

**VISUALISASI KUALITAS SISWA SEKOLAH
MENENGAH ATAS BERDASARKAN PEMAHAMAN
DASAR PADA KELURAHAN CILINCING
MENGUNAKAN METODE K - MEANS**

SKRIPSI SARJANA SISTEM INFORMASI

Oleh

Dimas Ghafi Irwansyah

227006516104



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA**

UNIVERSITAS NASIONAL

2026

**VISUALISASI KUALITAS SISWA SEKOLAH
MENENGAH ATAS BERDASARKAN PEMAHAMAN
DASAR PADA KELURAHAN CILINCING
MENGUNAKAN METODE K - MEANS
SKRIPSI SARJANA**

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar
Sarjana Sistem Informasi dari Fakultas Teknologi Komunikasi dan
Informatika

Oleh

Dimas Ghafi Irwansyah

227006516104



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
UNIVERSITAS NASIONAL**

2026

HALAMAN PENGESAHAN

TUGAS AKHIR

**VISUALISASI KUALITAS SISWA SEKOLAH MENENGAH ATAS
BERDASARKAN PEMAHAMAN DASAR PADA KELURAHAN
CILINCING MENGGUNAKAN METODE K - MEANS**



PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

**VISUALISASI KUALITAS SISWA SEKOLAH MENENGAH ATAS BERDASARKAN
PEMAHAMAN DASAR PADA KELURAHAN CILINCING MENGGUNAKAN
METODE K - MEANS**

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 3 Maret 2026


Diketahui dan disetujui
NPM 227006516104

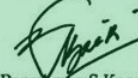
LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

VISUALISASI KUALITAS SISWA SEKOLAH MENENGAH ATAS BERDASARKAN PEMAHAMAN DASAR PADA KELURAHAN CILINCING MENGGUNAKAN METODE K – MEANS

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Review Akhir Semester Ganjil 2021-2022 pada tanggal 3 Maret Tahun 2026

Dosen Pembimbing 1



Dr. Benrahman, S.Kom., MMSI

NIDN: 0318096504

Ketua Program Studi



Dr. Ir. Andrianesih, S.Kom., MMSI

NIDN: 0303097902

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Dimas Ghafi Irwnasyah

NPM : 227006516104

Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika

Program Studi : Sistem Informasi

Tanggal Sidang : 24 Februari 2026

JUDUL DALAM BAHASA INDONESIA :

VISUALISASI KUALITAS SISWA SEKOLAH MENENGAH ATAS BERDASARKAN PEMAHAMAN DASAR PADA KELURAHAN CILINCING MENGUNAKAN METODE K - MEANS

JUDUL DALAM BAHASA INGGRIS :

VISUALIZATION OF HIGH SCHOOL STUDENT QUALITY BASED ON BASIC UNDERSTANDING IN CILINCING VILLAGE USING THE K-MEANS METHOD

TANDA TANGAN DAN TANGGAL		
Pembimbing 1	Ka-Prodi	Mahasiswa
TGL : 3 Maret 2026	TGL : 3 Maret 2026	TGL : 3 Maret 2026
		

ABSTRAK

Salah satu cara terbaik untuk mengetahui seberapa baik proses pembelajaran berlangsung adalah dengan melihat seberapa baik siswa SMA memahami dasar-dasarnya. Namun, fakta bahwa siswa di Kelurahan Cilincing memiliki tingkat kemampuan akademik yang berbeda menunjukkan bahwa mereka memahami mata pelajaran inti seperti Matematika, Sains, dan Bahasa Indonesia pada tingkat yang berbeda. Tujuan penelitian ini adalah untuk mengkategorikan kualitas siswa berdasarkan pemahaman dasar mereka dengan menggunakan metode K-Means Clustering dan untuk menyajikan hasil kategorisasi ini sebagai dasar pengambilan keputusan pendidikan. Penelitian ini menggunakan metodologi kuantitatif deskriptif yang memanfaatkan teknik data mining berbasis pembelajaran tak terawasi. Dataset terdiri dari 205 catatan siswa dari rapor semester 2 tahun ajaran 2022/2023, beserta data survei mengenai kelompok belajar. Proses penelitian meliputi pembersihan data, penggabungan, perubahan, normalisasi dengan StandardScaler, penggunaan metode Elbow dan Silhouette Score untuk menemukan jumlah cluster terbaik, dan penggunaan algoritma K-Means. Kualitas kluster menunjukkan bahwa siswa melakukan pekerjaan rata-rata di sekolah, tetapi terdapat banyak perbedaan. Kluster berkualitas tinggi, di sisi lain, selalu memiliki nilai akademik tinggi dan sesi bimbingan belajar yang lebih teratur. Plot sebaran, plot pasangan, dan dasbor interaktif menunjukkan pemisahan kluster yang jelas dan korelasi positif antar mata pelajaran. Studi ini menunjukkan bahwa metode K-Means Clustering bekerja dengan baik untuk memetakan kualitas siswa secara objektif dan sistematis. Anda dapat menggunakan hasil pengelompokan untuk membuat strategi pembelajaran yang lebih fleksibel, seperti program remedial untuk siswa yang membutuhkan bantuan tambahan dan program pengayaan untuk siswa yang berprestasi. Hal ini akan membantu meningkatkan kualitas sekolah di wilayah Kelurahan Cilincing.

Kata kunci: Data Mining, K-Means Clustering, Kualitas Siswa, Visualisasi Data, Silhouette Score, Pendidikan.

ABSTRACT

One of the best ways to assess how well a learning process is progressing is to assess how well high school students understand the fundamentals. However, the fact that students in Cilincing Village have varying levels of academic ability suggests that they understand core subjects such as Mathematics, Science, and Indonesian at varying levels. The purpose of this study is to categorize student quality based on their fundamental understanding using the K-Means Clustering method and to present the results of this categorization as a basis for educational decision-making. This study uses a descriptive quantitative methodology utilizing unsupervised learning-based data mining techniques. The dataset consists of 205 student records from semester 2 report cards of the 2022/2023 academic year, along with survey data regarding study groups. The research process included data cleaning, merging, transforming, normalization with StandardScaler, using the Elbow and Silhouette Score methods to find the best number of clusters, and using the K-Means algorithm. Cluster quality indicates that students perform averagely in school, but there is considerable variability. High-quality clusters, on the other hand, consistently have high academic grades and more regular tutoring sessions. Scatterplots, pair plots, and an interactive dashboard show clear cluster separation and positive correlations between subjects. This study demonstrates that the K-Means Clustering method works well for objectively and systematically mapping student quality. The clustering results can be used to create more flexible learning strategies, such as remedial programs for students who need additional support and enrichment programs for high-achieving students. This will help improve the quality of schools in the Cilincing Subdistrict area

Keywords: Data Mining, K-Means Clustering, Student Quality, Data Visualization, Silhouette Score, Education.

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat dan karunia sehingga penulis dapat menyelesaikan skripsi dengan judul “ **VISUALISASI KUALITAS SISWA SEKOLAH MENENGAH ATAS BERDASARKAN PEMAHAMAN DASAR PADA KELURAHAN CILINCING MENGGUNAKAN METODE K – MEANS** ” sebagai salah satu syarat kelulusan Program Studi Sarjana Fakultas Teknologi Komunikasi dan Informatika.

Penelitian dan penulisan skripsi ini tidak terlepas dari bantuan berbagai pihak, oleh karena itu penulis menyampaikan banyak terima kasih terutama kepada dosen pembimbing Tugas Akhir, Bapak Benrahman, S.Kom., M.Kom. yang telah meluangkan banyak waktu, tenaga, pikiran, arahan, motivasi serta memaklumi segala kekurangan penulis selama penelitian tugas akhir dan penyusunan skripsi. Penulis juga mengucapkan banyak terima kasih kepada :

1. Orang tua penulis yang telah banyak memberi dukungan dalam segala bentuk yang tak terhitung.
2. Seluruh dosen pengajar di Program Studi Sistem Informasi FTKI maupun dosen di Program Studi lain yang memberikan banyak ilmu.
3. Teman-teman seangkatan dan sehimpuan berbagai angkatan yang telah membantu dan mendukung.
4. Kepada saudari Jeii Weilian yang selalu memberikan nasihat dan dukungan yang begitu berarti kepada penulis.
5. Sahabat dan kerabat penulis yang sudah membantu memberikan dukungan secara mental dan finansial kepada penulis.
6. Arctic monkeys dengan album musiknya yang sudah memberikan semangat agar terus tetap menjalani tugas akhir ini hingga selesai.
7. FC Barcelona yang telah memberikan hiburan setiap minggu kepada penulis dikaloh jenuh dan Lelah dalam mengerjakan tugas akhir.

Akhir kata, semoga Tuhan Yang Maha Esa membalas kebaikan dan bantuan yang telah diberikan dengan hal yang lebih baik. Penulis mengharapkan kritik dan saran yang bersifat membangun dan semoga skripsi ini dapat memberikan manfaat di bidang Teknologi Informatika.

Jakarta, bulan dan tanggal skripsi

Penulis

Dimas Ghafi Irwansyah



DAFTAR ISI

KATA PENGANTAR.....	vii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR	xv
BAB i PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA	6
2.1 Penelitian Terdahulu.....	6
2.2 Landasan Teori	12
2.3 Technology Information System	15
2.3.1 Data Mining	15
2.3.2 Klasterisasi	16
2.3.3 K-Means Clustering	17
2.3.4 Python	18
2.3.5 Google Colab	18
2.3.6 Kualitas dan Pemahaman Dasar Siswa	18
2.3.7 Sistem Informasi dalam Pendidikan.....	19

2.4	Kerangka Berpikir	20
BAB III METODOLOGI PENELITIAN.....		22
3.1	Jenis dan Pendekatan Penelitian.....	22
3.2	Alur Penelitian.....	22
3.2.1	Pengumpulan Dataset.....	23
3.2.2	Data Pre processing.....	24
3.2.3	Tools dan software	29
3.2.4	Google Colaboratory (Google Colab).....	29
3.2.5	Phyton	30
3.2.6	Pandas	31
3.2.7	NumPy.....	31
3.2.8	Scikit-Learn (sklearn).....	31
3.2.9	Matplotlib.....	32
3.2.10	Seaborn.....	32
3.2.11	Microsoft Excel dan Google Sheets.....	32
BAB IV HASIL DAN DISKUSI		33
4.1	Implentasi Hasil Penelitian.....	33
4.1.1	Data Understanding.....	34
4.1.2	Data Processing.....	37
4.1.3	Visualisasi	41
4.2	Pembahasan	44
4.3	Hubungan Antar Variabel Akademik (Scatter Plot Antar-Fitur).....	47
4.4	Dashboard.....	48
4.5	Perhitungan Manual.....	50
BAB V Kesimpulan dan saran		59

5.1	Kesimpulan.....	59
5.2	Saran.....	59
DAFTAR PUSTAKA.....		61
LAMPIRAN.....		63



DAFTAR TABEL

Tabel 2. 1 Studi Literatur	6
Tabel 3. 1 Dataset.....	23
Tabel 4. 1 Perbandingan Silhouette Score.....	40



DAFTAR GAMBAR

Gambar 2. 1 Kerangka Berpikir	20
Gambar 3. 1 Alur Penelitian.....	22
Gambar 4. 1 Missing Counts Pada Data	35
Gambar 4. 2 Elbow Method dan Silhouette Score	39
Gambar 4. 3 Scatter Plot	41
Gambar 4. 4 Pair Plot.....	43
Gambar 4. 5 Dashboard Klasterisas	48



