*

Kode Sekolah	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	Daya Listrik	<i>Bandwidth Upload</i>	<i>Bandwidth Download</i>
S1	42	608	3	6	80	100	3	6	0	1	0	0	1	23000	100	100
S2	47	215	3	3	64	393	3	7	0	1	0	0	1	33000	366	419
S3	45	609	2	4	85	20	5	25	0	1	0	0	1	32000	20	20
S4	61	879	4	4	100	80	4	2	0	1	0	0	1	17000	80	80
S5	30	402	2	2	70	57	4	2	0	1	0	0	1	33000	19	95
S6	50	862	3	4	124	30	8	7	0	1	0	0	1	53000	30	30
S7	32	526	2	4	30	50	5	20	0	1	0	0	1	33000	20	80
S8	51	735	3	4	140	300	7	11	0	1	0	0	1	2640	300	300
S9	31	414	3	3	60	40	6	2	0	1	0	0	1	11000	40	40
S10	29	372	2	4	46	59	7	7	0	1	0	0	1	30000	18	99
S11	65	1073	3	3	120	10	6	10	0	1	0	0	1	22000	10	10
S12	32	513	2	2	40	3	2	7	0	1	0	0	1	22000	1	4
S13	55	862	4	4	130	60	10	5	0	1	0	0	1	23000	60	60
S14	55	755	5	6	130	150	10	20	0	1	0	0	1	66000	150	150
S15	37	532	3	3	72	19	7	12	0	1	0	0	1	18000	19	19
S16	41	630	2	3	64	57	6	4	0	0	0	0	1	13000	19	95
S17	51	883	3	8	120	150	5	30	0	1	0	0	1	88000	150	150
S18	42	739	3	3	105	60	7	1	0	1	0	0	1	11000	20	100
S19	38	641	3	3	65	20	6	7	0	1	0	0	1	92400	20	20
S20	13	199	1	1	19	6	1	1	1	1	0	0	0	33000	2	10

Kode Sekolah	X₁	X₂	X₃	X₄	X₅	X₆	X₇	X₈	X₉	X₁₀	X₁₁	X₁₂	X₁₃	Daya Listrik	Bandwidth Upload	Bandwidth Download
S21	26	313	1	2	25	58	6	4	1	1	1	0	0	1500	15	100
S22	16	40	1	1	15	16	1	1	1	1	0	0	0	4000	10	22
S23	19	125	1	2	28	20	5	5	1	1	0	0	0	16000	20	20
S24	22	295	0	4	63	30	5	3	1	0	0	0	0	47000	10	50
S25	17	77	1	1	20	18	2	1	1	1	0	0	0	1300	6	30
S26	16	28	1	1	20	71	2	3	1	0	0	1	0	1800	75	66
S27	27	92	1	3	30	9	2	6	1	1	0	0	0	10999	8	10
S28	15	29	1	1	10	5	1	1	1	1	0	0	0	1300	5	5
S29	27	110	1	2	30	16	1	0	1	1	0	0	0	5000	6	25
S30	14	41	1	1	23	40	24	2	1	1	0	0	0	82000	40	40
S31	3	6	1	2	40	23	2	1	1	0	0	3	0	3000	7	38
S32	11	12	1	1	6	2	2	2	1	0	0	1	0	22000	1	3
S33	16	67	1	1	22	10	3	33	1	1	0	0	0	900	10	10
S34	17	50	1	2	5	15	1	15	1	0	0	1	0	4000	10	20
S35	17	86	1	2	15	7	16	3	1	0	1	1	0	2200	3	10
S36	8	18	1	1	2	566	1	1	1	1	0	1	0	4500	586	546
S37	47	931	2	2	32	80	4	15	0	1	0	0	1	1300	80	80
S38	42	638	3	3	120	35	25	3	0	1	0	0	1	33000	30	40
S39	41	702	1	2	40	5	2	3	0	1	0	0	1	35000	1	8
S40	27	387	3	3	142	150	19	30	0	1	1	0	1	15000	150	150
S41	33	525	1	2	60	20	4	21	0	1	0	0	1	22000	20	20
S42	48	746	2	1	90	70	5	6	0	1	0	0	1	57000	70	70
S43	52	746	3	3	96	40	8	21	0	1	0	0	1	45000	40	40

Kode Sekolah	X₁	X₂	X₃	X₄	X₅	X₆	X₇	X₈	X₉	X₁₀	X₁₁	X₁₂	X₁₃	Daya Listrik	Bandwidth Upload	Bandwidth Download
S44	28	421	2	2	80	20	7	20	0	1	0	0	1	13000	20	20
S45	31	422	2	2	32	11	6	6	0	1	0	0	1	14300	9	12
S46	27	362	1	2	37	43	4	20	0	1	0	0	1	18700	35	50
S47	28	423	2	3	46	100	5	6	0	1	0	0	1	15000	100	100
S48	30	427	1	4	50	90	4	4	0	1	0	0	1	16500	90	90
S49	40	640	3	6	104	17	11	23	0	1	0	0	1	44450	14	20
S50	46	697	3	3	93	2	6	1	0	1	0	0	1	33000	1	2
S51	24	357	2	4	40	30	5	5	0	1	0	0	1	15000	10	50
S52	29	421	1	3	58	25	4	4	0	1	0	0	1	7700	25	25
S53	41	641	3	3	108	43	6	20	0	1	0	0	1	22000	35	50
S54	52	845	4	4	171	150	14	3	1	0	1	0	0	16000	150	150
S55	19	35	1	1	10	13	1	1	1	0	0	0	0	2200	11	15
S56	34	341	2	3	50	13	4	1	1	1	0	0	0	66000	12	14
S57	18	15	1	1	23	13	4	2	1	1	0	1	0	3200	8	17
S58	22	65	1	2	20	5	1	2	1	1	0	1	0	2200	5	5
S59	9	12	1	1	7	10	2	2	1	1	0	3	2	3000	10	10
S60	13	69	1	1	13	1	1	2	1	1	0	1	0	2200	1	1
S61	11	20	1	1	10	9	1	1	1	1	0	1	0	1200	7	10
S62	17	47	1	1	18	1	3	2	1	1	0	1	0	3300	1	1
S63	19	242	1	2	35	5	3	7	1	1	0	0	0	15999	5	5
S64	28	172	2	4	45	23	4	2	1	1	0	1	0	7700	15	30
S65	22	366	2	2	30	28	2	2	1	1	0	0	0	7700	15	40
S66	15	74	1	1	25	2	6	5	1	0	0	0	0	13000	1	2

Kode Sekolah	X₁	X₂	X₃	X₄	X₅	X₆	X₇	X₈	X₉	X₁₀	X₁₁	X₁₂	X₁₃	Daya Listrik	Bandwidth Upload	Bandwidth Download
S67	17	53	1	1	3	12	1	1	1	1	0	1	0	1300	12	12
S68	57	917	2	2	66	20	5	3	1	1	0	0	1	20000	20	20
S69	17	229	1	2	40	12	2	1	1	1	0	1	0	4399	4	20
S70	22	51	32	1	32	18	1	3	1	1	0	1	0	1500	6	29
S71	19	57	1	2	20	10	2	1	1	1	0	1	0	1300	4	16
S72	18	45	1	2	20	8	4	3	1	1	1	1	0	11000	2	14
S73	20	18	1	1	20	28	1	1	1	1	0	3	0	105000	31	24
S74	11	25	1	1	8	100	16	2	1	0	0	1	0	3300	100	100
S75	17	6	1	2	20	2	6	2	1	1	0	1	0	4400	2	2
S76	28	203	1	2	35	0	0	0	1	1	1	0	0	424000	0	0
S77	31	245	60	2	45	60	3	4	1	1	0	0	0	22000	20	100
S78	6	12	0	0	2	14	0	0	1	0	0	2	0	1300	14	14
S79	20	98	1	2	25	50	2	1	1	1	0	3	0	345000	50	50
S80	14	68	1	2	15	6	24	2	1	1	0	1	0	2200	7	5
S81	13	18	1	2	15	5	2	36	1	1	0	1	0	3500	2	8
S82	21	368	1	2	22	50	3	3	1	1	0	0	0	4000	50	50
S83	12	1	1	1	18	3	1	24	1	1	0	2	0	3000	2	4
S84	21	110	2	2	52	13	4	1	1	1	0	0	1	15000	5	20
S85	31	165	2	50	50	60	8	10	1	1	0	0	1	70000	60	60
S86	36	465	2	2	85	74	85	6	0	1	0	0	1	16700	70	77
S87	26	280	2	4	48	50	4	5	0	1	0	0	1	15000	50	50
S88	28	253	4	4	80	50	4	4	0	1	0	0	1	22000	50	50
S89	38	378	2	2	60	77	4	6	0	1	0	0	1	33000	87	66

Kode Sekolah	X₁	X₂	X₃	X₄	X₅	X₆	X₇	X₈	X₉	X₁₀	X₁₁	X₁₂	X₁₃	Daya Listrik	Bandwidth Upload	Bandwidth Download
S90	42	753	3	3	90	16	6	8	0	1	0	0	1	65000	9	22
S91	47	327	3	4	105	115	5	10	0	1	0	0	1	33000	30	200
S92	32	296	2	2	60	25	3	5	0	1	0	0	1	11000	20	30
S93	41	570	4	4	120	30	6	4	0	1	0	0	1	41500	10	50
S94	43	228	3	6	120	50	6	3	0	1	0	0	1	35000	25	75
S95	29	459	2	2	60	30	2	2	0	1	0	0	1	21997	17	42
S96	49	754	4	5	128	24	20	51	0	1	0	0	1	100000	9	38
S97	14	13	1	1	2	100	3	2	1	1	0	1	0	1200	100	100
S98	30	151	1	2	24	14	2	10	1	0	0	0	0	5000	7	20
S99	13	27	1	1	5	10	1	1	1	1	0	2	0	2199	3	16
S100	14	27	0	1	1	10	1	1	1	1	0	3	2	4400	10	10
S101	19	21	1	2	26	20	3	2	1	1	0	0	0	3500	10	30
S102	15	188	1	2	35	6	2	4	1	1	0	1	0	2200	4	7
S103	27	84	2	2	35	4	2	2	1	1	0	0	1	3000	2	5
S104	20	66	1	2	9	110	1	1	1	1	0	0	0	3500	100	120
S105	15	125	1	2	21	15	2	2	1	1	0	0	0	13000	9	20
S106	16	68	1	1	20	13	4	2	1	1	0	3	0	20200	4	21
S107	20	52	1	1	25	9	1	1	1	1	0	0	0	5000	7	11
S108	22	55	1	2	2	11	4	1	1	1	0	1	0	600	2	19
S109	30	384	1	2	62	18	2	2	0	1	0	0	1	22000	19	17
S110	28	68	1	2	50	8	3	3	0	1	0	0	1	11000	6	10
S111	43	384	3	3	70	18	8	20	0	1	1	0	1	30000	15	20
S112	23	240	1	2	40	45	2	4	0	1	0	0	1	10000	30	60

Kode Sekolah	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	Daya Listrik	<i>Bandwidth Upload</i>	<i>Bandwidth Download</i>
S113	28	343	3	3	80	95	3	12	0	1	0	0	1	15000	90	100
S114	22	30	1	2	30	100	4	3	0	1	1	0	1	2250	100	100
S115	20	132	1	2	25	360	2	5	0	1	0	0	1	2200	370	350
S116	46	612	3	4	105	24	10	4	0	1	0	0	1	66000	8	40
S117	20	529	3	4	90	31	8	12	0	1	0	0	1	40500	52	10
S118	38	466	1	1	30	40	4	5	0	1	0	0	0	99000	40	40
S119	51	756	3	6	142	100	7	4	0	1	0	0	1	33000	100	100
S120	17	22	1	2	18	2	22	126	1	1	0	3	0	1000	1	2
S121	14	7	1	0	1	0	0	0	1	1	0	1	0	1300	0	0
S122	23	92	1	2	7	8	4	0	1	1	0	1	2	12500	8	8
S123	22	94	2	2	5	70	2	2	1	1	1	0	0	100500	40	100
S124	20	67	1	1	0	9	1	1	1	1	0	1	0	5000	2	16
S125	11	41	1	2	15	10	1	3	1	1	0	1	0	2200	10	10
S126	71	932	1	2	80	30	5	40	0	1	0	0	1	30000	30	30
S127	50	863	2	3	90	91	12	18	0	1	0	0	1	120000	92	89
S128	55	825	3	3	90	80	9	1	0	1	0	0	1	16500	80	80
S129	61	958	3	3	100	89	10	27	0	1	1	0	1	22000	93	84
S130	54	753	2	3	120	90	8	30	0	1	0	0	1	98998	90	90
S131	50	861	1	3	100	85	6	18	0	1	0	0	1	50600	90	80
S132	58	970	2	2	64	160	8	4	0	1	1	0	1	49500	160	160
S133	53	753	2	3	86	100	8	32	0	1	0	0	1	75000	100	100
S134	53	864	3	3	129	21	10	12	0	1	0	0	1	6000	10	32
S135	56	851	3	3	100	71	4	7	0	1	0	0	1	144000	53	89

Kode Sekolah	X₁	X₂	X₃	X₄	X₅	X₆	X₇	X₈	X₉	X₁₀	X₁₁	X₁₂	X₁₃	Daya Listrik	Bandwidth Upload	Bandwidth Download
S136	40	617	3	3	86	8	6	8	0	1	0	0	1	60000	1	15
S137	33	460	2	4	80	50	5	8	1	1	0	0	1	44000	50	50
S138	6	18	0	0	2	15	0	0	1	1	0	2	2	1300	10	19
S139	41	267	3	3	60	4	6	2	1	1	0	0	1	32000	3	4
S140	35	324	2	3	70	19	5	10	1	1	0	0	1	34200	8	30
S141	19	0	2	2	20	28	8	2	1	0	0	1	0	2200	20	35
S142	20	107	2	3	50	50	4	4	1	1	0	1	0	33000	50	50
S143	49	646	5	6	155	135	10	5	1	1	1	0	1	380000	135	135
S144	48	628	3	3	34	30	9	0	1	1	0	0	0	197000	30	30
S145	14	25	0	0	2	5	0	1	1	1	0	1	2	1300	1	8
S146	199	118	1	2	33	15	2	3	1	1	0	0	0	6000	19	10
S147	20	232	1	1	30	30	15	8	1	1	1	0	0	32000	30	30
S148	18	25	1	1	20	5	25	1	1	1	0	1	0	1300	1	8
S149	10	13	1	1	12	7	2	2	1	0	0	1	0	7200	4	10
S150	50	832	3	8	100	60	4	12	1	1	0	0	1	171000	60	60
S151	17	31	1	2	16	40	2	1	1	1	0	0	0	11000	40	40
S152	15	31	1	2	5	10	2	2	1	1	0	2	0	2200	10	10
S153	8	16	1	1	8	5	2	1	1	1	0	2	0	44000	1	9
S154	20	76	1	1	25	62	2	8	1	0	0	0	0	24200	58	65
S155	14	9	1	1	5	3	0	0	1	1	0	3	2	998	1	4
S156	20	171	2	3	60	25	4	10	1	1	1	0	0	22499	25	25
S157	45	514	2	3	70	2	10	7	1	1	0	0	1	60000	2	2
S158	14	8	0	0	4	15	0	1	1	1	0	1	2	4000	10	20

Kode Sekolah	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	Daya Listrik	<i>Bandwidth Upload</i>	<i>Bandwidth Download</i>
S159	31	148	3	5	160	50	8	20	1	1	0	0	1	99998	50	50
S160	35	547	3	3	60	35	6	4	1	1	0	0	1	45000	20	50
S161	34	242	2	2	40	44	2	5	1	1	0	0	0	25000	58	29
S162	72	1157	4	4	160	110	30	71	1	0	0	0	1	345000	110	110
S163	27	41	1	1	20	21	2	2	1	1	0	0	1	5300	12	30
S164	19	56	1	1	20	6	3	1	1	1	0	1	1	7000	1	10
S165	11	57	1	0	9	50	24	2	1	1	0	1	0	7700	30	70

Lampiran 4 Ukuran Kedekatan Antar Objek Data Numerik

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	...	V165
1	8.00	6.96	7.58	7.42	7.49	7.25	7.31	7.01	7.43	7.35	6.96	7.34	7.20	7.20	7.56	7.68	7.14	7.52	7.62	6.78	...	6.43
2	6.96	8.00	6.67	6.50	7.04	6.36	6.69	6.82	7.03	6.99	6.10	6.79	6.29	6.40	6.88	6.94	6.34	6.60	6.89	6.72	...	6.37
3	7.58	6.67	8.00	7.26	7.36	7.32	7.45	6.89	7.35	7.25	7.11	7.34	7.15	7.14	7.65	7.57	7.16	7.43	7.63	6.76	...	6.35
4	7.42	6.50	7.26	8.00	7.14	7.60	6.90	7.08	7.08	6.94	7.45	6.91	7.65	7.29	7.15	7.34	7.37	7.64	7.27	6.41	...	6.08
5	7.49	7.04	7.36	7.14	8.00	6.99	7.44	6.60	7.84	7.71	6.74	7.56	6.95	6.71	7.62	7.65	6.65	7.36	7.56	7.27	...	6.94
6	7.25	6.36	7.32	7.60	6.99	8.00	6.88	7.27	7.04	6.93	7.62	6.94	7.83	7.43	7.25	7.27	7.44	7.61	7.34	6.36	...	6.06
7	7.31	6.69	7.45	6.90	7.44	6.88	8.00	6.53	7.51	7.62	6.63	7.67	6.78	6.77	7.55	7.49	6.72	7.10	7.46	7.20	...	6.90
8	7.01	6.82	6.89	7.08	6.60	7.27	6.53	8.00	6.59	6.55	6.97	6.45	7.29	7.46	6.83	6.88	7.24	7.22	6.86	5.87	...	5.62
9	7.43	7.03	7.35	7.08	7.84	7.04	7.51	6.59	8.00	7.75	6.79	7.60	6.94	6.70	7.67	7.68	6.61	7.34	7.66	7.28	...	6.97
10	7.35	6.99	7.25	6.94	7.71	6.93	7.62	6.55	7.75	8.00	6.62	7.63	6.87	6.65	7.52	7.55	6.55	7.18	7.49	7.32	...	7.07
11	6.96	6.10	7.11	7.45	6.74	7.62	6.63	6.97	6.79	6.62	8.00	6.76	7.50	7.15	7.07	7.02	7.25	7.33	7.13	6.18	...	5.76
12	7.34	6.79	7.34	6.91	7.56	6.94	7.67	6.45	7.60	7.63	6.76	8.00	6.77	6.56	7.61	7.53	6.50	7.13	7.60	7.41	...	6.88
13	7.20	6.29	7.15	7.65	6.95	7.83	6.78	7.29	6.94	6.87	7.50	6.77	8.00	7.57	7.08	7.23	7.39	7.58	7.17	6.22	...	6.04
14	7.20	6.40	7.14	7.29	6.71	7.43	6.77	7.46	6.70	6.65	7.15	6.56	7.57	8.00	6.95	6.99	7.60	7.34	6.96	5.98	...	5.79
15	7.56	6.88	7.65	7.15	7.62	7.25	7.55	6.83	7.67	7.52	7.07	7.61	7.08	6.95	8.00	7.69	6.85	7.44	7.81	7.03	...	6.66
16	7.68	6.94	7.57	7.34	7.65	7.27	7.49	6.88	7.68	7.55	7.02	7.53	7.23	6.99	7.69	8.00	6.90	7.60	7.86	6.99	...	6.71
17	7.14	6.34	7.16	7.37	6.65	7.44	6.72	7.24	6.61	6.55	7.25	6.50	7.39	7.60	6.85	6.90	8.00	7.23	6.88	5.92	...	5.62
18	7.52	6.60	7.43	7.64	7.36	7.61	7.10	7.22	7.34	7.18	7.33	7.13	7.58	7.34	7.44	7.60	7.23	8.00	7.53	6.64	...	6.37
19	7.62	6.89	7.63	7.27	7.56	7.34	7.46	6.86	7.66	7.49	7.13	7.60	7.17	6.96	7.81	7.86	6.88	7.53	8.00	7.02	...	6.63
20	6.78	6.72	6.76	6.41	7.27	6.36	7.20	5.87	7.28	7.32	6.18	7.41	6.22	5.98	7.03	6.99	5.92	6.64	7.02	8.00	...	7.43
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
165	6.43	6.37	6.35	6.08	6.94	6.06	6.90	5.62	6.97	7.07	5.76	6.88	6.04	5.79	6.66	6.71	5.62	6.37	6.63	7.43	...	8.00

Lampiran 5 Ukuran Kedekatan Antar Objek Data Kategorik

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	...	V165
1	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
2	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
3	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
6	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
7	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
8	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
9	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
10	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
11	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
12	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
13	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
14	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
15	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
16	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	4	4	4	2	...	1
17	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
18	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
19	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	3	...	2
20	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	3	3	3	5	...	4
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
165	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2	2	2	4	...	5

Lampiran 6 Ukuran Kedekatan Antar Objek Data Campuran

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	...	V165
1	1.00	0.92	0.97	0.96	0.96	0.94	0.95	0.92	0.96	0.95	0.92	0.95	0.94	0.94	0.97	0.90	0.93	0.96	0.97	0.75	...	0.65
2	0.92	1.00	0.90	0.88	0.93	0.87	0.90	0.91	0.93	0.92	0.85	0.91	0.87	0.88	0.91	0.84	0.87	0.89	0.91	0.75	...	0.64
3	0.97	0.90	1.00	0.94	0.95	0.95	0.96	0.91	0.95	0.94	0.93	0.95	0.93	0.93	0.97	0.89	0.94	0.96	0.97	0.75	...	0.64
4	0.96	0.88	0.94	1.00	0.93	0.97	0.92	0.93	0.93	0.92	0.96	0.92	0.97	0.95	0.93	0.87	0.95	0.97	0.94	0.72	...	0.62
5	0.96	0.93	0.95	0.93	1.00	0.92	0.96	0.89	0.99	0.98	0.90	0.97	0.92	0.90	0.97	0.90	0.90	0.95	0.97	0.79	...	0.69
6	0.94	0.87	0.95	0.97	0.92	1.00	0.91	0.94	0.93	0.92	0.97	0.92	0.99	0.96	0.94	0.87	0.96	0.97	0.95	0.72	...	0.62
7	0.95	0.90	0.96	0.92	0.96	0.91	1.00	0.89	0.96	0.97	0.89	0.97	0.91	0.91	0.97	0.88	0.90	0.93	0.96	0.78	...	0.68
8	0.92	0.91	0.91	0.93	0.89	0.94	0.89	1.00	0.89	0.89	0.92	0.88	0.95	0.96	0.91	0.84	0.94	0.94	0.91	0.68	...	0.59
9	0.96	0.93	0.95	0.93	0.99	0.93	0.96	0.89	1.00	0.98	0.91	0.97	0.92	0.90	0.97	0.90	0.89	0.95	0.97	0.79	...	0.69
10	0.95	0.92	0.94	0.92	0.98	0.92	0.97	0.89	0.98	1.00	0.89	0.97	0.91	0.90	0.96	0.89	0.89	0.94	0.96	0.79	...	0.70
11	0.92	0.85	0.93	0.96	0.90	0.97	0.89	0.92	0.91	0.89	1.00	0.90	0.96	0.93	0.93	0.85	0.94	0.95	0.93	0.71	...	0.60
12	0.95	0.91	0.95	0.92	0.97	0.92	0.97	0.88	0.97	0.97	0.90	1.00	0.91	0.89	0.97	0.89	0.88	0.93	0.97	0.80	...	0.68
13	0.94	0.87	0.93	0.97	0.92	0.99	0.91	0.95	0.92	0.91	0.96	0.91	1.00	0.97	0.93	0.86	0.95	0.97	0.94	0.71	...	0.62
14	0.94	0.88	0.93	0.95	0.90	0.96	0.91	0.96	0.90	0.90	0.93	0.89	0.97	1.00	0.92	0.85	0.97	0.95	0.92	0.69	...	0.60
15	0.97	0.91	0.97	0.93	0.97	0.94	0.97	0.91	0.97	0.96	0.93	0.97	0.93	0.92	1.00	0.90	0.91	0.96	0.99	0.77	...	0.67
16	0.90	0.84	0.89	0.87	0.90	0.87	0.88	0.84	0.90	0.89	0.85	0.89	0.86	0.85	0.90	1.00	0.84	0.89	0.91	0.69	...	0.59
17	0.93	0.87	0.94	0.95	0.90	0.96	0.90	0.94	0.89	0.89	0.94	0.88	0.95	0.97	0.91	0.84	1.00	0.94	0.91	0.69	...	0.59
18	0.96	0.89	0.96	0.97	0.95	0.97	0.93	0.94	0.95	0.94	0.95	0.93	0.97	0.95	0.96	0.89	0.94	1.00	0.96	0.74	...	0.64
19	0.97	0.91	0.97	0.94	0.97	0.95	0.96	0.91	0.97	0.96	0.93	0.97	0.94	0.92	0.99	0.91	0.91	0.96	1.00	0.77	...	0.66
20	0.75	0.75	0.75	0.72	0.79	0.72	0.78	0.68	0.79	0.79	0.71	0.80	0.71	0.69	0.77	0.69	0.69	0.74	0.77	1.00	...	0.88
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
165	0.65	0.64	0.64	0.62	0.69	0.62	0.68	0.59	0.69	0.70	0.60	0.68	0.62	0.60	0.67	0.59	0.59	0.64	0.66	0.88	...	1.00

Lampiran 7 Ukuran Jarak Antar Objek Data Campuran

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	...	V165
1	0	0.08	0.03	0.04	0.04	0.06	0.05	0.08	0.04	0.05	0.08	0.05	0.06	0.06	0.03	0.10	0.07	0.04	0.03	0.25	...	0.35
2	0.08	0	0.10	0.12	0.07	0.13	0.10	0.09	0.08	0.08	0.15	0.09	0.13	0.12	0.09	0.16	0.13	0.11	0.09	0.25	...	0.36
3	0.03	0.10	0	0.06	0.05	0.05	0.04	0.09	0.05	0.06	0.07	0.05	0.07	0.07	0.03	0.11	0.06	0.04	0.03	0.25	...	0.36
4	0.04	0.12	0.06	0	0.07	0.03	0.08	0.07	0.07	0.08	0.04	0.08	0.03	0.05	0.07	0.13	0.05	0.03	0.06	0.28	...	0.38
5	0.04	0.07	0.05	0.07	0	0.08	0.04	0.11	0.01	0.02	0.10	0.03	0.08	0.10	0.03	0.10	0.10	0.05	0.03	0.21	...	0.31
6	0.06	0.13	0.05	0.03	0.08	0	0.09	0.06	0.07	0.08	0.03	0.08	0.01	0.04	0.06	0.13	0.04	0.03	0.05	0.28	...	0.38
7	0.05	0.10	0.04	0.08	0.04	0.09	0	0.11	0.04	0.03	0.11	0.03	0.09	0.09	0.03	0.12	0.10	0.07	0.04	0.22	...	0.32
8	0.08	0.09	0.09	0.07	0.11	0.06	0.11	0	0.11	0.11	0.08	0.12	0.05	0.04	0.09	0.16	0.06	0.06	0.09	0.32	...	0.41
9	0.04	0.08	0.05	0.07	0.01	0.07	0.04	0.11	0	0.02	0.09	0.03	0.08	0.10	0.03	0.10	0.11	0.05	0.03	0.21	...	0.31
10	0.05	0.08	0.06	0.08	0.02	0.08	0.03	0.11	0.02	0	0.11	0.03	0.09	0.10	0.04	0.11	0.11	0.06	0.04	0.21	...	0.3
11	0.08	0.15	0.07	0.04	0.10	0.03	0.11	0.08	0.09	0.11	0	0.10	0.04	0.07	0.07	0.15	0.06	0.05	0.07	0.29	...	0.4
12	0.05	0.09	0.05	0.08	0.03	0.08	0.03	0.12	0.03	0.03	0.10	0	0.09	0.11	0.03	0.11	0.12	0.07	0.03	0.20	...	0.32
13	0.06	0.13	0.07	0.03	0.08	0.01	0.09	0.05	0.08	0.09	0.04	0.09	0	0.03	0.07	0.14	0.05	0.03	0.06	0.29	...	0.38
14	0.06	0.12	0.07	0.05	0.10	0.04	0.09	0.04	0.10	0.10	0.07	0.11	0.03	0	0.08	0.15	0.03	0.05	0.08	0.31	...	0.4
15	0.03	0.09	0.03	0.07	0.03	0.06	0.03	0.09	0.03	0.04	0.07	0.03	0.07	0.08	0	0.10	0.09	0.04	0.01	0.23	...	0.33
16	0.10	0.16	0.11	0.13	0.10	0.13	0.12	0.16	0.10	0.11	0.15	0.11	0.14	0.15	0.10	0	0.16	0.11	0.09	0.31	...	0.41
17	0.07	0.13	0.06	0.05	0.10	0.04	0.10	0.06	0.11	0.11	0.06	0.12	0.05	0.03	0.09	0.16	0	0.06	0.09	0.31	...	0.41
18	0.04	0.11	0.04	0.03	0.05	0.03	0.07	0.06	0.05	0.06	0.05	0.07	0.03	0.05	0.04	0.11	0.06	0	0.04	0.26	...	0.36
19	0.03	0.09	0.03	0.06	0.03	0.05	0.04	0.09	0.03	0.04	0.07	0.03	0.06	0.08	0.01	0.09	0.09	0.04	0	0.23	...	0.34
20	0.25	0.25	0.25	0.28	0.21	0.28	0.22	0.32	0.21	0.21	0.29	0.20	0.29	0.31	0.23	0.31	0.31	0.26	0.23	0	...	0.12
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
165	0.35	0.36	0.36	0.38	0.31	0.38	0.32	0.41	0.31	0.30	0.40	0.32	0.38	0.40	0.33	0.41	0.41	0.36	0.34	0.12	...	0

Lampiran 8 Ringkasan Hasil *Seeding* = 20 ($k = 3, 4, 5$)

Ringkasan Hasil *Seeding* = 20 untuk ($k = 3$)

<i>Seeding</i>	Iterasi	Medoid			Nilai <i>E</i>	<i>Seeding</i>	Iterasi	Medoid			Nilai <i>E</i>
		M1	M2	M3				M1	M2	M3	
1	Awal	84	33	3	14.7563	1	Awal	84	36	104	20.5979
	Akhir	140	71	19	12.7089		Akhir	19	61	23	11.2014
2	Awal	84	38	26	16.3491	2	Awal	84	130	22	19.1178
	Akhir	23	136	61	11.1999		Akhir	140	19	71	13.7329
3	Awal	84	10	72	16.5543	3	Awal	84	33	116	14.7099
	Akhir	23	19	61	11.2014		Akhir	140	71	19	12.7089
4	Awal	84	31	13	18.5806	4	Awal	84	139	116	19.8184
	Akhir	23	61	136	11.1999		Akhir	71	160	19	12.7277
5	Awal	84	80	82	18.4696	5	Awal	84	59	68	22.8438
	Akhir	19	61	23	11.2014		Akhir	23	61	136	11.1999
6	Awal	84	10	45	20.2748	6	Awal	84	128	79	15.872
	Akhir	25	43	52	12.9736		Akhir	140	19	71	12.7089
7	Awal	84	165	116	14.3465	7	Awal	84	62	109	13.5211
	Akhir	23	61	136	11.1999		Akhir	140	71	19	12.7089
8	Awal	84	85	49	20.4504	8	Awal	84	61	81	18.7719
	Akhir	71	140	19	12.7089		Akhir	23	136	61	11.1999
9	Awal	84	153	35	21.5334	9	Awal	84	14	20	15.3669
	Akhir	19	71	149	12.8484		Akhir	140	19	71	12.7089
10	Awal	84	116	85	19.9733	10	Awal	84	21	53	18.7234
	Akhir	71	19	140	12.7089		Akhir	61	23	136	11.1999

Ringkasan Hasil *Seeding* = 20 untuk ($k = 4$)

<i>Seeding</i>	Iterasi	Medoid				Nilai <i>E</i>	<i>Seeding</i>	Iterasi	Medoid				Nilai <i>E</i>
		M1	M2	M3	M4				M1	M2	M3	M4	
1	Awal	84	33	3	133	14.4215	11	Awal	84	128	79	54	15.4382
	Akhir	140	71	52	43	11.7687		Akhir	140	19	71	54	12.3422
2	Awal	84	26	9	138	14.2021	12	Awal	84	109	114	81	14.2631
	Akhir	23	71	19	138	10.7494		Akhir	23	136	111	61	10.721
3	Awal	84	31	13	121	13.9415	13	Awal	84	14	20	153	14.5038
	Akhir	29	154	136	61	10.6142		Akhir	140	136	23	61	10.2371
4	Awal	84	82	10	45	14.9038	14	Awal	84	53	17	134	19.7318
	Akhir	140	71	43	52	11.7687		Akhir	25	52	130	135	12.765
5	Awal	84	165	116	83	14.5199	15	Awal	84	77	131	60	13.3951
	Akhir	161	62	136	22	10.9927		Akhir	140	23	136	61	10.2371
6	Awal	84	49	24	120	17.605	16	Awal	84	123	163	43	19.0376
	Akhir	23	136	66	61	10.6112		Akhir	105	21	61	136	10.6831
7	Awal	84	116	85	32	15.1198	17	Awal	84	154	165	43	13.1418
	Akhir	23	136	140	61	10.2371		Akhir	140	23	61	136	10.2371
8	Awal	84	104	130	22	14.2158	18	Awal	84	86	102	46	14.6292
	Akhir	140	23	136	61	10.2371		Akhir	140	43	71	52	11.7687
9	Awal	84	33	116	139	14.5142	19	Awal	84	29	107	19	13.0485
	Akhir	23	61	136	140	10.2371		Akhir	140	23	61	136	10.2371
10	Awal	84	116	86	68	19.6311	20	Awal	84	60	13	150	13.9443
	Akhir	71	19	86	160	12.624		Akhir	23	61	13	140	10.2371

Ringkasan Hasil *Seeding* = 20 untuk ($k = 5$)

<i>Seeding</i>	Iterasi	Medoid					Nilai <i>E</i>	<i>Seeding</i>	Iterasi	Medoid					Nilai <i>E</i>
		M1	M2	M3	M4	M5				M1	M2	M3	M4	M5	
1	Awal	84	33	3	133	72	13.4257	11	Awal	84	114	148	85	124	19.2884
	Akhir	140	105	52	43	61	9.30261		Akhir	43	52	61	140	105	9.3026
2	Awal	84	10	72	36	130	14.8128	12	Awal	84	108	77	163	58	19.6945
	Akhir	23	52	71	61	43	14.7829		Akhir	136	61	161	25	71	10.8059
3	Awal	84	80	82	10	45	11.8587	13	Awal	84	154	165	43	111	12.7003
	Akhir	140	61	105	43	52	9.30261		Akhir	140	23	61	136	111	9.8046
4	Awal	84	165	116	83	62	13.3581	14	Awal	84	102	46	29	107	11.9067
	Akhir	23	80	136	99	711	10.5807		Akhir	140	62	136	161	22	10.0523
5	Awal	84	153	35	91	55	14.1566	15	Awal	84	19	125	13	150	12.0645
	Akhir	140	61	35	19	105	9.9784		Akhir	23	52	61	43	140	9.3075
6	Awal	84	36	104	148	22	18.007	16	Awal	84	149	7	159	117	14.5984
	Akhir	19	36	21	62	25	10.6113		Akhir	105	61	52	143	43	9.7843
7	Awal	84	33	116	139	91	14.1666	17	Awal	84	130	88	43	11	18.9464
	Akhir	23	61	43	140	52	9.30753		Akhir	25	130	52	90	6	12.6276
8	Awal	84	59	68	57	6	13.0151	18	Awal	84	101	59	136	16	12.2322
	Akhir	139	145	150	71	19	11.8664		Akhir	140	105	61	136	54	9.9050
9	Awal	84	62	109	114	13	11.9349	19	Awal	84	1	136	43	100	17.2557
	Akhir	23	61	52	111	43	11.6641		Akhir	23	89	52	135	61	10.1003
10	Awal	84	14	20	153	18	12.7748	20	Awal	84	86	42	55	31	16.3101
	Akhir	140	43	23	61	52	9.30753		Akhir	23	86	19	62	73	12.8619

Lampiran 9 Validasi Nilai *Silhouette Index* ($k = 2, 3, 4, 5$)

	$k = 2$		$k = 3$		$k = 4$		$k = 5$	
	Nilai SI	Klaster	Nilai SI	Klaster	Nilai SI	Klaster	Nilai SI	Klaster
S1	0.7591	1	0.7268	2	0.6231	2	0.0743	4
S2	0.6266	1	0.5652	2	0.4507	2	0.2310	3
S3	0.7585	1	0.7278	2	0.6221	2	0.1812	4
S4	0.7435	1	0.7149	2	0.5939	2	0.4381	4
S5	0.7362	1	0.6884	2	0.5982	2	0.4229	3
S6	0.7479	1	0.7207	2	0.5916	2	0.4942	4
S7	0.7062	1	0.6518	2	0.5699	2	0.3172	3
S8	0.6960	1	0.6633	2	0.5152	2	0.3713	4
S9	0.7366	1	0.6879	2	0.5964	2	0.4389	3
S10	0.7149	1	0.6574	2	0.5741	2	0.4714	3
S11	0.7126	1	0.6809	2	0.5329	2	0.4315	4
S12	0.7046	1	0.6474	2	0.568	2	0.4159	3
S13	0.7381	1	0.7100	2	0.5726	2	0.4933	4
S14	0.7218	1	0.6929	2	0.5494	2	0.4735	4
S15	0.7606	1	0.7245	2	0.6226	2	0.1390	3
S16	0.5864	1	0.5305	2	0.3954	2	0.0531	3
S17	0.7135	1	0.6834	2	0.5392	2	0.4697	4
S18	0.7638	1	0.7367	2	0.6317	2	0.3971	4
S19	0.7618	1	0.7264	2	0.629	2	0.0650	3
S20	0.5490	2	0.4339	1	0.476	1	0.5153	2
S21	0.3550	2	0.3986	1	0.4431	1	0.4613	2
S22	0.5850	2	0.3941	1	0.4366	1	0.4804	2
S23	0.5340	2	0.4764	1	0.5141	1	0.5501	2
S24	0.3591	2	0.3205	1	0.3681	1	0.3884	2
S25	0.5767	2	0.4315	1	0.4726	1	0.5161	2
S26	0.6075	2	0.3858	3	0.3593	4	0.3314	5
S27	0.5265	2	0.4705	1	0.5047	1	0.5403	2
S28	0.5870	2	0.3656	1	0.4086	1	0.4520	2
S29	0.5359	2	0.4731	1	0.5081	1	0.5459	2
S30	0.5287	2	0.3769	1	0.4143	1	0.4431	2
S31	0.5263	2	0.1192	3	0.0821	4	0.0465	5
S32	0.6205	2	0.4263	3	0.4022	4	0.3766	5
S33	0.5249	2	0.3887	1	0.4256	1	0.4540	2
S34	0.6050	2	0.3953	3	0.3698	4	0.3461	5
S35	0.5203	2	0.2283	3	0.1945	4	0.1710	5
S36	0.5533	2	0.3653	3	0.3478	4	0.3252	5
S37	0.7006	1	0.6543	2	0.5588	2	0.0608	4
S38	0.7293	1	0.6967	2	0.5681	2	0.3026	4
S39	0.7092	1	0.6592	2	0.5781	2	0.1705	3
S40	0.5641	1	0.5014	2	0.3635	2	0.0727	4
S41	0.7347	1	0.6899	2	0.5966	2	0.2974	3

	<i>k</i> = 2		<i>k</i> = 3		<i>k</i> = 4		<i>k</i> = 5	
	Nilai SI	Klaster	Nilai SI	Klaster	Nilai SI	Klaster	Nilai SI	Klaster
S42	0.7634	1	0.7349	2	0.6366	2	0.3280	4
S43	0.7631	1	0.7372	2	0.6288	2	0.4496	4
S44	0.7394	1	0.6976	2	0.6008	2	0.2441	3
S45	0.6903	1	0.6266	2	0.5557	2	0.4758	3
S46	0.6814	1	0.6157	2	0.5486	2	0.4656	3
S47	0.7168	1	0.6625	2	0.5832	2	0.4479	3
S48	0.7192	1	0.6656	2	0.5838	2	0.4478	3
S49	0.7489	1	0.7194	2	0.5985	2	0.3393	4
S50	0.7582	1	0.7289	2	0.623	2	0.2892	4
S51	0.6906	1	0.6257	2	0.5522	2	0.5000	3
S52	0.7266	1	0.6743	2	0.59	2	0.4796	3
S53	0.7653	1	0.7383	2	0.6321	2	0.3485	4
S54	0.0677	2	0.1936	1	0.101	1	0.1129	2
S55	0.5254	2	0.2335	1	0.2786	1	0.3160	2
S56	0.3743	2	0.4714	1	0.4528	1	0.4763	2
S57	0.6825	2	0.5065	3	0.4866	4	0.4554	5
S58	0.6821	2	0.4959	3	0.4758	4	0.4433	5
S59	0.4172	2	0.2121	3	0.2171	4	0.1914	5
S60	0.6869	2	0.5227	3	0.5032	4	0.4721	5
S61	0.6886	2	0.5390	3	0.5204	4	0.4915	5
S62	0.6868	2	0.5165	3	0.4969	4	0.4662	5
S63	0.4876	2	0.4817	1	0.5668	1	0.5899	2
S64	0.6038	2	0.3632	3	0.3429	4	0.3089	5
S65	0.4495	2	0.4830	1	0.5448	1	0.5669	2
S66	0.5014	2	0.2751	1	0.3178	1	0.3516	2
S67	0.6867	2	0.5307	3	0.5118	4	0.4815	5
S68	0.4430	1	0.2852	2	0.3448	3	0.3158	1
S69	0.6321	2	0.3995	3	0.3756	4	0.3413	5
S70	0.5952	2	0.3636	3	0.3438	4	0.3148	5
S71	0.6862	2	0.5048	3	0.4847	4	0.4521	5
S72	0.5733	2	0.2770	3	0.2459	4	0.2147	5
S73	0.6035	2	0.1531	3	0.1148	4	0.0692	5
S74	0.5856	2	0.3799	3	0.3548	4	0.3334	5
S75	0.6782	2	0.5066	3	0.4873	4	0.4588	5
S76	0.3940	2	0.3891	1	0.4233	1	0.4500	2
S77	0.3118	2	0.3355	1	0.345	1	0.3615	2
S78	0.5454	2	0.1952	3	0.1621	4	0.1307	5
S79	0.5839	2	-0.1274	1	-0.09	1	-0.0435	2
S80	0.6518	2	0.4656	3	0.4453	4	0.4234	5
S81	0.6436	2	0.4615	3	0.4415	4	0.4239	5
S82	0.4588	2	0.4719	1	0.5118	1	0.5410	2
S83	0.5727	2	0.1812	3	0.1501	4	0.1209	5
S84	0.1794	1	0.0726	1	0.4134	3	0.3959	1

	<i>k</i> = 2		<i>k</i> = 3		<i>k</i> = 4		<i>k</i> = 5	
	Nilai SI	Klaster	Nilai SI	Klaster	Nilai SI	Klaster	Nilai SI	Klaster
S85	0.2170	1	-0.0328	1	0.3653	3	0.2723	1
S86	0.6223	1	0.5695	2	0.4326	2	0.0576	3
S87	0.6904	1	0.6242	2	0.5426	2	0.5047	3
S88	0.7148	1	0.6624	2	0.5604	2	0.3441	3
S89	0.7342	1	0.6846	2	0.5972	2	0.4356	3
S90	0.7679	1	0.7404	2	0.6386	2	0.3323	4
S91	0.7368	1	0.6985	2	0.5848	2	0.0222	4
S92	0.7076	1	0.6477	2	0.5591	2	0.4960	3
S93	0.7562	1	0.7257	2	0.6106	2	0.2436	4
S94	0.7100	1	0.6638	2	0.5364	2	-0.0158	3
S95	0.7264	1	0.6748	2	0.5885	2	0.4490	3
S96	0.6928	1	0.6589	2	0.5033	2	0.4087	4
S97	0.6607	2	0.4990	3	0.4796	4	0.4513	5
S98	0.4678	2	0.3146	1	0.3541	1	0.3841	2
S99	0.6048	2	0.2248	3	0.1936	4	0.1544	5
S100	0.4170	2	0.2118	3	0.2165	4	0.1909	5
S101	0.5662	2	0.4361	1	0.4756	1	0.5155	2
S102	0.6461	2	0.4202	3	0.3968	4	0.3636	5
S103	0.1149	1	0.1403	1	0.3508	3	0.3314	1
S104	0.5436	2	0.3864	1	0.4256	1	0.4654	2
S105	0.5633	2	0.4502	1	0.4907	1	0.5322	2
S106	0.6033	2	0.1473	3	0.1087	4	-0.1062	2
S107	0.5732	2	0.4326	1	0.4726	1	0.5148	2
S108	0.6781	2	0.5124	3	0.4934	4	0.4634	5
S109	0.7120	1	0.6561	2	0.5716	2	0.4809	3
S110	0.6164	1	0.5364	2	0.4648	2	0.4727	3
S111	0.6013	1	0.5317	2	0.4111	2	0.0770	3
S112	0.6468	1	0.5681	2	0.5107	2	0.5164	3
S113	0.7324	1	0.6876	2	0.5924	2	0.2794	3
S114	0.4461	1	0.3434	2	0.32	2	0.2523	3
S115	0.5200	1	0.4351	2	0.3922	2	0.3488	3
S116	0.7628	1	0.7341	2	0.6237	2	0.3091	4
S117	0.7495	1	0.7144	2	0.6132	2	-0.0263	4
S118	0.2242	1	-0.0387	1	-0.081	1	0.2016	3
S119	0.7340	1	0.7050	2	0.5641	2	0.4493	4
S120	0.4635	2	-0.1180	1	-0.093	1	0.0601	5
S121	0.6783	2	0.5307	3	0.5126	4	0.4853	5
S122	0.4761	2	0.4096	3	0.4152	4	0.3940	5
S123	0.4359	2	0.3430	1	0.3797	1	0.4098	2
S124	0.6816	2	0.5179	3	0.4987	4	0.4680	5
S125	0.6865	2	0.5234	3	0.5039	4	0.4741	5
S126	0.7047	1	0.6684	2	0.536	2	0.3027	4
S127	0.7498	1	0.7227	2	0.6068	2	0.4484	4

	<i>k</i> = 2		<i>k</i> = 3		<i>k</i> = 4		<i>k</i> = 5	
	Nilai SI	Klaster	Nilai SI	Klaster	Nilai SI	Klaster	Nilai SI	Klaster
S128	0.7544	1	0.7266	2	0.6149	2	0.4214	4
S129	0.6132	1	0.5632	2	0.4174	2	0.3077	4
S130	0.7414	1	0.7145	2	0.592	2	0.4946	4
S131	0.7532	1	0.7270	2	0.6149	2	0.4719	4
S132	0.5932	1	0.5333	2	0.4012	2	0.1322	4
S133	0.7468	1	0.7183	2	0.6058	2	0.3994	4
S134	0.7391	1	0.7113	2	0.5755	2	0.4997	4
S135	0.7575	1	0.7307	2	0.6185	2	0.4556	4
S136	0.7669	1	0.7365	2	0.6363	2	0.1196	4
S137	0.4453	1	0.2579	2	0.422	3	0.3857	1
S138	0.4067	2	0.2261	3	0.2319	4	0.2080	5
S139	0.3359	1	0.0957	2	0.4615	3	0.3845	1
S140	0.3842	1	0.1620	2	0.4575	3	0.3893	1
S141	0.6025	2	0.3813	3	0.3559	4	0.3301	5
S142	0.6123	2	0.3818	3	0.362	4	0.3286	5
S143	0.3660	1	0.2300	2	0.1879	3	0.1120	1
S144	0.2709	2	0.4251	1	0.3609	1	0.3777	2
S145	0.4919	2	0.4394	3	0.4437	4	0.4241	5
S146	0.3874	2	0.3402	1	0.3643	1	0.3898	2
S147	0.3759	2	0.3839	1	0.4477	1	0.4642	2
S148	0.6512	2	0.4648	3	0.4449	4	0.4239	5
S149	0.6213	2	0.4221	3	0.3975	4	0.3712	5
S150	0.4904	1	0.3590	2	0.305	3	0.2048	1
S151	0.5737	2	0.4105	1	0.4523	1	0.4929	2
S152	0.6029	2	0.2134	3	0.1818	4	0.1422	5
S153	0.6025	2	0.2245	3	0.1938	4	0.1557	5
S154	0.4877	2	0.2940	1	0.3356	1	0.3685	2
S155	0.4179	2	0.2134	3	0.2181	4	0.1930	5
S156	0.3320	2	0.3861	1	0.3927	1	0.4137	2
S157	0.4373	1	0.2560	2	0.4201	3	0.3996	1
S158	0.4914	2	0.4385	3	0.4427	4	0.4227	5
S159	0.3458	1	0.1699	2	0.3844	3	0.3878	1
S160	0.4300	1	0.2326	2	0.4263	3	0.3808	1
S161	0.4329	2	0.4916	1	0.522	1	0.5450	2
S162	0.2832	1	0.1899	2	0.153	3	0.0679	1
S163	0.0355	1	0.1885	1	0.278	3	0.2561	1
S164	0.3033	2	0.3469	3	0.2347	4	0.2230	5
S165	0.6416	2	0.4600	3	0.4397	4	0.4183	5

	$k = 2$		$k = 3$		$k = 4$		$k = 5$	
	Nilai MSV	Klaster	Nilai MSV	Klaster	Nilai MSV	Klaster	Nilai MSV	Klaster
S1	0.9107	1	0.8954	2	0.7976	2	0.0038	4
S2	0.7424	1	0.6119	2	0.3551	2	0.3011	3
S3	0.9134	1	0.9217	2	0.8409	2	0.4289	4
S4	0.8418	1	0.8151	2	0.6786	2	0.5045	4
S5	0.8830	1	0.8026	2	0.5819	2	0.7570	3
S6	0.8596	1	0.8356	2	0.7083	2	0.5801	4
S7	0.8597	1	0.7494	2	0.5787	2	0.3691	3
S8	0.7789	1	0.7376	2	0.5664	2	0.4358	4
S9	0.9107	1	0.8179	2	0.6285	2	0.8367	3
S10	0.8639	1	0.7442	2	0.5200	2	0.7102	3
S11	0.8205	1	0.8008	2	0.6538	2	0.4736	4
S12	0.8902	1	0.8053	2	0.6577	2	0.6469	3
S13	0.8284	1	0.7951	2	0.6522	2	0.5312	4
S14	0.7951	1	0.7578	2	0.5924	2	0.6100	4
S15	0.9519	1	0.9176	2	0.8143	2	0.3080	3
S16	0.7752	1	0.6702	2	0.4790	2	0.1130	3
S17	0.7814	1	0.7404	2	0.5676	2	0.5403	4
S18	0.8935	1	0.8822	2	0.7836	2	0.5904	4
S19	1.0000	1	0.9356	2	0.8667	2	0.2862	3
S20	0.5984	2	0.7786	1	0.7786	1	0.8843	2
S21	0.3441	2	0.5003	1	0.4755	1	0.4440	2
S22	0.6459	2	0.7504	1	0.7504	1	0.8570	2
S23	0.5809	2	1.0000	1	1.0000	1	0.9046	2
S24	0.2577	2	0.4759	1	0.3554	1	0.3123	2
S25	0.6495	2	0.8379	1	0.8379	1	0.9221	2
S26	0.7700	2	0.4677	3	0.4677	4	0.4472	5
S27	0.5811	2	0.8789	1	0.8789	1	0.8486	2
S28	0.6397	2	0.6774	1	0.6774	1	0.8023	2
S29	0.5939	2	0.8790	1	0.8790	1	0.8813	2
S30	0.5698	2	0.7019	1	0.7019	1	0.7142	2
S31	0.5493	2	0.0181	3	0.0181	4	0.0054	5
S32	0.7693	2	0.5533	3	0.5533	4	0.5276	5
S33	0.5861	2	0.7288	1	0.7288	1	0.7544	2
S34	0.7602	2	0.4765	3	0.4765	4	0.4625	5
S35	0.6314	2	0.2906	3	0.2906	4	0.2857	5
S36	0.7677	2	0.5567	3	0.5567	4	0.5348	5
S37	0.8305	1	0.7186	2	0.5643	2	0.1591	4
S38	0.8635	1	0.8431	2	0.7220	2	0.3101	4
S39	0.9022	1	0.8213	2	0.7075	2	0.3387	3
S40	0.6029	1	0.5260	2	0.2969	2	0.0303	4
S41	0.9123	1	0.8292	2	0.6605	2	0.5494	3

	<i>k</i> = 2		<i>k</i> = 3		<i>k</i> = 4		<i>k</i> = 5	
	Nilai MSV	Klaster	Nilai MSV	Klaster	Nilai MSV	Klaster	Nilai MSV	Klaster
S42	0.8964	1	0.8838	2	0.7839	2	0.5616	4
S43	0.8870	1	0.8764	2	0.7672	2	1.0000	4
S44	0.8785	1	0.8443	2	0.6695	2	0.4198	3
S45	0.8653	1	0.7543	2	0.5814	2	0.7339	3
S46	0.8125	1	0.6732	2	0.4382	2	0.6048	3
S47	0.8604	1	0.7419	2	0.5338	2	0.7386	3
S48	0.8577	1	0.7416	2	0.5325	2	0.8043	3
S49	0.8913	1	0.8862	2	0.7837	2	0.5890	4
S50	0.9215	1	0.9341	2	0.8741	2	0.4781	4
S51	0.8469	1	0.7104	2	0.4874	2	0.7544	3
S52	0.9005	1	0.7993	2	0.6005	2	1.0000	3
S53	0.9076	1	0.9042	2	0.8142	2	0.6658	4
S54	0.0063	2	0.2141	1	0.1105	1	0.0892	2
S55	0.4886	2	0.3720	1	0.3720	1	0.4149	2
S56	0.3666	2	0.7196	1	0.6063	1	0.5735	2
S57	0.9713	2	0.8611	3	0.8611	4	0.8577	5
S58	0.9874	2	0.8444	3	0.8444	4	0.8321	5
S59	0.4787	2	0.1318	3	0.1318	4	0.0817	5
S60	0.9667	2	0.9286	3	0.9286	4	0.9207	5
S61	0.9607	2	1.0000	3	1.0000	4	1.0000	5
S62	0.9789	2	0.8817	3	0.8817	4	0.8729	5
S63	0.5184	2	0.8592	1	0.8581	1	0.8141	2
S64	0.8906	2	0.5579	3	0.5579	4	0.5771	5
S65	0.4750	2	0.7918	1	0.7790	1	0.7714	2
S66	0.4505	2	0.4926	1	0.4926	1	0.4768	2
S67	0.9644	2	0.9209	3	0.9209	4	0.9138	5
S68	0.5682	1	0.2629	2	0.5386	3	0.5463	1
S69	0.9259	2	0.6614	3	0.6614	4	0.6539	5
S70	0.8480	2	0.5545	3	0.5545	4	0.5596	5
S71	1.0000	2	0.8615	3	0.8615	4	0.8504	5
S72	0.7887	2	0.4522	3	0.4522	4	0.4430	5
S73	0.7303	2	0.0935	3	0.0935	4	0.0395	5
S74	0.7228	2	0.4680	3	0.4680	4	0.4594	5
S75	0.9695	2	0.8368	3	0.8368	4	0.8315	5
S76	0.3955	2	0.4735	1	0.4735	1	0.4762	2
S77	0.3109	2	0.4891	1	0.4293	1	0.4012	2
S78	0.5668	2	0.1360	3	0.1360	4	0.0885	5
S79	0.7146	2	0.1067	1	0.1067	1	0.1251	2
S80	0.9203	2	0.7260	3	0.7260	4	0.7142	5
S81	0.9119	2	0.7626	3	0.7626	4	0.7490	5
S82	0.4945	2	0.7762	1	0.7762	1	0.7889	2
S83	0.6894	2	0.1296	3	0.1296	4	0.0776	5
S84	0.2389	1	0.3349	1	0.6154	3	0.6275	1

	<i>k</i> = 2		<i>k</i> = 3		<i>k</i> = 4		<i>k</i> = 5	
	Nilai MSV	Klaster	Nilai MSV	Klaster	Nilai MSV	Klaster	Nilai MSV	Klaster
S85	0.2622	1	0.1420	1	0.4364	3	0.4470	1
S86	0.7284	1	0.6672	2	0.4550	2	0.1237	3
S87	0.8317	1	0.6914	2	0.4449	2	0.7277	3
S88	0.8381	1	0.7800	2	0.5638	2	0.5525	3
S89	0.8882	1	0.7853	2	0.5594	2	0.7245	3
S90	0.9352	1	0.9470	2	0.8949	2	0.5949	4
S91	0.8173	1	0.7920	2	0.5700	2	0.0059	4
S92	0.8735	1	0.7567	2	0.4998	2	0.7986	3
S93	0.8901	1	0.8887	2	0.7890	2	0.1429	4
S94	0.7931	1	0.7539	2	0.5515	2	0.1233	3
S95	0.8988	1	0.8035	2	0.6169	2	0.8480	3
S96	0.7856	1	0.7460	2	0.5770	2	0.5362	4
S97	0.9150	2	0.8244	3	0.8244	4	0.8133	5
S98	0.4167	2	0.5060	1	0.5060	1	0.5084	2
S99	0.7204	2	0.2335	3	0.2335	4	0.1534	5
S100	0.4787	2	0.1286	3	0.1286	4	0.0790	5
S101	0.6308	2	0.8765	1	0.8765	1	0.8672	2
S102	0.9342	2	0.6852	3	0.6852	4	0.6809	5
S103	0.1474	1	0.3781	1	0.5102	3	0.5075	1
S104	0.6088	2	0.6860	1	0.6860	1	0.7409	2
S105	0.6290	2	0.8945	1	0.8945	1	1.0000	2
S106	0.7309	2	0.0030	3	0.0030	4	0.0583	2
S107	0.6466	2	0.8244	1	0.8244	1	0.8602	2
S108	0.9640	2	0.8509	3	0.8509	4	0.8459	5
S109	0.8824	1	0.7803	2	0.5573	2	0.8589	3
S110	0.7622	1	0.6172	2	0.4205	2	0.6690	3
S111	0.7178	1	0.6196	2	0.3500	2	0.1055	3
S112	0.7818	1	0.6140	2	0.4039	2	0.6856	3
S113	0.8470	1	0.8024	2	0.5743	2	0.4526	3
S114	0.5194	1	0.3271	2	0.2126	2	0.3228	3
S115	0.6045	1	0.3905	2	0.2684	2	0.4121	3
S116	0.9082	1	0.9135	2	0.8356	2	0.3780	4
S117	0.8931	1	0.8906	2	0.7855	2	0.0153	4
S118	0.4297	1	0.1063	1	0.1063	1	0.1607	3
S119	0.8236	1	0.7898	2	0.6448	2	0.4841	4
S120	0.5533	2	0.0003	1	0.0003	1	0.0036	5
S121	0.9374	2	0.9076	3	0.9076	4	0.8987	5
S122	0.7127	2	0.4627	3	0.4627	4	0.4526	5
S123	0.4497	2	0.4205	1	0.4205	1	0.4511	2
S124	0.9606	2	0.8916	3	0.8916	4	0.8839	5
S125	0.9729	2	0.9330	3	0.9330	4	0.9265	5
S126	0.8207	1	0.7688	2	0.6006	2	0.4524	4
S127	0.8511	1	0.8288	2	0.6894	2	0.6807	4

	<i>k</i> = 2		<i>k</i> = 3		<i>k</i> = 4		<i>k</i> = 5	
	Nilai MSV	Klaster	Nilai MSV	Klaster	Nilai MSV	Klaster	Nilai MSV	Klaster
S128	0.8714	1	0.8531	2	0.7373	2	0.5921	4
S129	0.6769	1	0.6170	2	0.4280	2	0.3794	4
S130	0.8252	1	0.7954	2	0.6406	2	0.7134	4
S131	0.8515	1	0.8290	2	0.6921	2	0.6989	4
S132	0.6934	1	0.5759	2	0.3842	2	0.0934	4
S133	0.8553	1	0.8343	2	0.6977	2	0.7180	4
S134	0.8463	1	0.8228	2	0.6832	2	0.6086	4
S135	0.8707	1	0.8508	2	0.7298	2	0.6111	4
S136	0.9558	1	1.0000	2	1.0000	2	0.1840	4
S137	0.5386	1	0.2610	2	0.7917	3	0.7882	1
S138	0.4665	2	0.1349	3	0.1349	4	0.0872	5
S139	0.4575	1	0.0162	2	0.8304	3	0.8049	1
S140	0.5034	1	0.0896	2	1.0000	3	1.0000	1
S141	0.7661	2	0.4311	3	0.4311	4	0.4310	5
S142	0.8993	2	0.5684	3	0.5684	4	0.5972	5
S143	0.4042	1	0.2375	2	0.2105	3	0.2093	1
S144	0.1912	2	0.6248	1	0.5444	1	0.4964	2
S145	0.7084	2	0.5398	3	0.5398	4	0.5140	5
S146	0.4247	2	0.5468	1	0.5468	1	0.5426	2
S147	0.3570	2	0.5051	1	0.5044	1	0.4546	2
S148	0.9222	2	0.7328	3	0.7328	4	0.7275	5
S149	0.7752	2	0.5526	3	0.5526	4	0.5265	5
S150	0.5437	1	0.3690	2	0.4203	3	0.3334	1
S151	0.6427	2	0.7728	1	0.7728	1	0.8513	2
S152	0.7259	2	0.1617	3	0.1617	4	0.0701	5
S153	0.7179	2	0.2401	3	0.2401	4	0.1616	5
S154	0.4477	2	0.4800	1	0.4800	1	0.4699	2
S155	0.4779	2	0.1347	3	0.1347	4	0.0856	5
S156	0.3068	2	0.5005	1	0.4196	1	0.3734	2
S157	0.5784	1	0.3054	2	0.7459	3	0.7631	1
S158	0.7067	2	0.5379	3	0.5379	4	0.5118	5
S159	0.3310	1	0.0170	2	0.5631	3	0.5518	1
S160	0.5955	1	0.2398	2	0.7301	3	0.7099	1
S161	0.4491	2	0.7896	1	0.7539	1	0.7117	2
S162	0.3280	1	0.1878	2	0.1562	3	0.0603	1
S163	0.0487	1	0.4204	1	0.4063	3	0.3704	1
S164	0.6587	2	0.4746	3	0.3488	4	0.3488	5
S165	0.8895	2	0.7364	3	0.7364	4	0.7293	5

Kode Sekolah	Nama Sekolah	Klaster ($k = 2$)	Klaster ($k = 3$)	Klaster ($k = 4$)	Klaster ($k = 5$)
S1	SMA Negeri 1 Pundong	1	2	2	4
S2	SMA Negeri 3 Bantul	1	2	2	3
S3	SMA Negeri 1 Bambanglipuro	1	2	2	4
S4	SMA Negeri 1 Sedayu	1	2	2	4
S5	SMA Negeri 1 Kretek	1	2	2	3
S6	SMA Negeri 1 Jetis	1	2	2	4
S7	SMA Negeri 1 Pleret	1	2	2	3
S8	SMA Negeri 1 Banguntapan	1	2	2	4
S9	SMA Negeri 1 Dlingo	1	2	2	3
S10	SMA Negeri 1 Srandakan	1	2	2	3
S11	SMA Negeri 1 Sewon	1	2	2	4
S12	SMA Negeri 1 Pajangan	1	2	2	3
S13	SMA Negeri 1 Kasihan	1	2	2	4
S14	SMA Negeri 2 Bantul	1	2	2	4
S15	SMA Negeri 1 Piyungan	1	2	2	3
S16	SMA Negeri 1 Imogiri	1	2	2	3
S17	SMA Negeri 1 Bantul	1	2	2	4
S18	SMA Negeri 2 Banguntapan	1	2	2	4
S19	SMA Negeri 1 Sanden	1	2	2	3
S20	SMA Muhammadiyah Pleret	2	1	1	2
S21	SMA Ali Maksum	2	1	1	2
S22	SMA BOPKRI Banguntapan	2	1	1	2
S23	SMA UII Banguntapan	2	1	1	2
S24	SMA Pangudi Luhur Sedayu	2	1	1	2
S25	SMA Stella Duce Bambanglipuro	2	1	1	2
S26	SMA Tumbuh	2	3	4	5
S27	SMA Muhammadiyah Bantul	2	1	1	2
S28	SMA Muhammadiyah Piyungan	2	1	1	2
S29	SMA Muhammadiyah Imogiri	2	1	1	2
S30	SMAS Muhammadiyah Sewon	2	1	1	2
S31	SMA Mutiara Persada	2	3	4	5
S32	SMA 17 Bantul	2	3	4	5
S33	SMA PGRI Kasihan	2	1	1	2
S34	SMA Muhammadiyah Kasihan	2	3	4	5
S35	SMA Dharma Amiluhur	2	3	4	5
S36	SMA Patria Bantul	2	3	4	5
S37	SMAN 1 Seyegan	1	2	2	4
S38	SMAN 1 Ngaglik	1	2	2	4
S39	SMAN 1 Prambanan	1	2	2	3
S40	SMAN 1 Turi	1	2	2	4
S41	SMAN 1 Pakem	1	2	2	3

Kode Sekolah	Nama Sekolah	Klaster ($k = 2$)	Klaster ($k = 3$)	Klaster ($k = 4$)	Klaster ($k = 5$)
S42	SMAN 1 Sleman	1	2	2	4
S43	SMAN 1 Kalasan	1	2	2	4
S44	SMAN 1 Mlati	1	2	2	3
S45	SMAN 1 Cangkringan	1	2	2	3
S46	SMAN 1 Tempel	1	2	2	3
S47	SMAN 1 Gamping	1	2	2	3
S48	SMAN 2 Sleman	1	2	2	3
S49	SMAN 1 Godean	1	2	2	4
S50	SMAN 2 Ngaglik	1	2	2	4
S51	SMAN 1 Minggir Sleman	1	2	2	3
S52	SMAN 1 Ngemplak	1	2	2	3
S53	SMAN 1 Depok	1	2	2	4
S54	SMAS Kolese De Britto Depok	2	1	1	2
S55	SMAS Institut Indonesia Berbah	2	1	1	2
S56	SMA Islam Al Azhar 9 Yogyakarta	2	1	1	2
S57	SMAS Muhammadiyah Pakem	2	3	4	5
S58	SMAS Islam Prambanan	2	3	4	5
S59	SMA Insan Mulia Boarding School	2	3	4	5
S60	SMA Islam 3 Pakem	2	3	4	5
S61	SMAS Ma'arif 1 Sleman	2	3	4	5
S62	SMAS Immanuel Kalasan	2	3	4	5
S63	SMAS Muhammadiyah 1 Prambanan	2	1	1	2
S64	SMAS Islam 1 Sleman	2	3	4	5
S65	SMAS Terpadu Bina Umat	2	1	1	2
S66	SMAS Santo Mikael Mlati	2	1	1	2
S67	SMAS Muhammadiyah Kalasan	2	3	4	5
S68	SMAS Muhammadiyah Boarding School (MBS) Sleman	1	2	3	1
S69	SMA Islam Terpadu Baitussalam	2	3	4	5
S70	SMAS Sunan Kalijogo	2	3	4	5
S71	SMAS Muhammadiyah 1 Sleman	2	3	4	5
S72	SMAS Gama Depok Yogyakarta	2	3	4	5
S73	SMAS Cahaya Bangsa Utama	2	3	4	5
S74	SMAS Darul Hikmah	2	3	4	5
S75	SMAS Sulaiman Sleman	2	3	4	5
S76	SMA Budi Utama	2	1	1	2
S77	SMAS Angkasa Adisucipto	2	1	1	2
S78	SMAS Mandala Bhakti	2	3	4	5
S79	SMA Olifant	2	1	1	2
S80	SMAS IKIP Veteran	2	3	4	5
S81	SMAS Muhammadiyah Mlati	2	3	4	5
S82	SMA Sains Wahid Hasyim	2	1	1	2
S83	SMAS Dr Wahidin Mlati	2	3	4	5
S84	SMAS Kolombo Depok	1	1	3	1

Kode Sekolah	Nama Sekolah	Klaster ($k = 2$)	Klaster ($k = 3$)	Klaster ($k = 4$)	Klaster ($k = 5$)
S85	SMA Budi Mulia Dua	1	1	3	1
S86	SMAN 1 Tanjungsari	1	2	2	3
S87	SMAN 1 Rongkop	1	2	2	3
S88	SMAN 1 Semanu	1	2	2	3
S89	SMAN 1 Patuk	1	2	2	3
S90	SMAN 2 Wonosari	1	2	2	4
S91	SMAN 2 Playen	1	2	2	4
S92	SMAN 1 Playen	1	2	2	3
S93	SMAN 1 Semin	1	2	2	4
S94	SMAN 1 Karangmojo	1	2	2	3
S95	SMAN 1 Panggang	1	2	2	3
S96	SMAN 1 Wonosari	1	2	2	4
S97	SMA Gotong Royong Semin	2	3	4	5
S98	SMA IKIP Veteran Iii Tambakromo	2	1	1	2
S99	SMA PGRI Playen	2	3	4	5
S100	SMA Muhammadiyah Daarul Khoir Gunungkidul	2	3	4	5
S101	SMA Dominikus Wonosari	2	1	1	2
S102	SMA Muhammadiyah Ngawen	2	3	4	5
S103	SMA Pembangunan 2 Karangmojo	1	1	3	1
S104	SMA Muhammadiyah Ponjong	2	1	1	2
S105	SMA Muhammadiyah Al Mujahidin Wonosari	2	1	1	2
S106	SMA Baitul Qur'an Yogyakarta	2	3	4	2
S107	SMA Pembangunan 3 Ponjong	2	1	1	2
S108	SMAS Pembangunan 1 Wonosari	2	3	4	5
S109	SMA Negeri 1 Temon	1	2	2	3
S110	SMA Negeri 1 Galur	1	2	2	3
S111	SMA Negeri 1 Lendah	1	2	2	3
S112	SMA Negeri 1 Girimulyo	1	2	2	3
S113	SMA Negeri 1 Kalibawang	1	2	2	3
S114	SMA Negeri 1 Kokap	1	2	2	3
S115	SMA Negeri 1 Samigaluh	1	2	2	3
S116	SMA Negeri 1 Pengasih	1	2	2	4
S117	SMA Negeri 1 Sentolo	1	2	2	4
S118	SMA Negeri 2 Wates	1	1	1	3
S119	SMA Negeri 1 Wates	1	2	2	4
S120	SMA IT Al Ghifari Kulon Progo	2	1	1	5
S121	SMA Muhammadiyah Al Manar Galur	2	3	4	5
S122	SMA Muhammadiyah Wates	2	3	4	5
S123	SMA IT Abu Bakar Boarding School Kulon Progo	2	1	1	2
S124	SMA Maarif Wates	2	3	4	5
S125	SMA Sanjaya XIV Nanggulan	2	3	4	5
S126	SMAN 1 Yogyakarta	1	2	2	4
S127	SMAN 8 Yogyakarta	1	2	2	4

Kode Sekolah	Nama Sekolah	Klaster ($k = 2$)	Klaster ($k = 3$)	Klaster ($k = 4$)	Klaster ($k = 5$)
S128	SMAN 6 Yogyakarta	1	2	2	4
S129	SMAN 2 Yogyakarta	1	2	2	4
S130	SMAN 9 Yogyakarta	1	2	2	4
S131	SMAN 5 Yogyakarta	1	2	2	4
S132	SMAN 11 Yogyakarta	1	2	2	4
S133	SMAN 3 Yogyakarta	1	2	2	4
S134	SMAN 4 Yogyakarta	1	2	2	4
S135	SMAN 7 Yogyakarta	1	2	2	4
S136	SMAN 10 Yogyakarta	1	2	2	4
S137	SMAS Stella Duce 2	1	2	3	1
S138	SMAS Gotong Royong	2	3	4	5
S139	SMAS Muhammadiyah 4 Yogyakarta	1	2	3	1
S140	SMAS Muhammadiyah 7 Yogyakarta	1	2	3	1
S141	SMAS Gadjah Mada	2	3	4	5
S142	SMAS Budaya Wacana	2	3	4	5
S143	SMAS Muhammadiyah 2 Yogyakarta	1	2	3	1
S144	SMAS BOPKRI 1	2	1	1	2
S145	SMA 17 I Yogyakarta	2	3	4	5
S146	SMAS Katolik Sang Timur	2	1	1	2
S147	SMAS Muhammadiyah 5 Yogyakarta	2	1	1	2
S148	SMAS Maarif Yogyakarta	2	3	4	5
S149	SMAS Santo Thomas	2	3	4	5
S150	SMAS Stella Duce 1	1	2	3	1
S151	SMAS Muhammadiyah 6 Yogyakarta	2	1	1	2
S152	SMAS Budi Luhur	2	3	4	5
S153	SMAS Sultan Agung	2	3	4	5
S154	SMAS Piri 1 Yogyakarta	2	1	1	2
S155	SMAS Perak	2	3	4	5
S156	SMAS Santa Maria	2	1	1	2
S157	SMAS Muhammadiyah 3 Yogyakarta	1	2	3	1
S158	SMAS Bhinneka Tunggal Ika	2	3	4	5
S159	SMAS Pangudi Luhur	1	2	3	1
S160	SMAS IT Abu Bakar	1	2	3	1
S161	SMAS BOPKRI 2	2	1	1	2
S162	SMAS Muhammadiyah 1 Yogyakarta	1	2	3	1
S163	SMAS Taman Madya IP	1	1	3	1
S164	SMAS Taman Madya Jetis	2	3	4	5
S165	SMAS IT Bina Anak Sholeh	2	3	4	5

```

skm.w <- function(distdata, ncluster, seeding = 20,
iterate = 10) {

  if(any(is.na(distdata))) stop("Cannot handle missing
values!")

  if((is.matrix(distdata)||inherits(distdata,
"dist"))==FALSE)
    stop("The distdata must be a matrix or a dist
object!")

  if(is.matrix(distdata)==TRUE) {
    nr <- nrow(distdata)
    nc <- ncol(distdata)
    if (nc!=nr) stop("The distdata is not an x n
distance matrix!")
  }

  if(inherits(distdata, "dist")) distdata <-
as.matrix(distdata)

  if(is.null(rownames(distdata))==TRUE)
rownames(distdata) <- 1:nrow(distdata)

  n <- nrow(distdata)
  index <- 1:n
  colnames(distdata) <- rownames(distdata) <- 1:n

  distsum <- colSums(distdata/sum(distdata))
  names(distsum) <- index
  iduniq <- as.numeric(names(distsum
[!duplicated(distsum)]))
  distuniq <- unique(distsum)
  names(distuniq) <- iduniq
  sorted_object <- as.numeric(names(sort(distuniq)))
  init1 <- sorted_object[1]
  noninit1 <- sorted_object[-1]
  rownames(distdata)<-NULL
  colnames(distdata)<-NULL
  z <- vector("list", seeding)

```

```

    for (i in 1:seeding) {
      medinit <- c(init1, sample(noninit1, ncluster-1))
      z[[i]] <- fastkmed.x(distdata, ncluster, iterate,
init = medinit)
    }

    min_dist <- numeric(seeding)
    for (i in 1:seeding) {
      min_dist[i] <- z[[i]]$minimum_distance
    }
    result <- z[[which.min(min_dist)]]
    return(result)
}

```

```

park.fastkmed<-function(distdata, ncluster, iterate =
10, init = NULL) {

  if(any(is.na(distdata))) stop("Cannot handle missing
values!")

  if((is.matrix(distdata)||inherits(distdata,
"dist"))==FALSE)
    stop("The distdata must be a matrix or a dist
object!")

  if(is.matrix(distdata)==TRUE) {
    nr <- nrow(distdata)
    nc <- ncol(distdata)
    if (nc!=nr) stop("The distdata is not an x n
distance matrix!")
  }

  if(inherits(distdata, "dist")) distdata <-
as.matrix(distdata)

  if(is.null(rownames(distdata))==TRUE)
rownames(distdata) <- 1:nrow(distdata)

  n <- nrow(distdata)
  index <- 1:n

  if (is.null(init) == TRUE) {
    distsum <- colSums(distdata/sum(distdata))
    names(distsum) <- index
    iduniq <- as.numeric(names(distsum
[!duplicated(distsum)]))
    distuniq <- unique(distsum)
    names(distuniq) <- iduniq
    sorted_object <- as.numeric(names(sort(distuniq)))
    medoid_init <- sorted_object[1:ncluster]
    if (sum(!is.na(medoid_init)) < ncluster)
      stop("An empty cluster is present, change the
number of clusters or initial medoids")

  } else {
    if (ncluster != length(init))
      stop("The initial medoids must equal to the
number of clusters")
    if (sum(is.na(init)) > 0)
      stop("The initial medoids consist of NA")
  }
}

```

```

    medoid_init <- init
  }

  dist_0 <- distdata[,medoid_init, drop = FALSE]
  for (i in 1:ncluster) {
    dist_0[names(dist_0[,i]) == colnames(dist_0)[i],
i] <- -1
  }
  member_0 <- apply(dist_0, 1, which.min)
  E0 <- sum(apply(dist_0, 1, min))

  if (length(unique(member_0)) != ncluster)
    stop("An empty cluster is present, change the
number of clusters or initial medoids")

  row_0 <- row_1 <- medoid_0 <- medoid_1 <-
numeric(ncluster)

  for (i in 1:ncluster) {
    row_0[i] <-
which.min(rowSums(distdata[member_0==i,member_0==i,
drop = FALSE]))
    medoid_0[i] <- index[member_0==i][row_0[i]]
  }

  dist_1 <- distdata[,medoid_0, drop = FALSE]
  for (i in 1:ncluster) {
    dist_1[names(dist_1[,i]) == colnames(dist_1)[i],
i] <- -1
  }
  member_1 <- apply(dist_1, 1, which.min)
  E1 <- sum(apply(dist_1, 1, min))

  for (i in 1:ncluster) {
    row_1[i] <-
which.min(rowSums(distdata[member_1==i,member_1==i,
drop = FALSE]))
    medoid_1[i] <- index[member_1==i][row_1[i]]
  }

  if (sum(sort(medoid_0) != sort(medoid_1))!=0) {

    x <- 1
    medoid <- vector("list", iterate)
    E <- numeric(iterate)
    repeat {
      dist_0 <- dist_1
      member_0 <- member_1

```

```

E0 <- E1
row_0 <- row_1
medoid_0 <- medoid_1

dist_1 <- distdata[,medoid_0, drop = FALSE]
for (i in 1:ncluster) {
  dist_1[names(dist_1[,i]) ==
colnames(dist_1)[i], i] <- -1
}
member_1 <- apply(dist_1, 1, which.min)

for (i in 1:ncluster) {
  row_1[i] <-
which.min(rowSums(distdata[member_1==i,member_1==i,
drop = FALSE]))
  medoid_1[i] <- index[member_1==i][row_1[i]]
}
E1 <- sum(apply(dist_1, 1, min))

medoid[[x]] <- medoid_1
E[x] <- E1
if (x == iterate || sum(sort(medoid_0) !=
sort(medoid_1))==0) { break }
x <- x + 1

}
medoid <- do.call(rbind, medoid)
E <- E[1:x]
index_E <- which.min(E)
medoid_2 <- medoid[index_E,]
dist_2 <- distdata[,medoid_2, drop = FALSE]
for (i in 1:ncluster) {
  dist_2[names(dist_2[,i]) == colnames(dist_2)[i],
i] <- -1
}
member_2 <- apply(dist_2, 1, which.min)
E2 <- sum(apply(dist_2, 1, min))

names(member_2) <- rownames(distdata)

result <- list(cluster = member_2, medoid =
medoid_2, minimum_distance = E2)#,

}
else {
  dist_2 <- distdata[,medoid_1, drop = FALSE]
  member_1 <- apply(dist_2, 1, which.min)
  E2 <- sum(apply(dist_2, 1, min))
  medoid <- medoid_1

```

```
    result <- list(cluster = member_1, medoid = medoid,  
minimum_distance = E2)  
  }  
  return(list(medoid_init, medoid))  
}
```

```

simplekm <- function(distdata, ncluster, seeding = 20,
iterate = 10) {

  if(any(is.na(distdata))) stop("Cannot handle missing
values!")

  if((is.matrix(distdata)||inherits(distdata,
"dist"))==FALSE)
    stop("The distdata must be a matrix or a dist
object!")

  if(is.matrix(distdata)==TRUE) {
    nr <- nrow(distdata)
    nc <- ncol(distdata)
    if (nc!=nr) stop("The distdata is not an x n
distance matrix!")
  }

  if(inherits(distdata, "dist")) distdata <-
as.matrix(distdata)

  if(is.null(rownames(distdata))==TRUE)
rownames(distdata) <- 1:nrow(distdata)

  n <- nrow(distdata)
  index <- 1:n
  colnames(distdata) <- rownames(distdata) <- 1:n

  distsum <- colSums(distdata/sum(distdata))
  names(distsum) <- index
  iduniq <- as.numeric(names(distsum
[!duplicated(distsum)]))
  distuniq <- unique(distsum)
  names(distuniq) <- iduniq
  sorted_object <- as.numeric(names(sort(distuniq)))
  init1 <- sorted_object[1]
  noninit1 <- sorted_object[-1]
  rownames(distdata) <- NULL
  colnames(distdata) <- NULL
  z <- vector("list", seeding)

```

```

    for (i in 1:seeding) {
      medinit <- c(init1, sample(noninit1, ncluster-1))
      z[[i]] <- park.fastkmed(distdata, ncluster,
iterate, init = medinit)
    }
    return(z)
  }
}

```

MENGINPUT DATA

```
skripsiku=read.delim("clipboard")
View(skripsiku)
#Melihat data missing
View(is.na(skripsiku))
#Melihat ringkasan statistika deskriptif
summary(skripsiku)
```

ALGORITMA SIMPLE K-MEDOIDS PARTITIONING

```
#Mengubah data menjadi matriks
mat.skripsiku<-as.matrix(skripsiku[2:14])
View(mat.skripsiku)

#Mengaktifkan package "kmed"
install.packages("kmed")
library(kmed)
#Perhitungan jarak data campuran : distmix
#idnum = kolom data numerik : kolom 1 2 3 4 5 6 7 8
#idbin = kolom data biner : kolom 9 10 11
#idcat = kolom data kategorik : kolom 12 13
skripsiku.dist<-dismix(mat.skripsiku, method =
"gower", idnum = 1:8, idbin = 9:11, idcat = 12:13)
View(skripsiku.dist)
#menyimpan nilai jarak dalam bentuk file csv
write.csv(skripsiku.dist, file="E://KULIAH//TUGAS
AKHIR//SKRIPSI IRSA//SEEDING = 20, ITERATE =
10//Matriks Jarak Skripsiku (2 Klaster, 13
Variabel).csv")

#k=2
set.seed(1)
skm.skripsi.2 <- skm.w(skripsiku.dist, ncluster = 2,
seeding=20, iterate=10)
skm.skripsi.2
#Melihat rata-rata data numerik setiap klaster
install.packages("tidyverse")
library(tidyverse)
skripsiku[,3:10] %>%
  mutate(Cluster = skm.skripsi.2$cluster) %>%
  group_by(Cluster) %>%
```

```

    summarise_all("mean")
#simplekm : untuk melihat medoid awal dan akhir
set.seed(1)
simplekm.skripsi.2 <- simplekm(skripsiku.dist,
ncluster = 2, seeding=20, iterate=10)
simplekm.skripsi.2

#k=3
set.seed(1)
skm.skripsi.3 <- skm.w(skripsiku.dist, ncluster = 3,
seeding=20, iterate=10)
skm.skripsi.3
set.seed(1)
simplekm.skripsi.3 <- simplekm(skripsiku.dist,
ncluster = 3, seeding=20, iterate=10)
simplekm.skripsi.3

#k=4
set.seed(1)
skm.skripsi.4 <- skm.w(skripsiku.dist, ncluster = 4,
seeding=20, iterate=10)
skm.skripsi.4
set.seed(1)
simplekm.skripsi.4 <- simplekm(skripsiku.dist,
ncluster = 4, seeding=20, iterate=10)
simplekm.skripsi.4

#k=5
set.seed(1)
skm.skripsi.5 <- skm.w(skripsiku.dist, ncluster = 5,
seeding=20, iterate=10)
skm.skripsi.5
set.seed(1)
simplekm.skripsi.5 <- simplekm(skripsiku.dist,
ncluster = 5, seeding=20, iterate=10)
simplekm.skripsi.5

#menyimpan HASIL SKM (k=2) dalam bentuk file csv
skripsi.final<-data.frame(skripsiku,
skm.skripsi.2$cluster)
write.csv(skripsi.final, file="E://KULIAH//TUGAS
AKHIR//SKRIPSI IRSA//SEEDING = 20, ITERATE =

```

```
10//Hasil SKM Seeding=20, Iterate=10 (2 Klaster, 13
Variabel).csv")
```

VALIDASI : SILHOUETTE INDEX (SI)

```
#k=2
sil.skripsi.2<-sil(skripsiku.dist,
skm.skripsi.2$medoid, skm.skripsi.2$cluster,
                      title = "Silhouette Index Plot
dengan SKM (k = 2)")
sil.skripsi.2$result
sil.skripsi.2$plot
#menyimpan nilai silhouette dalam bentuk file csv
write.csv(sil.skripsi.2$result,
file="E://KULIAH//TUGAS AKHIR//SKRIPSI IRSA//SEEDING
= 20, ITERATE = 10//EXCEL//Silhouette untuk 2
Klaster.csv")

#k=3
sil.skripsi.3<-sil(skripsiku.dist,
skm.skripsi.3$medoid, skm.skripsi.3$cluster,
                      title = "Silhouette Index Plot
dengan SKM (k = 3)")
sil.skripsi.3$result
sil.skripsi.3$plot
#menyimpan nilai silhouette dalam bentuk file csv
write.csv(sil.skripsi.3$result,
file="E://KULIAH//TUGAS AKHIR//SKRIPSI IRSA//SEEDING
= 20, ITERATE = 10//EXCEL//Silhouette untuk 3
Klaster.csv")

#k=4
sil.skripsi.4<-sil(skripsiku.dist,
skm.skripsi.4$medoid, skm.skripsi.4$cluster,
                      title = "Silhouette Index Plot
dengan SKM (k=4)")
sil.skripsi.4$result
sil.skripsi.4$plot
#menyimpan nilai silhouette dalam bentuk file csv
write.csv(sil.skripsi.4$result,
file="E://KULIAH//TUGAS AKHIR//SKRIPSI IRSA//SEEDING
= 20, ITERATE = 10//EXCEL//Silhouette untuk 4
Klaster.csv")
```

```
#k=5
sil.skripsi.5<-sil(skripsiku.dist,
skm.skripsi.5$medoid, skm.skripsi.5$cluster,
                      title = "Silhouette Index Plot
dengan SKM (k=5)")
sil.skripsi.5$result
sil.skripsi.5$plot
#menyimpan nilai silhouette dalam bentuk file csv
write.csv(sil.skripsi.5$result,
file="E://KULIAH//TUGAS AKHIR//SKRIPSI IRSA//SEEDING
= 20, ITERATE = 10//EXCEL//Silhouette untuk 5
Klaster.csv")
```

VALIDASI : MEDOID SHADOW VALUE (MSV)

```
#k=2
MSV.skripsi.2<-msv(skripsiku.dist,
skm.skripsi.2$medoid, skm.skripsi.2$cluster,
                      title = "Medoid Shadow Value Plot
dengan SKM (k=2)")
MSV.skripsi.2$result
MSV.skripsi.2$plot
#menyimpan nilai MSV dalam bentuk file csv
write.csv(MSV.skripsi.2$result,
file="E://KULIAH//TUGAS AKHIR//SKRIPSI IRSA//SEEDING
= 20, ITERATE = 10//EXCEL//MSV untuk 2 Klaster.csv")
```

```
#k=3
MSV.skripsi.3<-msv(skripsiku.dist,
skm.skripsi.3$medoid, skm.skripsi.3$cluster,
                      title = "Medoid Shadow Value Plot
dengan SKM (k=3)")
MSV.skripsi.3$result
MSV.skripsi.3$plot
#menyimpan nilai MSV dalam bentuk file csv
write.csv(MSV.skripsi.3$result,
file="E://KULIAH//TUGAS AKHIR//SKRIPSI IRSA//SEEDING
= 20, ITERATE = 10//EXCEL//MSV untuk 3 Klaster.csv")
```

```
#k=4
MSV.skripsi.4<-msv(skripsiku.dist,
skm.skripsi.4$medoid, skm.skripsi.4$cluster,
```

```

                                title = "Medoid Shadow Value Plot
dengan SKM (k=4)")
MSV.skripsi.4$result
MSV.skripsi.4$plot
#menyimpan nilai MSV dalam bentuk file csv
write.csv(MSV.skripsi.4$result,
file="E://KULIAH//TUGAS AKHIR//SKRIPSI IRSA//SEEDING
= 20, ITERATE = 10//EXCEL//MSV untuk 4 Klaster.csv")

#k=5
MSV.skripsi.5<-msv(skripsiku.dist,
skm.skripsi.5$medoid, skm.skripsi.5$cluster,
                                title = "Medoid Shadow Value Plot
dengan SKM (k=5)")
MSV.skripsi.5$result
MSV.skripsi.5$plot
#menyimpan nilai MSV dalam bentuk file csv
write.csv(MSV.skripsi.5$result,
file="E://KULIAH//TUGAS AKHIR//SKRIPSI IRSA//SEEDING
= 20, ITERATE = 10//EXCEL//MSV untuk 5 Klaster.csv")

```

Uji Rata-Rata Dua Populasi Independen

```

#Menggunakan statistik uji Z (var.equal=TRUE)
klaster1.TA<-read.delim("clipboard")
View(klaster1.TA)
klaster2.TA<-read.delim("clipboard")
View(klaster2.TA)

t.test(x=klaster1.TA$X1, y=klaster2.TA$X1,
       alternative="two.sided", var.equal=TRUE,
       conf.level = 0.95, paired = FALSE)
t.test(x=klaster1.TA$X2, y=klaster2.TA$X2,
       alternative="two.sided", var.equal=TRUE,
       conf.level = 0.95, paired = FALSE)
t.test(x=klaster1.TA$X3, y=klaster2.TA$X3,
       alternative="two.sided", var.equal=TRUE,
       conf.level = 0.95, paired = FALSE)
t.test(x=klaster1.TA$X4, y=klaster2.TA$X4,
       alternative="two.sided", var.equal=TRUE,
       conf.level = 0.95, paired = FALSE)
t.test(x=klaster1.TA$X5, y=klaster2.TA$X5,
       alternative="two.sided", var.equal=TRUE,

```

```

        conf.level = 0.95, paired = FALSE)
t.test(x=klaster1.TA$X6, y=klaster2.TA$X6,
       alternative="two.sided", var.equal=TRUE,
       conf.level = 0.95, paired = FALSE)
t.test(x=klaster1.TA$X7, y=klaster2.TA$X7,
       alternative="two.sided", var.equal=TRUE,
       conf.level = 0.95, paired = FALSE)
t.test(x=klaster1.TA$X8, y=klaster2.TA$X8,
       alternative="two.sided", var.equal=TRUE,
       conf.level = 0.95, paired = FALSE)
t.test(x=klaster1.TA$DayaListrik,
       y=klaster2.TA$DayaListrik,
       alternative="two.sided", var.equal=TRUE,
       conf.level = 0.95, paired = FALSE)
t.test(x=klaster1.TA$BwUpload,
       y=klaster2.TA$BwUpload,
       alternative="two.sided", var.equal=TRUE,
       conf.level = 0.95, paired = FALSE)
t.test(x=klaster1.TA$BwDownload,
       y=klaster2.TA$BwDownload,
       alternative="two.sided", var.equal=TRUE,
       conf.level = 0.95, paired = FALSE)

```

VALIDASI METODE: KM, SFKM

```
#SFKM
```

```
set.seed(1)
```

```
sfkm.skripsi.2<-fastkmed(skripsiku.dist, ncluster =
2, iterate = 10, init=NULL)
```

```
sfkm.skripsi.2
```

```
#KM
```

```
set.seed(1)
```

```
(kminit<-sample(1:nrow(skripsiku),2))
```

```
km.skripsi.2<-fastkmed(skripsiku.dist, ncluster = 2,
iterate = 10, init=kminit)
```

```
km.skripsi.2
```

- **Uji Rata – Rata Dua Populasi Independen untuk X_1**

Pengujian hipotesis dilakukan untuk mengetahui ada atau tidak adanya perbedaan yang signifikan pada nilai rata-rata jumlah guru antara klaster 1 dan 2 sebagai berikut:

- i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak terdapat perbedaan yang signifikan pada rata – rata jumlah guru antara klaster 1 dan klaster 2)

$H_1 : \mu_1 \neq \mu_2$ (Terdapat perbedaan yang signifikan pada rata – rata jumlah guru antara klaster 1 dan klaster 2)

- ii. Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

- iii. Daerah Kritis

Tolak H_0 jika $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

- iv. Statistik Uji

Diperoleh nilai $Z_{hitung} = 7.4087$

- v. Keputusan

Tolak H_0 karena nilai $Z_{hitung} > Z_{\frac{\alpha}{2}}$ atau $7.4087 > 1.96$

- vi. Kesimpulan

Menggunakan tingkat kepercayaan sebesar 95% maka data yang ada mendukung untuk tolak H_0 artinya terdapat perbedaan yang signifikan pada rata – rata jumlah guru antara klaster 1 dan klaster 2.

- **Uji Rata – Rata Dua Populasi Independen untuk X_2**

Pengujian hipotesis dilakukan untuk mengetahui ada atau tidaknya perbedaan yang signifikan pada nilai rata-rata jumlah siswa antara klaster 1 dan 2 sebagai berikut:

- i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak terdapat perbedaan yang signifikan pada rata – rata jumlah siswa antara klaster 1 dan klaster 2)

$H_1 : \mu_1 \neq \mu_2$ (Terdapat perbedaan yang signifikan pada rata – rata jumlah siswa antara klaster 1 dan klaster 2)

ii. Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii. Daerah Kritis

Tolak H_0 jika $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

iv. Statistik Uji

Diperoleh nilai $Z_{hitung} = 13.811$

v. Keputusan

Tolak H_0 karena nilai $Z_{hitung} > Z_{\frac{\alpha}{2}}$ atau $13.811 > 1.96$

vi. Kesimpulan

Menggunakan tingkat kepercayaan sebesar 95% maka data yang ada mendukung untuk tolak H_0 artinya terdapat perbedaan yang signifikan pada rata – rata jumlah siswa antara klaster 1 dan klaster 2.

• **Uji Rata – Rata Dua Populasi Independen untuk X_3**

Pengujian hipotesis dilakukan untuk mengetahui ada atau tidak adanya perbedaan yang signifikan pada nilai rata – rata jumlah laboratorium komputer antara klaster 1 dan 2 sebagai berikut:

i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak terdapat perbedaan yang signifikan pada rata – rata jumlah laboratorium komputer antara klaster 1 dan klaster 2)

$H_1 : \mu_1 \neq \mu_2$ (Terdapat perbedaan yang signifikan pada rata – rata jumlah laboratorium komputer antara klaster 1 dan klaster 2)

ii. Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii. Daerah Kritis

Tolak H_0 jika $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

iv. Statistik Uji

Diperoleh nilai $Z_{hitung} = 0.32395$

v. Keputusan

Gagal tolak H_0 karena $Z_{hitung} < Z_{\frac{\alpha}{2}}$ atau $0.32395 < 1.96$

vi. Kesimpulan

Menggunakan tingkat kepercayaan sebesar 95% maka data yang ada mendukung untuk gagal tolak H_0 artinya tidak terdapat perbedaan yang signifikan pada rata – rata jumlah laboratorium komputer antara klaster 1 dan klaster 2.

• **Uji Rata – Rata Dua Populasi Independen untuk X_4**

Pengujian hipotesis dilakukan untuk mengetahui ada atau tidak adanya perbedaan yang signifikan pada nilai rata – rata jumlah komputer utama antara klaster 1 dan 2 sebagai berikut:

i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak terdapat perbedaan yang signifikan pada rata – rata jumlah komputer utama antara klaster 1 dan klaster 2)

$H_1 : \mu_1 \neq \mu_2$ (Terdapat perbedaan yang signifikan pada rata – rata jumlah komputer utama antara klaster 1 dan klaster 2)

ii. Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii. Daerah Kritis

Tolak H_0 jika $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

iv. Statistik Uji

Diperoleh nilai $Z_{hitung} = 3.37781$

v. Keputusan

Tolak H_0 karena nilai $Z_{hitung} < Z_{\frac{\alpha}{2}}$ atau $3.7781 > 1.96$

vi. Kesimpulan

Menggunakan tingkat kepercayaan sebesar 95% maka data yang ada mendukung untuk tolak H_0 artinya terdapat perbedaan yang signifikan pada rata – rata jumlah komputer utama / proktor antara klaster 1 dan klaster 2.

• **Uji Rata – Rata Dua Populasi Independen untuk X_5**

Pengujian hipotesis dilakukan untuk mengetahui ada atau tidaknya perbedaan yang signifikan pada nilai rata – rata jumlah komputer sesuai spesifikasi (milik) antara klaster 1 dan 2 sebagai berikut:

i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak terdapat perbedaan yang signifikan pada rata – rata jumlah komputer sesuai spesifikasi (milik) antara klaster 1 dan klaster 2)

$H_1 : \mu_1 \neq \mu_2$ (Terdapat perbedaan yang signifikan pada rata – rata jumlah komputer sesuai spesifikasi (milik) antara klaster 1 dan klaster 2)

ii. Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii. Daerah Kritis

Tolak H_0 jika $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

iv. Statistik Uji

Diperoleh nilai $Z_{hitung} = 13.182$

v. Keputusan

Tolak H_0 karena nilai $Z_{hitung} > Z_{\frac{\alpha}{2}}$ atau $13.182 > 1.96$

vi. Kesimpulan

Menggunakan tingkat kepercayaan sebesar 95% maka data yang ada mendukung tolak H_0 artinya terdapat perbedaan yang signifikan pada rata – rata jumlah komputer sesuai spesifikasi (milik) antara klaster 1 dan klaster 2.

- **Uji Rata – Rata Dua Populasi Independen untuk X_6**

Pengujian hipotesis dilakukan untuk mengetahui ada atau tidaknya perbedaan yang signifikan pada nilai rata – rata untuk rata – rata *bandwidth* antara klaster 1 dan 2 sebagai berikut:

i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak terdapat perbedaan yang signifikan pada rata – rata nilai rata – rata *bandwidth* antara klaster 1 dan klaster 2)

$H_1 : \mu_1 \neq \mu_2$ (Terdapat perbedaan yang signifikan pada rata – rata nilai rata – rata *bandwidth* antara klaster 1 dan klaster 2)

ii. Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii. Daerah Kritis

Tolak H_0 jika $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

iv. Statistik Uji

Diperoleh nilai $Z_{hitung} = 3.1875$

v. Keputusan

Tolak H_0 karena nilai $Z_{hitung} > Z_{\frac{\alpha}{2}}$ atau $3.1875 > 1.96$

vi. Kesimpulan

Dengan menggunakan tingkat kepercayaan sebesar 95% maka data yang ada mendukung untuk tolak H_0 yang artinya terdapat perbedaan yang signifikan pada rata – rata nilai rata – rata *bandwidth* antara klaster 1 dan klaster 2.

• **Uji Rata – Rata Dua Populasi Independen untuk X_7**

Pengujian hipotesis dilakukan untuk mengetahui ada atau tidaknya perbedaan yang signifikan nilai rata – rata jumlah *switch hub* antara klaster 1 dan 2 sebagai berikut:

i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak terdapat perbedaan yang signifikan pada rata – rata jumlah *switch hub* antara klaster 1 dan klaster 2)

$H_1 : \mu_1 \neq \mu_2$ (Terdapat perbedaan yang signifikan pada rata – rata jumlah *switch hub* antara klaster 1 dan klaster 2)

ii. Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii. Daerah Kritis

Tolak H_0 jika $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

iv. Statistik Uji

Diperoleh nilai $Z_{hitung} = 2.5919$

v. Keputusan

Tolak H_0 karena nilai $Z_{hitung} > Z_{\frac{\alpha}{2}}$ atau $2.5919 > 1.96$

vi. Kesimpulan

Menggunakan tingkat kepercayaan sebesar 95% maka data yang ada mendukung untuk tolak H_0 artinya terdapat perbedaan yang signifikan pada rata – rata jumlah *switch hub* antara klaster 1 dan klaster 2.

• **Uji Rata – Rata Dua Populasi Independen untuk X_8**

Pengujian hipotesis dilakukan untuk mengetahui ada atau tidak adanya perbedaan yang signifikan pada nilai rata – rata jumlah WIFI antara klaster 1 dan 2 sebagai berikut:

i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak terdapat perbedaan yang signifikan pada rata – rata jumlah WIFI antara klaster 1 dan klaster 2)

$H_1 : \mu_1 \neq \mu_2$ (Terdapat perbedaan yang signifikan pada rata – rata jumlah WIFI antara klaster 1 dan klaster 2)

ii. Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii. Daerah Kritis

Tolak H_0 jika $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

iv. Statistik Uji

Diperoleh nilai $Z_{hitung} = 2.9504$

v. Keputusan

Tolak H_0 karena $Z_{hitung} > Z_{\frac{\alpha}{2}}$ atau $2.9504 > 1.96$

vi. Kesimpulan

Menggunakan tingkat kepercayaan sebesar 95% maka data yang ada mendukung untuk tolak H_0 artinya terdapat perbedaan secara signifikan pada rata – rata jumlah WIFI antara klaster 1 dan klaster 2.

• **Uji Rata – Rata Dua Populasi Independen untuk Daya Listrik**

Pengujian hipotesis dilakukan untuk mengetahui ada atau tidak adanya perbedaan yang signifikan pada nilai rata – rata daya listrik antara klaster 1 dan 2 sebagai berikut:

i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak terdapat perbedaan yang signifikan pada rata – rata daya listrik antara klaster 1 dan klaster 2)

$H_1 : \mu_1 \neq \mu_2$ (Terdapat perbedaan yang signifikan pada rata – rata daya listrik antara klaster 1 dan klaster 2)

ii. Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii. Daerah Kritis

Tolak H_0 jika $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

iv. Statistik Uji

Diperoleh nilai $Z_{hitung} = 2.2328$

v. Keputusan

Tolak H_0 karena $Z_{hitung} > Z_{\frac{\alpha}{2}}$ atau $2.2328 > 1.96$

vi. Kesimpulan

Menggunakan tingkat kepercayaan sebesar 95% maka data yang ada mendukung untuk tolak H_0 artinya terdapat perbedaan secara signifikan pada rata – rata daya listrik antara klaster 1 dan klaster 2.

• **Uji Rata – Rata Dua Populasi Independen untuk *Bandwidth Upload***

Pengujian hipotesis dilakukan untuk mengetahui ada atau tidak adanya perbedaan yang signifikan pada nilai rata – rata *bandwidth upload* antara klaster 1 dan 2 sebagai berikut:

i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak terdapat perbedaan yang signifikan pada rata – rata *bandwidth upload* antara klaster 1 dan klaster 2)

$H_1 : \mu_1 \neq \mu_2$ (Terdapat perbedaan yang signifikan pada rata – rata *bandwidth upload* antara klaster 1 dan klaster 2)

ii. Tingkat Signifikansi

$\alpha = 5\% = 0.05$

iii. Daerah Kritis

Tolak H_0 jika $Z_{hitung} < -Z_{\frac{\alpha}{2}}$ atau $Z_{hitung} > Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

iv. Statistik Uji

Diperoleh nilai $Z_{hitung} = 2.9302$

v. Keputusan

Tolak H_0 karena $Z_{hitung} > Z_{\frac{\alpha}{2}}$ atau $2.9302 > 1.96$

vi. Kesimpulan

Menggunakan tingkat kepercayaan sebesar 95% maka data yang ada mendukung untuk tolak H_0 artinya terdapat perbedaan secara signifikan pada rata – rata *bandwidth upload* antara klaster 1 dan klaster 2.

- **Uji Rata – Rata Dua Populasi Independen untuk *Bandwidth Download***

Pengujian hipotesis dilakukan untuk mengetahui ada atau tidak adanya perbedaan yang signifikan pada nilai rata – rata *bandwidth download* antara klaster 1 dan 2 sebagai berikut:

i. Hipotesis

$H_0 : \mu_1 = \mu_2$ (Tidak terdapat perbedaan yang signifikan pada rata – rata *bandwidth download* antara klaster 1 dan klaster 2)

$H_1 : \mu_1 \neq \mu_2$ (Terdapat perbedaan yang signifikan pada rata – rata *bandwidth download* antara klaster 1 dan klaster 2)

ii. Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

iii. Daerah Kritis

$$\text{Tolak } H_0 \text{ jika } Z_{hitung} < -Z_{\frac{\alpha}{2}} \text{ atau } Z_{hitung} > Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$$

iv. Statistik Uji

$$\text{Diperoleh nilai } Z_{hitung} = 3.3635$$

v. Keputusan

$$\text{Tolak } H_0 \text{ karena } Z_{hitung} > Z_{\frac{\alpha}{2}} \text{ atau } 3.3635 > 1.96$$

vi. Kesimpulan

Menggunakan tingkat kepercayaan sebesar 95% maka data yang ada mendukung untuk tolak H_0 artinya terdapat perbedaan secara signifikan pada rata – rata *bandwidth download* antara klaster 1 dan klaster 2.

Dipunyai suatu data yang disajikan dalam **Tabel 1**. Data simulasi ini merupakan modifikasi data yang dilakukan oleh peneliti yang bersumber dari salah satu website perhitungan jarak data campuran *Gower* yaitu:

<https://jamesmccaffrey.wordpress.com/2020/04/21/example-of-calculating-the-gower-distance/>. Terdapat 8 objek dengan 6 variabel yang akan digunakan yaitu variabel numerik adalah *Age* dan *Income*, variabel biner adalah *IsMale* (*TRUE* or *FALSE*), serta variabel kategorik adalah *Race* (1, 2, 3), *Height* (1, 2, 3) dan *Politic* (*Moderate*, *Liberal* dan *Conservative*).

Tabel 1 Data Contoh Studi Kasus

Objek ke- <i>i</i>	<i>Age</i>	<i>Income</i>	<i>IsMale</i>	<i>Race</i>	<i>Height</i>	<i>Politic</i>
1	22	0.39	<i>TRUE</i>	1	3	<i>Moderate</i>
2	33	0.34	<i>TRUE</i>	3	1	<i>Liberal</i>
3	52	0.51	<i>FALSE</i>	1	2	<i>Moderate</i>
4	46	0.63	<i>TRUE</i>	2	3	<i>Conservative</i>
5	35	0.44	<i>FALSE</i>	2	3	<i>Liberal</i>
6	29	0.76	<i>TRUE</i>	3	1	<i>Liberal</i>
7	24	0.48	<i>FALSE</i>	1	2	<i>Moderate</i>
8	38	0.53	<i>TRUE</i>	3	1	<i>Liberal</i>

Sebelum menghitung jarak antar objek maka akan dilakukan perhitungan nilai kedekatan antar objek. Kedekatan antara objek ke-*i* dan objek ke-*j* sesuai dengan jenis data yaitu numerik, biner dan kategorik dengan perhitungan manual sebagai berikut.

- **Kedekatan Antar Objek Data Numerik**

Perhitungan kedekatan antar objek numerik menggunakan rumus pada **Persamaan 3.1**.

$$\begin{aligned}
 s_{num(1,1)} &= \left(1 - \frac{|22 - 22|}{(52 - 22)}\right) + \left(1 - \frac{|0.39 - 0.39|}{(0.76 - 0.34)}\right) = 2 \\
 s_{num(1,2)} &= \left(1 - \frac{|22 - 33|}{(52 - 22)}\right) + \left(1 - \frac{|0.39 - 0.34|}{(0.76 - 0.34)}\right) = 1.51429 \\
 s_{num(1,3)} &= \left(1 - \frac{|22 - 52|}{(52 - 24)}\right) + \left(1 - \frac{|0.39 - 0.51|}{(0.76 - 0.34)}\right) = 0.71429
 \end{aligned}$$

$$S_{num(7,7)} = \left(1 - \frac{|24 - 24|}{(52 - 22)}\right) + \left(1 - \frac{|0.48 - 0.48|}{(0.76 - 0.34)}\right) = 2$$

$$S_{num(7,8)} = \left(1 - \frac{|24 - 38|}{(52 - 22)}\right) + \left(1 - \frac{|0.48 - 0.53|}{(0.76 - 0.34)}\right) = 1.41429$$

$$S_{num(8,8)} = \left(1 - \frac{|38 - 38|}{(52 - 22)}\right) + \left(1 - \frac{|0.53 - 0.53|}{(0.76 - 0.34)}\right) = 2$$

Hasil perhitungan kedekatan antar objek data numerik disajikan dalam bentuk matriks pada **Tabel 2**.

Tabel 2 Contoh: Matriks Kedekatan Antar Objek Data Numerik

	1	2	3	4	5	6	7	8
1	2	1.51429	0.71429	0.62857	1.44762	0.88571	1.71905	1.13333
2	1.51429	2	0.96190	0.87619	1.69524	0.86667	1.36667	1.38095
3	0.71429	0.96190	2	1.51429	1.26667	0.63810	0.99524	1.48571
4	0.62857	0.87619	1.51429	2	1.18095	1.12381	0.90952	1.49524
5	1.44762	1.69524	1.26667	1.18095	2	1.03810	1.53810	1.68571
6	0.88571	0.86667	0.63810	1.12381	1.03810	2	1.16667	1.15238
7	1.71905	1.36667	0.99524	0.90952	1.53810	1.16667	2	1.41429
8	1.13333	1.38095	1.48571	1.49524	1.68571	1.15238	1.41429	2

- **Kedekatan Antar Objek Data Biner dan Kategorik**

Perhitungan kedekatan antar objek biner dan kategorik menggunakan rumus pada **Persamaan 3.2**.

$$S_{cat(1,1)} = 1 + 1 + 1 + 1 = 4$$

$$S_{cat(1,2)} = 1 + 0 + 0 + 0 = 1$$

$$S_{cat(1,3)} = 0 + 1 + 0 + 1 = 2$$

$$S_{cat(7,7)} = 1 + 1 + 1 + 1 = 4$$

$$S_{cat(7,8)} = 0 + 0 + 0 + 0 = 0$$

$$S_{cat(8,8)} = 1 + 1 + 1 + 1 = 4$$

Hasil perhitungan kedekatan antar objek data biner dan kategorik disajikan dalam bentuk matriks pada **Tabel 3**.

Tabel 3 Contoh: Matriks Kedekatan Antar Objek Data Biner & Kategorik

	1	2	3	4	5	6	7	8
1	4	1	2	2	1	1	2	1
2	1	4	0	1	1	4	0	4
3	2	0	4	0	1	0	4	0
4	2	1	0	4	2	1	0	1

	1	2	3	4	5	6	7	8
5	1	1	1	2	4	1	1	1
6	1	4	0	1	1	4	0	4
7	2	0	4	0	1	0	4	0
8	1	4	0	1	1	4	0	4

- **Kedekatan *Gower* Antar Objek Data Campuran**

Perhitungan kedekatan *Gower* antar objek data campuran menggunakan rumus pada **Persamaan 3.3**.

$$\begin{aligned}
 d_{gab(1,1)} &= \left(\frac{2+4}{6} \right) &= 1 \\
 d_{gab(1,2)} &= \left(\frac{1.51429+1}{6} \right) &= 0.41905 \\
 d_{gab(1,3)} &= \left(\frac{0.71429+2}{6} \right) &= 0.45238 \\
 d_{gab(7,7)} &= \left(\frac{2+4}{6} \right) &= 0 \\
 d_{gab(7,8)} &= \left(\frac{1.41429+0}{6} \right) &= 0.23571 \\
 d_{gab(8,8)} &= \left(\frac{2+4}{6} \right) &= 0
 \end{aligned}$$

Hasil perhitungan kedekatan antar objek data campuran disajikan dalam bentuk matriks pada **Tabel 4**.

Tabel 4 Contoh: Kedekatan Antar Objek Data Campuran

	1	2	3	4	5	6	7	8
1	1.00000	0.41905	0.45238	0.43810	0.40794	0.31429	0.61984	0.35556
2	0.41905	1.00000	0.16032	0.31270	0.44921	0.81111	0.22778	0.89683
3	0.45238	0.16032	1.00000	0.25238	0.37778	0.10635	0.83254	0.24762
4	0.43810	0.31270	0.25238	1.00000	0.53016	0.35397	0.15159	0.41587
5	0.40794	0.44921	0.37778	0.53016	1.00000	0.33968	0.42302	0.44762
6	0.31429	0.81111	0.10635	0.35397	0.33968	1.00000	0.19444	0.85873
7	0.61984	0.22778	0.83254	0.15159	0.42302	0.19444	1.00000	0.23571
8	0.35556	0.89683	0.24762	0.41587	0.44762	0.85873	0.23571	1.00000

- **Jarak *Gower* Antar Objek Data Campuran**

Perhitungan jarak *Gower* antar objek data campuran menggunakan rumus pada **Persamaan 3.4**.

$$\begin{aligned}
 d_{gab(1,1)} &= 1 - 1 &= 0 \\
 d_{gab(1,2)} &= 1 - 0.41905 &= 0.58095
 \end{aligned}$$

$$\begin{aligned}
d_{gab(1,3)} &= 1 - 0.45238 &= 0.54762 \\
d_{gab(7,7)} &= 1 - 1 &= 0 \\
d_{gab(7,8)} &= 1 - 0.23571 &= 0.76429 \\
d_{gab(8,8)} &= 1 - 1 &= 0
\end{aligned}$$

Berdasarkan perhitungan kedekatan campuran di atas, maka akan didapatkan matriks jarak antar objek yang disajikan pada **Tabel 5**.

Tabel 5 Contoh: Matriks Jarak Antar Objek

	1	2	3	4	5	6	7	8
1	0	0.58095	0.54762	0.56190	0.59206	0.68571	0.38016	0.64444
2	0.58095	0	0.83968	0.68730	0.55079	0.18889	0.77222	0.10317
3	0.54762	0.83968	0	0.74762	0.62222	0.89365	0.16746	0.75238
4	0.56190	0.68730	0.74762	0	0.46984	0.64603	0.84841	0.58413
5	0.59206	0.55079	0.62222	0.46984	0	0.66032	0.57698	0.55238
6	0.68571	0.18889	0.89365	0.64603	0.66032	0	0.80556	0.14127
7	0.38016	0.77222	0.16746	0.84841	0.57698	0.80556	0	0.76429
8	0.64444	0.10317	0.75238	0.58413	0.55238	0.14127	0.76429	0

Hasil perhitungan jarak **Tabel 5** merupakan jarak *Gower* yang digunakan pada pengelompokan menggunakan algoritma *Simple K-Medoids Partitioning*.

Pengelompokan Menggunakan Algoritma *Simple K-Medoids Partitioning*

Pengelompokan menggunakan algoritma SKM untuk data dalam **Tabel 1** dilakukan kedalam dua klaster. Dari matriks jarak pada **Tabel 5** diperoleh nilai a_i menggunakan **Persamaan 3.10** Contoh perhitungan manual nilai a_i :

$$\begin{aligned}
a_1 &= 0 + 0.58095 + 0.54762 + 0.56190 + 0.59206 + 0.68571 + 0.38016 \\
&\quad + 0.64444 = 3.99286
\end{aligned}$$

$$\begin{aligned}
a_2 &= 0.58095 + 0 + 0.83968 + 0.68730 + 0.55079 + 0.18889 + 0.77222 \\
&\quad + 0.75238 = 3.72302
\end{aligned}$$

$$\begin{aligned}
a_8 &= 0.64444 + 0.10317 + 0.75238 + 0.58413 + 0.55238 + 0.14127 + \\
&\quad 0.76429 + 0 = 3.54206
\end{aligned}$$

Setelah diperoleh nilai a_i seluruh objek, lalu diurutkan untuk memperoleh 2 nilai a_i terkecil sebagai medoid awal. Diperoleh nilai a_i yang disajikan pada **Tabel 6**.

Tabel 6 Contoh: Nilai a_i Matriks Jarak

Objek ke- i	a_i
8	3.54206
2	3.72302
1	3.99286

Objek ke- i	a_i
6	4.02143
5	4.02460
7	4.31508
4	4.54524
3	4.57064

Berdasarkan **Tabel 6** didapatkan nilai a_i terkecil yang merupakan nilai a_i dari objek ke-8 yang akan menjadi medoid awal kluster 1. Karena pengelompokan dilakukan dalam 2 kluster maka untuk medoid kluster 2 dipilih secara random. Randomisasi dapat dilakukan dengan =RANDBETWEEN () dalam *Microsoft Excel*. *Output* yang dihasilkan dalam *R Studio*, diperoleh objek ke-2 yang menjadi medoid awal kluster 2. Maka medoid awal yang didapatkan disajikan pada **Tabel 7**.

Tabel 7 Contoh: Medoid Awal

Objek ke- i	Keterangan	Age	Income	IsMale	Race	Height	Politic
8	Data ke-8 sebagai medoid awal kluster 1	38	0.53	0	3	1	1
2	Data ke-2 sebagai medoid awal kluster 2	33	0.34	0	3	1	1

Jarak objek menuju medoid dapat dilihat pada matriks jarak yang disajikan pada **Tabel 5**. Pilih jarak pada objek ke-8 dan ke-2 yang akan menjadi jarak antara objek ke masing – masing medoid awalnya. Perhitungan ini merupakan iterasi 1 dalam proses pengelompokan. Nilai tersebut disajikan pada **Tabel 8**.

Tabel 8 Contoh: Iterasi 1

Objek ke- i	Jarak ke Medoids		Jarak Terdekat	Kluster
	$M1 = 8$	$M2 = 2$		
1	0.64444	0.58905	0.58905	2
2	0.10317	0	0	2
3	0.75238	0.83968	0.75238	1
4	0.58413	0.68730	0.58413	1
5	0.55238	0.55079	0.55079	2
6	0.14127	0.18889	0.14127	1
7	0.76429	0.77222	0.76429	1
8	0	0.10317	0	1
Total Jarak (E)			3.37381	

Penentuan jarak terdekat dengan medoid menggunakan rumus : $MIN d(x_n, m_p)$. Pada **Tabel 8** diperoleh nilai total jarak (E) sebesar 3.37381, total

jarak (E) diperoleh dengan menjumlahkan kedua jumlah M1 dan M2. Adapun hasil penentuan jarak terdekat menuju medoid telah ditabulasikan pada **Tabel 8**. Jika dilihat pada **Tabel 8**, kluster 1 terdiri dari objek ke-3, 4, 6, 7 dan 8 sedangkan kluster 2 terdiri dari objek ke-1, 2 dan 5. Maka akan dibuat matriks jarak dari setiap kluster untuk mendapat nilai a_i minimumnya. Nilai jarak antar objek pada kluster 1 dibuat dalam bentuk matriks jarak disajikan pada **Tabel 9**.

Tabel 9 Contoh: Matriks Jarak *Reduce* Kluster 1 Iterasi 1

	3	4	6	7	8	a_i
3	0	0.74762	0.89365	0.16746	0.75238	2.56111
4	0.74762	0	0.64603	0.84841	0.58413	2.82619
6	0.89365	0.64603	0	0.80556	0.14127	2.48651
7	0.16746	0.84841	0.80556	0	0.76429	2.58571
8	0.75238	0.58413	0.14127	0.76429	0	2.24206

Nilai a_i didapatkan dari penjumlahan objek yang berada dalam satu baris. Pada **Tabel 9** diperoleh nilai a_i terkecil kluster 1 merupakan nilai a_i dari objek ke-8 dengan nilai a_i sebesar 2.24206. Jadi medoid baru kluster 1 yaitu objek ke-8. Dengan langkah yang sama dilakukan perhitungan untuk medoid baru pada kluster 2. Matriks jarak objek yang masuk dalam kluster 2 disajikan pada **Tabel 10**.

Tabel 10 Contoh: Matriks Jarak *Reduce* Kluster 2 Iterasi 1

	1	2	5	a_i
1	0	0.58095	0.59206	1.17302
2	0.58095	0	0.55079	1.13175
5	0.59206	0.55079	0	1.14286

Pada **Tabel 10** diperoleh nilai a_i terkecil dari objek dalam kluster 2 sebesar 1.13175 yang merupakan nilai a_i dari objek ke-2. Setelah diperoleh set medoid baru pada setiap kluster, lalu dilakukan perhitungan iterasi kedua menggunakan set medoid baru. Medoid baru iterasi kedua disajikan pada **Tabel 11**.

Tabel 11 Contoh: Medoid Baru Iterasi 2

Objek ke- i	Keterangan	Age	Income	IsMale	Race	Height	Politic
8	Data ke-8 sebagai medoid baru kluster 1	38	0.53	0	3	1	1
2	Data ke-2 sebagai medoid baru kluster 2	33	0.34	0	3	1	1

Kemudian akan dilakukan perhitungan jarak objek menuju medoid baru. Pilih jarak pada objek ke-8 dan ke-2 yang menjadi jarak antara objek ke masing-masing medoid barunya. Hasil perhitungan iterasi kedua menggunakan medoid yang baru tersebut disajikan pada **Tabel 12**.

Tabel 12 Contoh: Iterasi 2

Objek ke- <i>i</i>	Jarak ke <i>Medoids</i>		Jarak Terdekat	Klaster
	<i>M1</i> = 8	<i>M2</i> = 2		
1	0.64444	0.58905	0.58905	2
2	0.10317	0	0	2
3	0.75238	0.83968	0.75238	1
4	0.58413	0.68730	0.58413	1
5	0.55238	0.55079	0.55079	2
6	0.14127	0.18889	0.14127	1
7	0.76429	0.77222	0.76429	1
8	0	0.10317	0	1
Total Jarak (E)			3.37381	

Pada iterasi kedua didapatkan nilai *E* sebesar 3.37381. Apabila nilai *E* bernilai sama dengan nilai *E* sebelumnya atau himpunan medoid tidak berubah maka proses iterasi dihentikan. Karena nilai *E* pada *seeding* 1 untuk iterasi 1 dan iterasi 2 bernilai sama serta himpunan medoid tetap maka proses iterasi dihentikan. Didapatkan hasil pengelompokan menggunakan SKM disajikan pada **Tabel 13**.

Tabel 13 Contoh: Hasil Pengelompokan *Seeding* 1

Klaster	Anggota Klaster (Objek ke- <i>i</i>)	Jumlah Anggota Klaster
1	3, 4, 6, 7, 8	5
2	1, 2, 5	3

Selanjutnya akan dilakukan perhitungan kembali dengan langkah yang sama untuk 20 kali *seeding* di mana pada setiap kali *seeding* dibatasi dengan maksimum *iterate* adalah 10 kali. Proses *seeding* dilakukan dengan tahapan yang sama dengan yang telah dijelaskan sebelumnya. Pada setiap kali *seeding* akan menggunakan *most centrally located object* atau medoid yang terletak paling pusat dengan meminimumkan jarak. Selain *most centrally located object* maka medoid lainnya akan terpilih secara random. Penentuan medoid akhir dilihat pada nilai total jarak (*E*) yang minimum. Diperoleh ringkasan hasil *seeding* pada **Tabel 14**.

Tabel 14 Contoh: Hasil Perhitungan dengan *Seeding* = 20

<i>Seeding</i>	Iterasi	Medoid		Nilai <i>E</i>	<i>Seeding</i>	Iterasi	Medoid		Nilai <i>E</i>
		M1	M2				M1	M2	
1	1	8	2	3.37381	13	1	8	7	1.92857
	2	8	2	3.37381		2	8	7	1.92857
2	1	8	1	2.28651	14	1	8	6	3.40079
	2	8	7	1.92857		2	1	6	2.4119
	3	8	7	1.92857		3	7	8	1.92857
3	1	8	6	3.40079		4	7	8	1.92857
	2	1	6	2.4119	15	1	8	4	2.7881
	3	7	8	1.92857		2	8	5	2.50556
	4	7	8	1.92857		3	8	7	1.92857
4	1	8	7	1.92857		4	8	7	1.92857
	2	8	7	1.92857	16	1	8	5	2.50556
5	1	8	3	2.09603		2	8	7	1.92857
	2	8	7	1.92857		3	8	7	1.92857
	3	8	7	1.92857	17	1	8	4	2.7881
6	1	8	1	2.28651		2	8	5	2.50556
	2	8	7	1.92857		3	8	7	1.92857
	3	8	7	1.92857		4	8	7	1.92857
7	1	8	3	2.09603	18	1	8	3	2.09603
	2	8	7	1.92857		2	8	7	1.92857
	3	8	7	1.92857	19	3	8	7	1.92857
8	1	8	3	2.09603		1	8	6	3.40079
	2	8	7	1.92857		2	1	6	2.4119
	3	8	7	1.92857		3	7	8	1.92857
9	1	8	7	1.92857		4	7	8	1.92857
	2	8	7	1.92857	20	1	8	4	2.7881
10	1	8	7	1.92857		2	8	5	2.50556
	2	8	7	1.92857		3	8	7	1.92857
11	1	8	1	2.28651		4	8	7	1.92857
	2	8	7	1.92857	12				
	3	8	7	1.92857					
12	1	8	1	2.28651					
	2	8	7	1.92857					
	3	8	7	1.92857					

Berdasarkan **Tabel 14** didapatkan hasil dari 20 kali *seeding* di mana objek yang menjadi *initial* medoid adalah objek ke-8 karena objek meminimalkan jumlah jarak antar objek yang berada dalam satu klaster. Hasil dari 20 kali *seeding* yang dilakukan diperoleh nilai *E* yang paling minimum sebesar 1.92857. Nilai *E* tersebut dihasilkan dari medoid akhir yaitu objek ke-8 dan ke-7. Sehingga hasil akhir

pengelompokan data pada **Tabel 1** disajikan pada **Tabel 13** yang mana kluster 1 terdiri dari 5 anggota dan kluster 2 terdiri dari 3 anggota.

Validasi Kluster Menggunakan *Silhouette Index* (SI) dan *Medoid Shadow Value* (MSV)

Berdasarkan hasil pengelompokan untuk data contoh pada **Tabel 1** maka akan dilakukan validasi kluster menggunakan nilai SI dan MSV. Validasi hanya dilakukan dengan menggunakan percobaan $k = 2$.

a. *Silhouette Index* (SI)

Adapun perhitungan secara manual SI memiliki tahapan sebagai berikut.

1. Langkah Awal

Menghitung jarak kedekatan antar objek pada masing-masing kluster menggunakan jarak *Gower*. Hasil perhitungan jarak untuk kluster 1 disajikan pada **Tabel 9** dan kluster 2 disajikan pada **Tabel 10**.

2. Langkah Kedua

Menghitung nilai a_i untuk kedua kluster menggunakan **Persamaan 3.13**.

Contoh menghitung nilai a_i pada kluster 1 dan 2:

$$\begin{aligned}
 a_2^1 &= \frac{1}{m_j - 1} \sum (d(x_2^1, x_2^1) + d(x_2^1, x_4^1) + d(x_2^1, x_5^1) + d(x_2^1, x_6^1) \\
 &\quad + d(x_2^1, x_8^1)) \\
 &= \frac{1}{5 - 1} \sum (0 + 0.68730 + 0.55079 + 0.18889 + 0.10317) \\
 &= 0.38254 \\
 a_1^2 &= \frac{1}{m_j - 1} \sum (d(x_1^2, x_1^2) + d(x_1^2, x_3^2) + d(x_1^2, x_7^2)) \\
 &= \frac{1}{3 - 1} \sum (0 + 0.54762 + 0.38016) \\
 &= 0.46389
 \end{aligned}$$

Adapun hasil perhitungan nilai a_i untuk masing – masing objek pada setiap kluster nya disajikan pada **Tabel 15**.

Tabel 15 Perhitungan SI : Nilai a_i Klaster 1 dan Klaster 2

Klaster 1 (C1)		Klaster 2 (C2)	
Objek	Nilai a_i	Objek	Nilai a_i
2	0.38254	1	0.46389
4	0.59683	3	0.35754
5	0.55833	7	0.27381
6	0.40913		

Klaster 1 (C1)		Klaster 2 (C2)	
Objek	Nilai a_i	Objek	Nilai a_i
8	0.34524		

3. Langkah Ketiga

Menghitung jarak pada klaster yang berbeda menggunakan *Gower* serta menghitung rata – rata nilai jarak. Hasil perhitungan jarak antara objek pada klaster 1 dengan objek pada klaster 2 disajikan pada **Tabel 16**.

Tabel 16 Perhitungan SI : Hasil Jarak Objek Klaster 1 dengan Klaster 2

	2	4	5	6	8
1	0.58095	0.56190	0.59206	0.68571	0.64444
3	0.83968	0.74762	0.62222	0.89365	0.75238
7	0.77222	0.84841	0.57698	0.80556	0.76429
mean	0.73095	0.71931	0.59709	0.79497	0.72037

Sedangkan hasil perhitungan jarak antara objek pada klaster 2 dengan objek pada klaster 1 disajikan pada **Tabel 17**.

Tabel 17 Perhitungan SI : Hasil Jarak Objek Klaster 2 dengan Klaster 1

	1	3	7
2	0.58095	0.83968	0.77222
4	0.56190	0.74762	0.84841
5	0.59206	0.62222	0.57698
6	0.68571	0.89365	0.80556
8	0.64444	0.75238	0.76429
mean	0.61302	0.77111	0.75349

4. Langkah Keempat

Menghitung nilai b_i untuk kedua klaster menggunakan **Persamaan 3.14**.

Hasil nilai b_i pada setiap klaster yang diperoleh disajikan pada **Tabel 18**.

Tabel 18 Perhitungan SI : Nilai b_i Klaster 1 dan Klaster 2

Klaster 1 (C1)		Klaster 2 (C2)	
Objek	Nilai b_i	Objek	Nilai b_i
2	0.73095	1	0.61302
4	0.71931	3	0.77111
5	0.59709	7	0.75349
6	0.79497		
8	0.72037		

5. Langkah Kelima

Selanjutnya adalah menghitung nilai *Silhouette Index* objek ke-*i* menggunakan **Persamaan 3.15**. Contoh perhitungan nilai SI untuk klaster 1:

$$\begin{aligned}
 SI_2^1 &= \frac{b_2^1 - a_2^1}{\max(a_2^1; b_2^1)} = \frac{0.73095 - 0.38254}{0.73095} = 0.47665 \\
 SI_4^1 &= \frac{b_4^1 - a_4^1}{\max(a_4^1; b_4^1)} = \frac{0.71931 - 0.59683}{0.71931} = 0.17028 \\
 SI_3^2 &= \frac{b_3^2 - a_3^2}{\max(a_3^2; b_3^2)} = \frac{0.77111 - 0.35754}{0.77111} = 0.53633 \\
 SI_7^2 &= \frac{b_7^2 - a_7^2}{\max(a_7^2; b_7^2)} = \frac{0.75349 - 0.27381}{0.75349} = 0.63661
 \end{aligned}$$

Hasil perhitungan nilai SI masing-masing objek disajikan pada **Tabel 19**.

Tabel 19 Perhitungan SI : Nilai *Silhouette Index* Setiap Objek

Klaster 1 (C1)		Klaster 2 (C2)	
Objek	Nilai SI	Objek	Nilai SI
2	0.47665	1	0.24326
4	0.17028	3	0.53633
5	0.06490	7	0.63661
6	0.48536		
8	0.52075		

6. Langkah Keenam

Menghitung nilai *Silhouette Index* pada setiap klaster menggunakan **Persamaan 3.16**.

$$\begin{aligned}
 SI_1 &= \frac{1}{m_1} (SI_2^1 + SI_4^1 + SI_5^1 + SI_6^1 + SI_8^1) \\
 &= \frac{1}{5} (0.47665 + 0.17028 + 0.06490 + 0.48536 + 0.52075) \\
 &= 0.34359 \\
 SI_2 &= \frac{1}{m_2} (SI_3^2 + SI_7^2) \\
 &= \frac{1}{3} (0.24326 + 0.53633 + 0.63661) \\
 &= 0.47207
 \end{aligned}$$

Sehingga diperoleh nilai *Silhouette Index* dari setiap klaster yang disajikan pada **Tabel 20**.

Tabel 20 Perhitungan SI: Nilai *Silhouette Index* Setiap Klaster

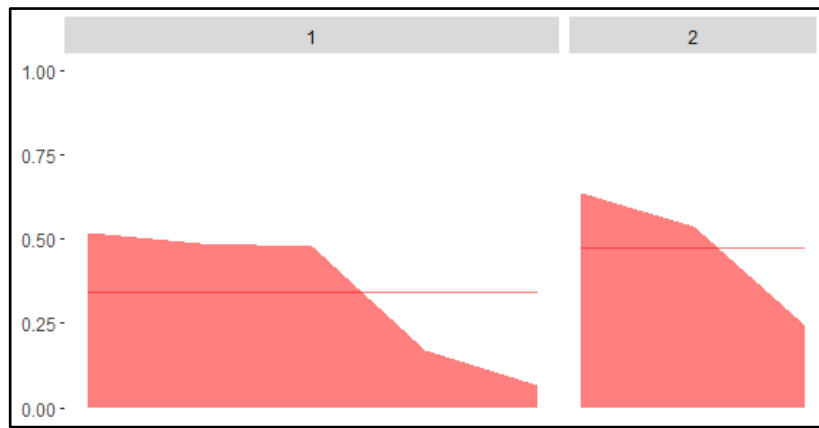
Klaster (<i>j</i>)	Nilai SI_j
1	0.34359
2	0.47207

7. Langkah Ketujuh

Pada langkah terakhir, untuk mendapatkan nilai SI global dihitung dengan mencari nilai rata – rata SI dari seluruh klaster. Berdasarkan pada **Persamaan 3.17** maka perhitungan nilai SI global sebagai berikut:

$$\begin{aligned} SI &= \frac{(0.24326 + 0.47665 + \dots + 0.63661 + 0.52075)}{8} \\ &= \frac{(3.13415)}{8} \\ &= 0.39177 \end{aligned}$$

Berdasarkan hasil perhitungan didapatkan nilai SI global sebesar 0.39177, maka dapat dikatakan bahwa struktur hasil pengelompokan data pada **Tabel 1** memiliki struktur yang lemah. Adapun *plot* SI disajikan pada **Gambar 1**.



Gambar 1 Grafik *Silhouette Index*

b. Medoid Shadow Value (MSV)

Selanjutnya, akan dilakukan validasi menggunakan *Medoid Shadow Value* (MSV). Adapun perhitungan manual dari MSV adalah sebagai berikut:

$$msv_1 = \frac{(0.64444 - 0.38016)}{0.64444} = 0.41010$$

$$msv_2 = \frac{(0.77222 - 0.10317)}{0.77222} = 0.86639$$

$$msv_3 = \frac{(0.75238 - 0.16746)}{0.75238} = 0.77743$$

$$msv_4 = \frac{(0.84841 - 0.58413)}{0.84841} = 0.31151$$

$$msv_5 = \frac{(0.57698 - 0.55238)}{0.57698} = 0.04264$$

$$msv_6 = \frac{(0.80556 - 0.14127)}{0.80556} = 0.82463$$

$$msv_7 = \frac{(0.76429 - 0)}{0.76429} = 1$$

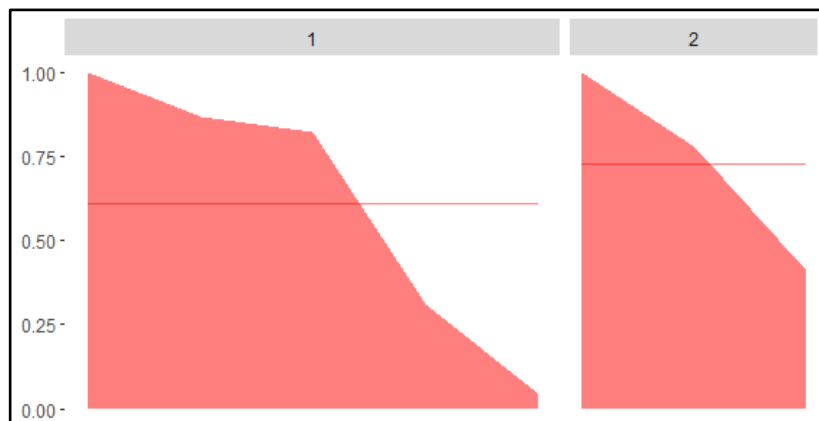
$$msv_8 = \frac{(0.76429 - 0)}{0.76429} = 1$$

Berdasarkan hasil perhitungan tersebut, maka diperoleh nilai MSV masing – masing objek yang disajikan pada **Tabel 21**.

Tabel 21 Nilai *Medoid Shadow Value* (MSV)

Objek	MSV	Klaster
1	0.41010	2
2	0.86639	1
3	0.77743	2
4	0.31151	1
5	0.04264	1
6	0.82463	1
7	1	2
8	1	1

Hasil nilai MSV masing-masing objek yang diperoleh, dapat digunakan untuk mencari nilai MSV global untuk pengelompokan data pada **Tabel 1** yaitu sebesar 0.66910. Nilai MSV tersebut dapat dikatakan dalam kategori hasil pengelompokan yang baik karena angka tersebut lebih dekat menuju 1. Adapun grafik nilai MSV disajikan pada **Gambar 2**.



Gambar 2 Grafik *Medoid Shadow Value*