

Analisis dan Visualisasi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma *K-Means Clustering* dan *Silhouette Score* Berdasarkan Asal Daerah

DOI: <http://dx.doi.org/10.35889/progresif.v22i2.3663>

Creative Commons License 4.0 (CC BY –NC)



Jermias Victor Manuhutu¹, Yoakhina Nicole Makaruku^{2*}, Daniel Halauwet³

Sistem Informasi, Institut Agama Kristen Negeri Ambon, Ambon, Indonesia

*e-mail *Corresponding Author*: y.n.makaruku@gmail.com

Abstract

Academic achievement of graduates serves as a key indicator of higher education quality and graduates' readiness to enter the workforce. This study examines the Christian Religious Education Study Program at the State Christian Institute of Ambon (IAKN Ambon) during the 2010–2020 period, focusing on the influence of students' district of origin on graduation rates and Grade Point Average (GPA). Employing a descriptive quantitative approach based on secondary data, the analysis was conducted using descriptive statistics and the K-Means clustering method, with the optimal number of clusters evaluated through the silhouette score. The findings reveal a concentration of students from several major districts, including Central Maluku, Ambon City, and West Seram Regency. Four performance clusters were identified: high-performing, moderate-performing, challenging, and low-performing/no graduates. An Average Silhouette Score of 0.77 indicates good cluster quality and strong separation among groups. The study highlights disparities in academic achievement based on students' geographical origins, suggesting the need for more targeted interventions and stronger collaboration with students' home districts to improve academic success and graduation outcomes.

Keywords: *Academic Performance; District of Origin; Graduation Rate; K-Means Clustering; Silhouette Score*

Abstrak

Prestasi akademik lulusan menjadi indikator utama mutu pendidikan tinggi dan kesiapan lulusan menghadapi dunia kerja. Penelitian ini mengkaji Program Studi Pendidikan Agama Kristen (PAK) Institut Agama Kristen Negeri Ambon periode 2010–2020, dengan fokus pada pengaruh asal kabupaten terhadap tingkat kelulusan dan IPK. Menggunakan pendekatan kuantitatif deskriptif berbasis data sekunder, analisis dilakukan melalui statistik deskriptif dan metode *K-Means clustering*, dengan evaluasi jumlah kluster menggunakan *silhouette score*. Hasil menunjukkan konsentrasi mahasiswa pada beberapa kabupaten utama seperti Maluku Tengah, Kota Ambon, dan Seram Bagian Barat. Ditemukan empat kluster kinerja, yaitu tinggi, moderat, menantang, dan rendah/tanpa lulusan. Nilai *Average Silhouette Score* sebesar 0,77 menunjukkan kualitas kluster yang baik. Studi ini mengungkap adanya perbedaan capaian akademik berdasarkan asal daerah, sehingga diperlukan intervensi yang lebih terarah serta kerja sama dengan daerah asal mahasiswa untuk meningkatkan keberhasilan studi.

Kata kunci: *Asal Kabupaten; Clustering; K-Means; Silhouette Score; Tingkat Kelulusan*

1. Pendahuluan

Tingkat kelulusan dan capaian akademik mahasiswa bukan sekadar deretan angka statistik di atas kertas, melainkan parameter utama yang merefleksikan mutu riil dari keseluruhan proses edukasi di perguruan tinggi [1]. Indikator prestasi akademik ini tidak hanya menggarisbawahi sejauh mana keberhasilan institusi dalam mentransfer ilmu pengetahuan secara efektif [2], tetapi juga menjadi bukti kesiapan serta daya saing lulusan saat terjun ke

dalam dunia kerja [3]. Oleh karena itu, melakukan pemetaan tingkat kelulusan berdasarkan latar belakang wilayah geografis asal mahasiswa merupakan langkah yang sangat krusial [4]. Melalui pemetaan regional yang akurat, kami di institusi pendidikan dapat mendeteksi wilayah mana saja yang memengaruhi retensi studi dan memerlukan penguatan intervensi akademik khusus [5]. Dengan demikian, kebijakan strategis yang diambil ke depan dapat berjalan lebih berkeadilan dan tepat sasaran demi menjaga standar mutu lulusan secara berkesinambungan.

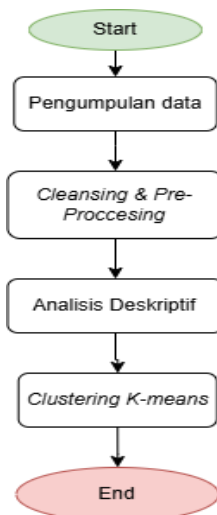
Tantangan nyata terkait pengelolaan data kelulusan ini ditemukan pada Program Studi Pendidikan Agama Kristen (PAK) Institut Agama Kristen Negeri Ambon, yang merupakan program studi dengan jumlah mahasiswa terbesar di institusi ini. Sepanjang rentang waktu tahun 2010 hingga 2020, kami mencermati adanya berbagai dinamika kurikulum, pergeseran regulasi pendidikan, hingga variasi kondisi sosio-ekonomi daerah asal yang memicu prokrastinasi maupun stres akademik pada mahasiswa rantau [6];[7]. Masalah psikologis dan adaptasi budaya dari wilayah asal yang berbeda ini secara tidak langsung memengaruhi kelancaran serta masa studi mereka selama di perkuliahan [8]. Masalah utama yang dihadapi dalam situasi ini adalah besarnya volume data historis kelulusan selama satu dekade yang masih tersimpan dalam bentuk basis data konvensional yang pasif, menumpuk, dan belum diolah secara optimal [9]. Akibatnya, kami selaku pengelola program studi mengalami kesulitan dalam mengidentifikasi pola tersembunyi (*hidden pattern*) mengenai kabupaten mana saja yang memiliki performa kelulusan tinggi, moderat, maupun rendah, sehingga keputusan intervensi akademik yang diambil selama ini cenderung bersifat umum dan kurang menyentuh akar masalah.

Dalam upaya memecahkan persoalan ini, beberapa riset terdahulu telah menerapkan metode ekstraksi basis data (*data mining*) dan visualisasi untuk mengolah data akademik [10]. Teknik pengelompokan (*clustering*) banyak digunakan untuk mengelompokkan objek berdasarkan kemiripan karakteristiknya sehingga mampu mengidentifikasi pola dan struktur pada data yang heterogen maupun multitype [11]. Pada ranah ini, pengujian kekompakan objek data telah dievaluasi menggunakan algoritma *K-Means* dan *Silhouette Score*, termasuk melalui kombinasi dengan metode *Elbow*. Selaras dengan hal tersebut, algoritma serupa juga diimplementasikan untuk mengelompokkan faktor eksternal penentu capaian studi, yang kemudian diperluas melalui pemanfaatan metrik *Within-Cluster Sum of Squares (WCSS)* demi mengidentifikasi pola hasil belajar ke dalam kelompok homogen [12]. Guna mengatasi keterbatasan visualisasi data akademik yang bersifat non-spasial, penggunaan Sistem Informasi Geografis (SIG) berbasis *Quantum GIS (QGIS)* mulai diterapkan untuk memetakan sebaran demografi mahasiswa lintas wilayah secara dinamis [13], yang kemudian dilengkapi pada level manajemen program studi menggunakan platform *Tableau* demi memantau tren performansi secara *real-time* [14]. Meskipun rangkaian riset tersebut berkontribusi besar dalam menunjang keputusan yang didasarkan pada data (*data-driven decision making*), masih terdapat celah penelitian (*gap*) yang belum terselesaikan; literatur terdahulu umumnya baru mengkaji performa akademik internal secara parsial atau visualisasi demografi makro secara terpisah, tanpa menyatukan analisis deskriptif multidimensi *K-Means* tervalidasi *Silhouette* ke dalam peta klaster geospasial asal kabupaten secara spesifik, serta belum mengeksplorasi tren historis jangka panjang satu dekade penuh demi menyajikan visualisasi performa wilayah yang interaktif bagi pimpinan institusi.

Untuk mengisi kesenjangan tersebut, konsep (*state of the art*) yang diusulkan dalam penelitian ini adalah pengembangan model visualisasi tingkat kelulusan melalui integrasi analisis deskriptif multidimensi dan algoritma *K-Means Clustering* [15] yang divalidasi dengan *Silhouette Analysis* [16], lalu dipetakan secara spasial berdasarkan asal daerah kabupaten mahasiswa selama periode sepuluh tahun (2010–2020) [17]. Rasionalisasi logis mengapa konsep ini dipandang sangat efektif karena algoritma *K-Means* memiliki efisiensi yang tinggi dalam pengelompokan data numerik kelulusan, sementara penggunaan *Silhouette Analysis* mampu menjamin objektivitas dan kekuatan pemisahan antar-klaster secara valid. Sisi kebaruan (*novelty*) yang ditawarkan dalam riset ini terletak pada transformasi data tekstual-historis satu dekade menjadi sebuah peta klaster performa akademis regional yang interaktif. Pendekatan ini tidak lagi memandang mahasiswa sebagai individu yang terisolasi, melainkan mengelompokkan karakteristik daerah asal mereka ke dalam empat klaster kinerja, sehingga menghasilkan sebuah instrumen evaluasi geospasial yang mutakhir bagi pimpinan institusi dalam menunjang keputusan yang didasarkan pada data (*data-driven decision making*) [18].

2. Metodologi

Penelitian ini dilakukan mengikuti tahapan sistematis untuk memastikan hasil analisis yang akurat dan terukur. Alur penelitian dimulai dari tahap awal hingga akhir ditunjukkan dengan diagram alir (*flowchart*) dalam gambar alur penelitian. Proses ini diawali dengan simbol "Start", dilanjutkan dengan pengumpulan data, pembersihan data, analisis, hingga pengelompokan data, dan diakhiri dengan simbol "End".



Gambar 1. Alur Penelitian

Alur Penelitian Berdasarkan Gambar 1, penelitian bermula dengan tahap Pengumpulan Data, di mana data mentah diperoleh dari sumber terkait. Selanjutnya, data masuk ke tahap *Cleansing & Pre-processing* untuk membersihkan data dari nilai yang kosong atau tidak relevan agar siap dianalisis. Setelah data bersih, dilakukan analisis deskriptif untuk mendapatkan gambaran umum mengenai statistik dasar seperti jumlah mahasiswa dan rata-rata IPK. Tahap inti dari penelitian ini adalah *Clustering K-Means*, di mana algoritma digunakan untuk mengelompokkan data daerah asal mahasiswa berdasarkan kemiripan polanya.

3. Hasil dan Pembahasan

3.1 Pengumpulan data

Tahap awal penelitian dimulai dengan proses perolehan data mentah yang menjadi dasar analisis. Penelitian ini menggunakan jenis data sekunder, yaitu data yang tidak diambil langsung dari responden melainkan dari basis data institusi. Dataset diperoleh dari UPT TIPD Institut Agama Kristen Negeri Ambon yang mencakup rekam jejak akademik mahasiswa Program Studi Pendidikan Agama Kristen (PAK) untuk periode tahun 2010 hingga 2020. Dengan mencakup beberapa variabel utama yaitu Asal Daerah (Kabupaten/Kota), Status Kelulusan, Indeks Prestasi Kumulatif (IPK) yang diperlukan dalam proses analisis.

3.2 *Cleansing & pre-processing*

Setelah data yang diperlukan sudah diperoleh, tahapan selanjutnya adalah melakukan pembersihan dan pra-pemrosesan data dengan tujuan untuk memastikan konsistensi dan akurasi data sebelum tahap analisis. Dalam tahap ini terdapat beberapa langkah yang dilakukan langkah awal yaitu dengan melakukan penanganan data seperti penanganan terhadap nilai yang hilang (*missing value*) serta penghapusan spasi untuk standarisasi format teks pada *DataFrame*.

Setelah data yang diperlukan sudah diperoleh, tahapan selanjutnya adalah melakukan pembersihan dan pra-pemrosesan data dengan tujuan untuk memastikan konsistensi dan akurasi data sebelum tahap analisis. Terdapat beberapa langkah yang dilakukan pada tahap ini yaitu dimulai dengan melakukan penanganan data seperti penanganan terhadap nilai yang hilang (*missing value*) serta penghapusan spasi untuk standarisasi format teks pada *DataFrame*.

Selanjutnya dilakukan transformasi pada kolom, seperti Beberapa nama kolom diubah agar lebih representatif untuk analisis, seperti 'Tempat lahir' menjadi 'Daerah Asal', 'Status' menjadi 'Status Kelulusan', dan 'IPK Lulus' menjadi 'IPK' yang ditunjukkan dalam gambar 2.

	TempatLahir	Status	IPK Lulus
0	WERWAWAN	Lulus	3.14
1	LETI	Lulus	3.28
2	LUTUR	Lulus	3.05
3	PENCADO	Lulus	3.22
4	PAPERU	Lulus	3.37
...
771	PASSO	Lulus	3.84
772	WAEMULANG	Lulus	3.65
773	MASOHI	Lulus	3.97
774	BAL-BALU	Lulus	3.72
775	AMBON	Lulus	3.71

	DaerahAsal	StatusKelulusan	IPK
0	Werwawan	Lulus	3.14
1	Leti	Lulus	3.28
2	Lutur	Lulus	3.05
3	Pencado	Lulus	3.22
4	Paperu	Lulus	3.37
...
771	Passo	Lulus	3.84
772	Waemulang	Lulus	3.65
773	Masohi	Lulus	3.97
774	Bal-Balu	Lulus	3.72
775	Ambon	Lulus	3.71

Gambar 2. atribut sebelum dan sesudah dibersihkan

Selanjutnya dilakukan perubahan nilai pada kolom status kelulusan, nilai 'Aktif' dalam kolom Status Kelulusan diganti menjadi 'Belum/Tidak Lulus'. Ini menyederhanakan kategorisasi status kelulusan menjadi dua kelompok utama yaitu 'Lulus' dan 'Belum/Tidak Lulus'.

	Nama	Status
47	FINI AFYONITA RISAKOTTA	AKTIF
48	MARNEIX A. MASBAITUBUN	AKTIF
49	MERRY FIONA SIPAKOLY	AKTIF
50	POPI SALEKY	AKTIF
51	YONAS LUANMASAR	AKTIF
...
734	YOHANA PATOTNEM	AKTIF
735	YULINDA BERHARD	AKTIF
736	DESI DIANA MOFU	AKTIF
737	NOVENSKA ALEXSANDRA	AKTIF
738	ANASTASYA TRIFENA PATTILEAMONIA	AKTIF

	Nama	StatusKelulusan
47	FINI AFYONITA RISAKOTTA	Belum/Tidak Lulus
48	MARNEIX A. MASBAITUBUN	Belum/Tidak Lulus
49	MERRY FIONA SIPAKOLY	Belum/Tidak Lulus
50	POPI SALEKY	Belum/Tidak Lulus
51	YONAS LUANMASAR	Belum/Tidak Lulus
...
734	YOHANA PATOTNEM	Belum/Tidak Lulus
735	YULINDA BERHARD	Belum/Tidak Lulus
736	DESI DIANA MOFU	Belum/Tidak Lulus
737	NOVENSKA ALEXSANDRA	Belum/Tidak Lulus
738	ANASTASYA TRIFENA PATTILEAMONIA	Belum/Tidak Lulus

Gambar 3. Atribut status kelulusan sebelum dan sesudah dibersihkan

Langkah terakhir adalah *mapping* (pemetaan) kabupaten berdasarkan daerah asal. Pada langkah ini dilakukan penambahan kolom yaitu kolom kabupaten untuk mengelompokkan kabupaten berdasarkan daerah asal mahasiswa dengan mengisi nilai berdasarkan pencarian pada kamus kabupaten yang sudah dibuat. Serta daerah asal yang tidak teridentifikasi dalam kamus menjadi kategori 'Lainnya'.

	DaerahAsal	Kabupaten
0	Werwawan	Kabupaten Maluku Barat Daya
1	Leti	Kabupaten Maluku Barat Daya
2	Lutur	Kabupaten Kepulauan Aru
3	Pencado	Kabupaten Maluku Barat Daya
4	Paperu	Kabupaten Maluku Tengah
...
95	Ambon	Kota Ambon
96	Haria	Kabupaten Maluku Tengah
97	Kamal	Kabupaten Seram Bagian Barat
98	Kojjabi	Kabupaten Kepulauan Aru
99	Waeturen	Kabupaten Buru Selatan

Gambar 5. Hasil Mapping

DaerahAsal	Kabupaten
-	Lainnya
Kaktuan	Lainnya
Kaktuan	Lainnya
Nan	Lainnya
Nan	Lainnya
Nan	Lainnya
Sabuai	Lainnya
Nan	Lainnya
Mantarara	Lainnya
Waurtahait	Lainnya
Letpei	Lainnya
Waewali	Lainnya
Wamesa	Lainnya

Gambar 6. Daerah asal yang tidak teridentifikasi

3.3 Analisis deskriptif

Pada Tahapan ini disajikan gambaran umum statistik dari data yang telah dibersihkan. Bertujuan untuk melihat pola sebaran dan profil akademik mahasiswa.

3.3.1 Sebaran mahasiswa

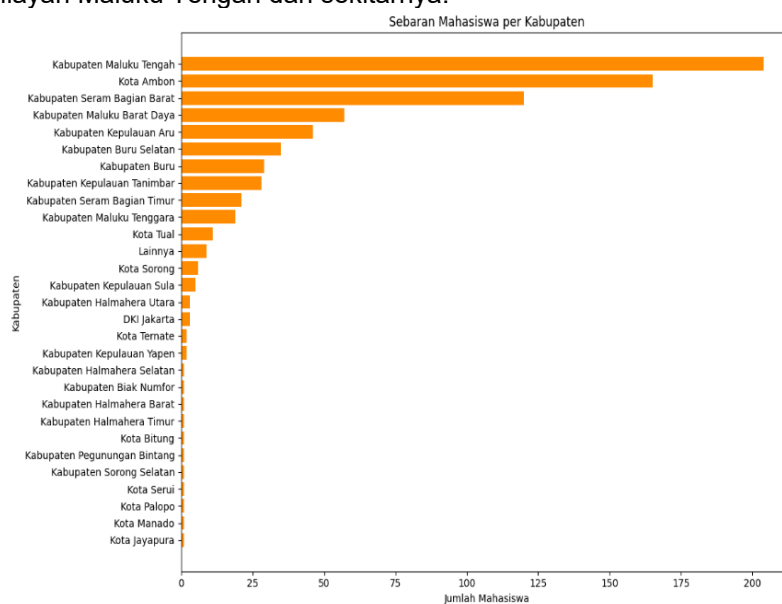
Berdasarkan data yang diolah, terlihat bahwa konsentrasi mahasiswa terpusat pada beberapa wilayah utama di Maluku. Tabel 1 memperlihatkan data jumlah mahasiswa per kabupaten. Dapat dilihat bahwa Kabupaten Maluku Tengah menjadi penyumbang mahasiswa terbanyak dengan 204 orang, diikuti oleh Kota Ambon dengan 165 orang, dan Kabupaten Seram Bagian Barat dengan 120 orang. Sebaliknya, terdapat daerah-daerah dengan jumlah mahasiswa yang sangat sedikit, seperti dari wilayah Papua atau Sulawesi yang hanya berjumlah 1 orang per kabupaten.

Tabel 1. Sebaran Mahasiswa Per Kabupaten

No	Kabupaten	Jumlah Mahasiswa
1	Kabupaten Maluku Tengah	204
2	Kota Ambon	165
3	Kabupaten Seram Bagian Barat	120
4	Kabupaten Maluku Barat Daya	57
5	Kabupaten Kepulauan Aru	46
6	Kabupaten Buru Selatan	35
7	Kabupaten Buru	29
8	Kabupaten Kepulauan Tanimbar	28
9	Kabupaten Seram Bagian Timur	21
10	Kabupaten Maluku Tenggara	19
11	Kota Tual	11
12	Lainnya	9
13	Kota Sorong	6
14	Kabupaten Kepulauan Sula	5
15	Kabupaten Halmahera Utara	3
16	DKI Jakarta	3
17	Kota Ternate	2
18	Kabupaten Kepulauan Yapen	2
19	Kabupaten Halmahera Selatan	1
20	Kabupaten Biak Numfor	1
21	Kabupaten Halmahera Barat	1

No	Kabupaten	Jumlah Mahasiswa
22	Kabupaten Halmahera Timur	1
23	Kota Bitung	1
24	Kabupaten Pegunungan Bintang	1
25	Kabupaten Sorong Selatan	1
26	Kota Serui	1
27	Kota Palopo	1
28	Kota Manado	1
29	Kota Jayapura	1

Untuk mempermudah pemahaman mengenai dominasi daerah asal mahasiswa, visualisasi data disajikan pada Gambar 7. Grafik ini mempertegas bahwa mayoritas mahasiswa berasal dari wilayah Maluku Tengah dan sekitarnya.



Gambar 7. Sebaran Mahasiswa

3.3.2 Persentase Kelulusan

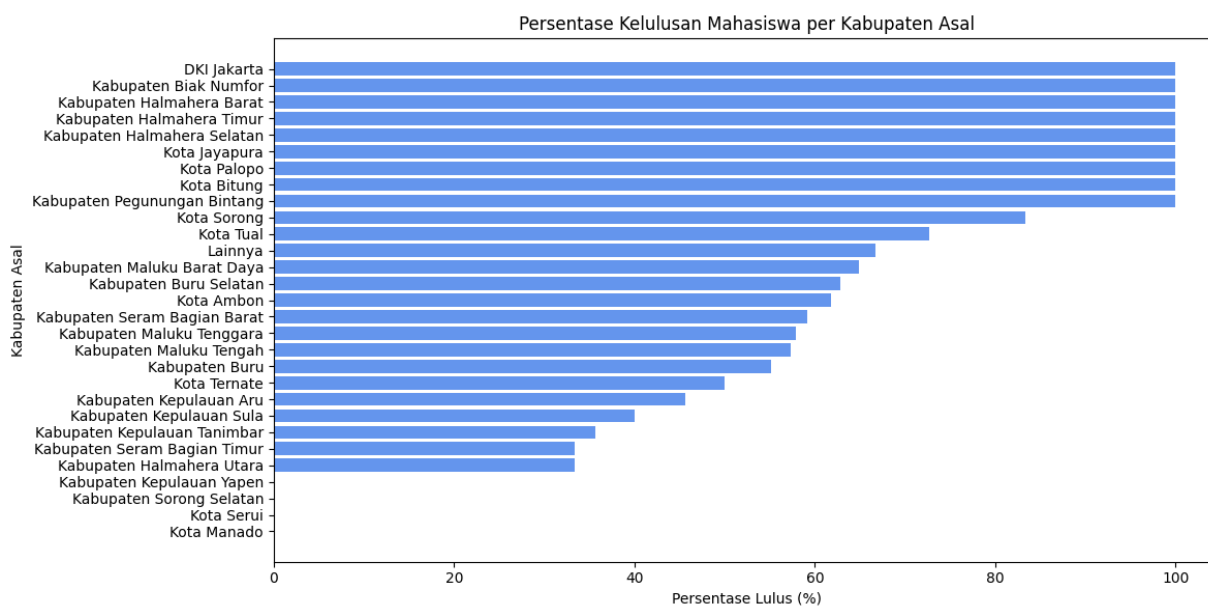
Setelah mengetahui jumlah mahasiswa berdasarkan asal daerah, analisis dilanjutkan dengan melihat seberapa besar peluang kelulusan mahasiswa dari masing-masing daerah tersebut. Data rinci mengenai status kelulusan disajikan pada Tabel 2. Beberapa daerah seperti DKI Jakarta, Biak Numfor, dan Halmahera Barat mencatatkan persentase kelulusan sempurna (100%), meskipun jumlah mahasiswanya relatif sedikit. Di sisi lain, daerah dengan jumlah mahasiswa besar seperti Kota Ambon dan Maluku Tengah memiliki tingkat kelulusan di kisaran 57% - 61%. Ada pula daerah dengan tingkat kelulusan 0%, seperti Kota Manado dan Kota Serui.

Tabel 2. Persentase Kelulusan Mahasiswa Per Kabupaten

No	Kabupaten	Belum/Tidak Lulus	Lulus	Persentase Lulus
1	DKI Jakarta	0	3	100%
2	Kabupaten Biak Numfor	0	1	100%
3	Kabupaten Halmahera Barat	0	1	100%
4	Kabupaten Halmahera Timur	0	1	100%
5	Kabupaten Halmahera Selatan	0	1	100%
6	Kota Jayapura	0	1	100%
7	Kota Palopo	0	1	100%
8	Kota Bitung	0	1	100%
9	Kabupaten Pegunungan Bintang	0	1	100%

No	Kabupaten	Belum/Tidak Lulus	Lulus	Persentase Lulus
10	Kota Sorong	1	5	83,3%
11	Kota Tual	3	8	72,7%
12	Lainnya	3	6	66,6%
13	Kabupaten Maluku Barat Daya	20	37	64,9%
14	Kabupaten Buru Selatan	13	22	62,8%
15	Kota Ambon	63	102	61,8%
16	Kabupaten Seram Bagian Barat	49	71	59,1%
17	Kabupaten Maluku Tenggara	8	11	57,8%
18	Kabupaten Maluku Tengah	87	117	57,3%
19	Kabupaten Buru	13	16	55,1%
20	Kota Ternate	1	1	50%
21	Kabupaten Kepulauan Aru	25	21	45,6%
22	Kabupaten Kepulauan Sula	3	2	40%
23	Kabupaten Kepulauan Tanimbar	18	10	35,7%
24	Kabupaten Seram Bagian Timur	14	7	33,3%
25	Kabupaten Halmahera Utara	2	1	33,3%
26	Kabupaten Kepulauan Yapen	2	0	0%
27	Kabupaten Sorong Selatan	1	0	0%
28	Kota Serui	1	0	0%
29	Kota Manado	1	0	0%

Perbandingan persentase ini divisualisasikan dalam bentuk grafik batang pada Gambar 8, yang memudahkan kita melihat urutan daerah dari tingkat kelulusan tertinggi hingga terendah.



Gambar 8. Visualisasi Persentase Kelulusan

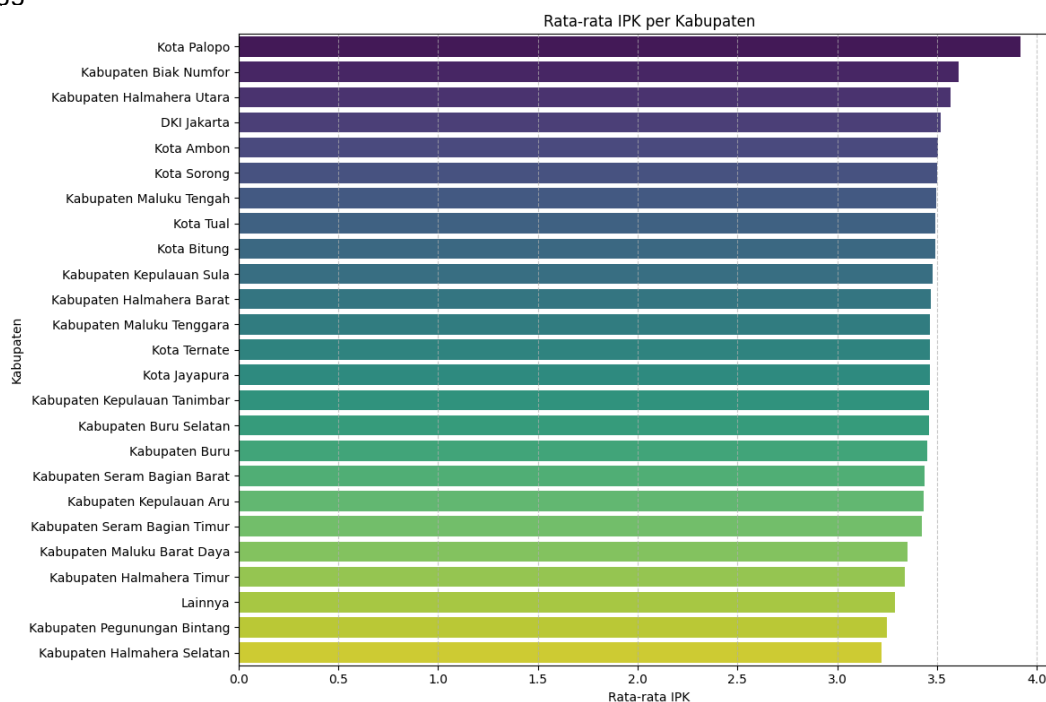
3.3.3 Rata-rata indeks prestasi kumulatif (IPK)

Selain status kelulusan, kualitas akademik juga diukur melalui rata-rata IPK. Secara umum, daerah dengan rata-rata IPK yang tinggi cenderung memiliki tingkat kelulusan yang lebih baik, yang mengindikasikan kesiapan akademik yang matang. Tabel 3 menunjukkan peringkat rata-rata IPK berdasarkan kabupaten asal. Kota Palopo memimpin dengan rata-rata IPK 3.92, diikuti oleh Kabupaten Biak Numfor dengan 3.61. Kabupaten-kabupaten di Maluku seperti Kota Ambon dan Maluku Tengah berada pada kisaran IPK 3.49 – 3.50, yang menunjukkan performa yang cukup baik.

Tabel 3. Rata-Rata IPK

No	Kabupaten	Rata-Rata IPK
1	Kota Palopo	3.92
2	Kabupaten Biak Numfor	3.61
3	Kabupaten Halmahera Utara	3.56
4	DKI Jakarta	3.51
5	Kota Ambon	3.50
6	Kota Sorong	3.50
7	Kabupaten Maluku Tengah	3.49
8	Kota Tual	3.49
9	Kota Bitung	3.49
10	Kabupaten Kepulauan Sula	3.47
11	Kabupaten Halmahera Barat	3.47
12	Kabupaten Maluku Tenggara	3.46
13	Kota Ternate	3.46
14	Kota Jayapura	3.46
15	Kabupaten Kepulauan Tanimbar	3.46
16	Kabupaten Buru Selatan	3.46
17	Kabupaten Buru	3.45
18	Kabupaten Seram Bagian Barat	3.43
19	Kabupaten Kepulauan Aru	3.43
20	Kabupaten Seram Bagian Timur	3.42
21	Kabupaten Maluku Barat Daya	3.35
22	Kabupaten Halmahera Timur	3.34
23	Lainnya	3.28
24	Kabupaten Pegunungan Bintang	3.25
25	Kabupaten Halmahera Selatan	3.22

Distribusi rata-rata IPK ini dapat dilihat lebih jelas pada Gambar 9. Grafik ini membantu mengidentifikasi daerah mana yang mahasiswanya secara konsisten meraih nilai akademik tinggi.



Gambar 9. Visualisasi Rata-Rata IPK

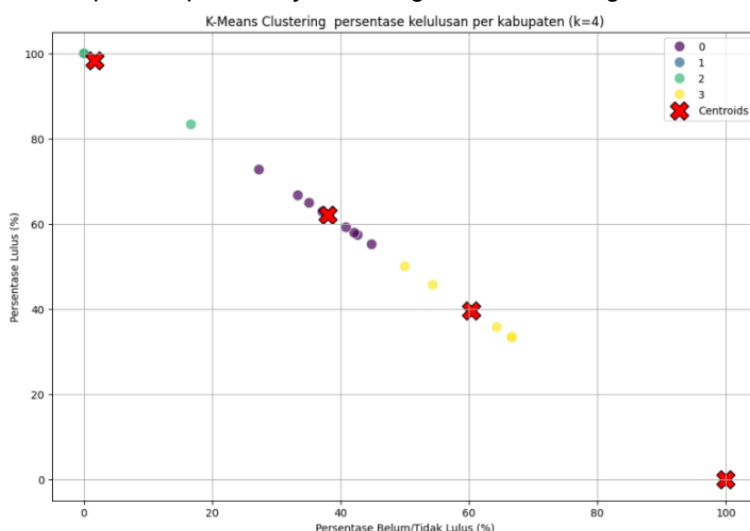
3.4 Analisis Clustering (K-Means)

Untuk mendapatkan wawasan yang lebih strategis, penelitian ini menggunakan metode K-Means Clustering. Tujuannya adalah membentuk kelompok dari kabupaten-kabupaten menjadi beberapa kategori berdasarkan kemiripan pola kelulusannya. Penelitian ini menetapkan 4 klaster (k=4) sebagai pembagian yang paling optimal. Pusat penyebaran data atau centroid untuk setiap klaster yang terbentuk ditunjukkan pada Tabel 4. Tabel ini menjadi acuan untuk memahami karakteristik setiap kelompok.

Tabel 4. Nilai Centroid (pusat)

Klaster	Belum/Tidak Lulus (%)	Persentase_Lulus (%)
0	37.94	62.06
1	100.00	0.00
2	1.67	98.33
3	60.33	39.67

Proses pengelompokan ini divisualisasikan dalam Gambar 10, di mana setiap titik mewakili satu kabupaten/kota dan warnanya menunjukkan keanggotaan klasternya. Titik pusat atau *centroid* untuk setiap kelompok ditunjukkan dengan tandan silang berwarna merah.



Gambar 10. Clustering

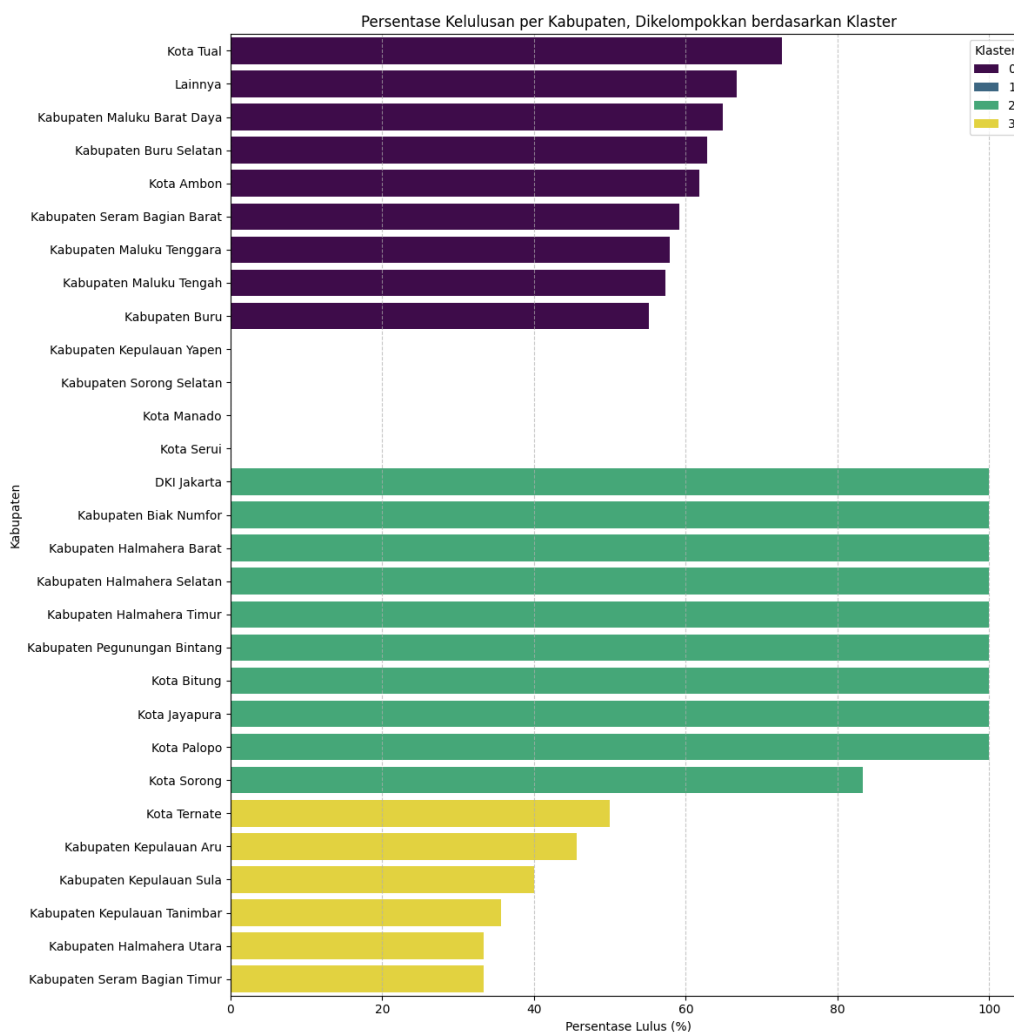
Proses klasterisasi ini menghasilkan empat kategori kinerja kabupaten sebagaimana dijelaskan secara rinci pada Tabel 5.

1. **Klaster 0 (Kinerja Moderat):** Kelompok ini berisi daerah penyumbang mahasiswa terbesar seperti Kota Ambon dan Maluku Tengah. Tingkat kelulusannya sedang (sekitar 62%), artinya mayoritas lulus namun masih ada sepertiga yang belum tuntas.
2. **Klaster 1 (Kinerja Rendah/Tanpa Lulusan):** Berisi daerah yang mahasiswanya belum ada yang lulus sama sekali, seperti Kota Manado dan Kepulauan Yapen.
3. **Klaster 2 (Kinerja Tinggi):** Ini adalah kelompok "bintang", berisi daerah dengan tingkat kelulusan hampir sempurna (98%), seperti DKI Jakarta dan Biak Numfor.
4. **Klaster 3 (Kinerja Menantang):** Kelompok ini butuh perhatian khusus karena tingkat ketidaklulusannya (60%) lebih tinggi daripada yang lulus, contohnya Kabupaten Kepulauan Aru dan Tanimbar.

Tabel 5. Hasil Akhir *Clustering*

Klaster	Kategori	Lulus (%)	Belum/Tidak Lulus (%)	Rata-Rata IPK	Keterangan
0	Kinerja Moderat	62.06	37.94	3.44	Klaster ini mewakili kabupaten dengan tingkat kelulusan sedang dan proporsi mahasiswa yang belum/tidak lulus yang relatif moderat. Ini adalah daerah di mana sebagian besar mahasiswa lulus. Diantaranya (Kabupaten Buru Selatan, Kabupaten Buru, Kabupaten Maluku Barat Daya, Kabupaten Maluku Tengah, Kabupaten Seram Bagian Barat, Kabupaten Maluku Tenggara, Kota Ambon, Kota Tual, Lainnya).
1	Kinerja Rendah/Tanpa Lulusan	0.00	100.00	-	Klaster ini berisi kabupaten di mana hampir semua mahasiswa 'Belum/Tidak Lulus'. Diantaranya (Kabupaten Kepulauan Yapen, Kabupaten Sorong Selatan, Kota Manado, Kota Serui).
2	Kinerja Tinggi	98.33	1.67	3.50	Klaster ini menunjukkan kabupaten dengan tingkat kelulusan yang sangat tinggi dan persentase mahasiswa yang 'Belum/Tidak Lulus' yang sangat rendah. Ini adalah wilayah yang sangat berhasil dalam hal hasil kelulusan mahasiswa. Diantaranya (DKI Jakarta, Kabupaten Biak Numfor, Kabupaten Halmahera Barat, Kabupaten Halmahera Timur, Kabupaten Halmahera Selatan, Kabupaten Pegunungan Bintang, Kota Bitung, Kota Jayapura, Kota Palopo, Kota Sorong).
3	Kinerja Menantang	39.67	60.33	3.47	Klaster ini mewakili kabupaten di mana proporsi mahasiswa yang 'Belum/Tidak Lulus' lebih tinggi dibandingkan dengan mereka yang telah lulus. Ini mungkin daerah yang menghadapi tantangan dalam kemajuan dan kelulusan mahasiswa, membutuhkan perhatian lebih dekat. Diantaranya (Kabupaten Halmahera Utara, Kabupaten Kepulauan Aru, Kabupaten Kepulauan Sula, Kabupaten Kepulauan Tanimbar, Kabupaten Seram Bagian Timur, Kota Ternate).

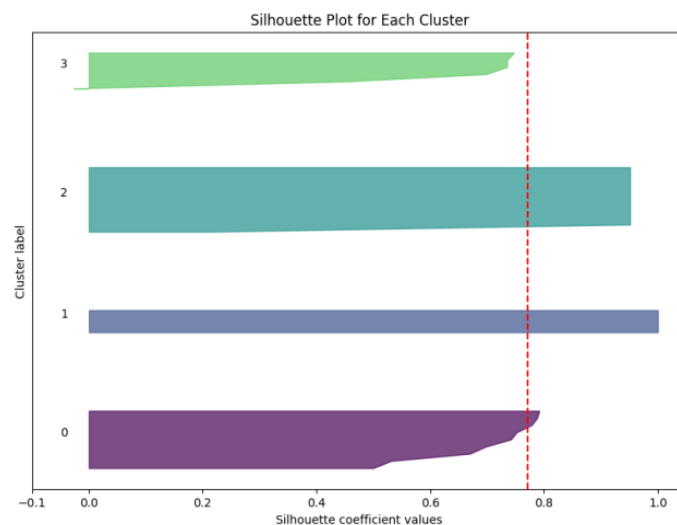
Untuk melihat perbandingan antar klaster secara lebih mudah, Gambar 10 menyajikan grafik persentase kelulusan yang sudah dikelompokkan berdasarkan warnanya. Untuk memastikan bahwa pengelompokan di atas valid dan objektif, dilakukan pengujian menggunakan metode *Silhouette Analysis*. Metode ini mengukur seberapa rapi pemisahan antar kelompok. Hasil pengujian menunjukkan nilai rata-rata *Silhouette Score* sebesar 0.77. Berdasarkan hasil pengujian, diperoleh nilai *Silhouette Score* rata-rata sebesar 0,77 yang mengindikasikan bahwa struktur klaster yang terbentuk sudah sangat baik, padat, dan tingkat separasi yang jelas antar klaster. Grafik *Silhouette* dapat dilihat pada Gambar 11.



Gambar 10. Persentase Kelulusan Mahasiswa per Kabupaten berdasarkan Klaster

Temuan dari analisis ini mengindikasikan adanya diferensiasi tingkat kelulusan mahasiswa yang berasal dari kabupaten yang berbeda, khususnya di antara wilayah dengan kontribusi mahasiswa yang signifikan. Hasil clustering memperkuat observasi ini dengan menunjukkan bahwa kabupaten/kota di Maluku terbagi ke dalam empat klaster dengan karakteristik kelulusan yang berbeda. Klaster 0, dengan tingkat kelulusan rata-rata 62.06%, termasuk dalam kategori Kinerja Moderat karena masih terdapat proporsi mahasiswa yang belum lulus sebesar 37.94% dengan rata-rata IPK 3.44. Sementara itu, Klaster 1 menunjukkan Kinerja Rendah/Tanpa Lulusan dengan persentase kelulusan 0% dan seluruh mahasiswa masuk kategori belum/tidak lulus. Di sisi lain, Klaster 2 menunjukkan Kinerja Tinggi dengan tingkat kelulusan yang sangat baik, yaitu 98.33%, serta persentase mahasiswa belum/tidak lulus yang sangat rendah (1.67%) dengan rata-rata IPK 3.50. Terakhir, Klaster 3 termasuk dalam kategori Kinerja Menantang dengan rata-rata kelulusan 39.67% dan proporsi belum/tidak

lulus yang cukup tinggi (60.33%) dengan rata-rata IPK 3.47, menandakan bahwa wilayah dalam klaster ini masih menghadapi tantangan signifikan terkait penyelesaian studi mahasiswa.



Gambar 11. Silhouette Score

Berdasarkan hasil temuan dapat dilihat bahwa Kabupaten dengan rata-rata IPK lebih tinggi cenderung berada di klaster dengan persentase kelulusan yang lebih tinggi, mengindikasikan bahwa performa akademik secara umum di wilayah tersebut lebih baik, namun terdapat anomali pada Klaster 3 (Kinerja Menantang) dengan memiliki tingkat kelulusan terendah (selain Klaster 1) Namun, rata-rata IPK-nya secara mengejutkan lebih tinggi daripada Klaster 0. Ini menegaskan poin yang sebelumnya diangkat yaitu mahasiswa di klaster ini mungkin memiliki kemampuan akademik (IPK) yang baik, tetapi terhambat oleh faktor non-akademik atau struktural lain dalam menyelesaikan studi yang memerlukan penelitian lebih lanjut. Perbedaan tingkat kelulusan antar klaster ini mengisyaratkan adanya potensi pengaruh faktor-faktor eksternal, terkait dengan asal geografis mahasiswa, terhadap keberhasilan studi mereka di perguruan tinggi.

3.5 Pengujian Performa Algoritma

Tahap pengujian merupakan langkah krusial untuk memvalidasi apakah pengelompokan yang dihasilkan oleh algoritma *K-Means* memiliki tingkat kohesi dan separasi yang memadai. Mengingat penelitian ini menggunakan pendekatan *unsupervised learning*, evaluasi dilakukan dengan menggunakan metode *Silhouette Coefficient*. Proses pengujian ini dilakukan dengan mengukur seberapa serupa sebuah objek dengan klasternya sendiri dibandingkan dengan klaster lainnya.

Proses perhitungan dilakukan secara sistematis dengan mengevaluasi nilai *silhouette* untuk setiap titik data dalam empat klaster yang telah terbentuk. Hasil pengujian menunjukkan nilai *Average Silhouette Score* sebesar 0,77. Berdasarkan skala interpretasi kualitas klaster, nilai yang berada pada rentang 0,71 hingga 1,00 dikategorikan sebagai *Strong Structure* (Struktur Kuat). Hal ini membuktikan bahwa algoritma telah berhasil memisahkan data mahasiswa berdasarkan asal daerah ke dalam kelompok-kelompok yang sangat kontras dan memiliki keanggotaan yang solid. Dengan demikian, hasil pengelompokan ini tidak terjadi secara acak, melainkan merepresentasikan pola akademik yang nyata dari data historis tahun 2010–2020, sehingga model ini memiliki reliabilitas yang tinggi untuk dijadikan instrumen evaluasi program studi.

3.6 Hasil temuan

Hasil analisis data menunjukkan adanya stratifikasi performa akademik yang cukup jelas berdasarkan asal daerah mahasiswa. Temuan ini tidak hanya memberikan gambaran statistik, tetapi juga memberikan ruang diskusi mendalam mengenai dinamika pendidikan di tingkat regional.

3.6.1 Relevansi dan Integrasi Temuan dengan Penelitian Terdahulu

Temuan dalam penelitian ini memberikan penguatan empiris terhadap hasil riset terdahulu yang menekankan pentingnya pemetaan kelompok dalam dunia pendidikan. Nilai Silhouette Score sebesar 0,77 menunjukkan bahwa struktur klaster yang terbentuk memiliki tingkat kohesi dan separasi yang baik. Temuan ini konsisten dengan penelitian Bagirov et al. yang menjelaskan bahwa *Silhouette Coefficient* merupakan indeks validitas klaster yang efektif untuk mengevaluasi kekompakan dan keterpisahan klaster, sehingga menghasilkan pengelompokan yang lebih reliabel dan terukur [19].

Lebih lanjut, adanya identifikasi kelompok dengan kinerja rendah (seperti pada Klaster 1 dan 3) memberikan validasi terhadap studi yang dilakukan, yang menggarisbawahi bahwa perbedaan latar belakang geografis seringkali berkorelasi dengan aksesibilitas informasi dan kesiapan akademik mahasiswa di perguruan tinggi [20] [21]. Integrasi temuan ini ke dalam literatur yang ada menunjukkan bahwa fenomena kesenjangan akademik antar-wilayah adalah pola yang konsisten, dan penggunaan *data mining* seperti *K-Means* [22] merupakan solusi saintifik yang efektif untuk memetakan kondisi tersebut secara objektif.

3.6.2 Kontribusi Temuan terhadap Pengembangan Ilmu Pengetahuan

Kontribusi signifikan yang dapat diberikan penelitian ini terhadap bidang *Educational Data Mining* (EDM) dan Sistem Informasi, khususnya dalam pemanfaatan teknik analitik data untuk mendukung pengambilan keputusan yang lebih efektif dan berbasis data. Secara teoritis, penelitian ini memperluas penerapan konsep *knowledge discovery in databases* (KDD) dengan mengintegrasikan aspek geospasial (asal kabupaten) ke dalam analisis performa kelulusan mahasiswa. Hal ini memberikan perspektif baru bagi pengembangan ilmu manajemen pendidikan berbasis data (*data-driven management*).

Secara praktis, kontribusi utama penelitian ini adalah terciptanya sebuah kerangka kerja evaluasi yang dapat diadopsi oleh institusi untuk merancang sistem peringatan dini (*early warning system*). Dengan mengetahui klaster asal daerah yang memiliki risiko kelulusan rendah, pihak program studi dapat merumuskan kebijakan intervensi yang lebih spesifik, seperti program matrikulasi atau pendampingan akademik intensif bagi mahasiswa dari wilayah tertentu. Hal ini menggeser paradigma evaluasi pendidikan yang sebelumnya bersifat umum (pemerataan perlakuan) menjadi evaluasi yang berbasis pada kebutuhan dan karakteristik spesifik kelompok mahasiswa.

4. Simpulan

Penelitian ini menganalisis tingkat kelulusan mahasiswa tahun 2010-2020 berdasarkan asal kabupaten. Sebaran mahasiswa terkonsentrasi di Kabupaten Maluku Tengah, Kota Ambon, dan Kabupaten Seram Bagian Barat. Analisis klastering pada kabupaten dengan k-4 mengidentifikasi 4 kelompok utama berdasarkan persentase kelulusan: Klaster 0 Kinerja Moderat (rata-rata ~62,06% lulus, mencakup Kabupaten Buru Selatan, Kabupaten Buru, Kabupaten Maluku Tengah, Kabupaten Maluku Barat Daya, Kabupaten Maluku Tenggara, Kabupaten Seram Bagian Barat, Kota Ambon, Kota Tual, Lainnya), Klaster 1 kinerja rendah/tanpa lulus (rata-rata ~0% lulus, mencakup Kabupaten Kepulauan Yapen, Kabupaten Sorong Selatan, Kota Manado, Kota Serui), Klaster 2 Kinerja Tinggi (rata-rata ~98,33% lulus, mencakup DKI Jakarta, Kabupaten Biak Numfor, Kabupaten Halmahera Barat, Kabupaten Halmahera Timur, Kabupaten Halmahera Selatan, Kabupaten Pegunungan Bintang, Kota Bitung, Kota Jayapura, Kota Palopo, Kota Sorong). Temuan ini menunjukkan adanya diferensiasi tingkat kelulusan yang berkorelasi dengan performa akademik mahasiswa di daerah tersebut berdasar pada data 2010-2020. Hasil ini dapat menjadi dasar bagi institusi untuk merancang program dukungan mahasiswa yang lebih bertarget. Untuk penelitian mendatang, disarankan menggunakan data yang lebih baru dan variabel yang lebih lengkap untuk analisis yang lebih mendalam.

Daftar Referensi

- [1] E. Alyahyan and D. Düşteğör, "Predicting academic success in higher education: literature review and best practices," *International Journal of Educational Technology in Higher Education*, vol. 17, no. 1, pp. 57-77, Dec. 2020, doi: 10.1186/s41239-020-0177-7.
- [2] S. Rizky Rahayu and D. Murni, "Analisis Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif Mahasiswa Dapartemen Matematika FMIPA UNP Dengan Pendekatan Metode CHAID," *Jurnal Pendidikan Tambusai*, vol. 8, no. 2, pp. 29858–29868, Aug. 2024.
- [3] Hadi Putra, Khairunnisa Nasution, and Elkin Rilvani, "Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Decision Tree: Studi Perbandingan Algoritma ID3 DAN C4.5," *JURNAL MEDIA AKADEMIK*, vol. 3, no. 7, pp.49-63, Jul. 2025.
- [4] Destu Satya Widyaningsih, "Hubungan Antara Motivasi Belajar Dan Asal Daerah Dengan Prestasi Belajar Mahasiswa Di Akademi Manajemen Administrasi Yogyakarta," *JURNAL AMA YOGYAKARTA*, vol. 8, pp.29-38, 2017.
- [5] H. Ermewaningsih, F. E. Renni Susanti, and Y. Yus, "Klasterisasi Kinerja Akademik Mahasiswa Berdasarkan IPK Menggunakan K-Means dan Metode Elbow," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 6, no. 3, pp. 619–625, Apr. 2026.
- [6] R. A. H. Annafi, "Pengaruh tingkat prokrastinasi akademik pada mahasiswa ditinjau dari asal daerah," Skripsi, Prodi Psikologi, Universitas Islam Negeri Maulana Malik Ibrahim, Malang, 2025.
- [7] M. W. Agustina and Putri Deastuti Pracintya Wahyu, "Hardiness Dan Stress Akademik Pada Mahasiswa Rantau," *IDEA: Jurnal Psikologi*, vol. 7, no. 1, pp. 34–45, Apr. 2023.
- [8] C. P. Lawrence and M. Mudzakkir, "Penyesuaian Diri Mahasiswa Rantau Indonesia Timur di Universitas Negeri Surabaya," *The Sociology Journal Paradigma*, vol. 13, no. 1, pp. 11–20, Jan. 2024.
- [9] W. Atika, "Analisis Efisiensi Dan Produktivitas Program Studi S – 1 Di Fakultas Teknik Universitas Negeri Yogyakarta," Skripsi, Prodi Pendidikan Teknik Elektronika, Universitas Negeri Yogyakarta, Yogyakarta, 2017.
- [10] H. A. Madni, Z. Anwar, and M. A. Shah, "Data mining techniques and applications — A decade review," in *23rd International Conference on Automation and Computing (ICAC)*, IEEE, Sep. 2017, pp. 1–7. doi: 10.23919/IconAC.2017.8082090.
- [11] S. Harikumar and P. v. Surya, "K-Medoid Clustering for Heterogeneous DataSets," in *4th International Conference on Eco-friendly Computing and Communication Systems (ICECCS)*, Kuruksheetra, pp. 226–237, 7-8 Dec. 2015, doi: 10.1016/j.procs.2015.10.077.
- [12] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, vol. 3, no. 1, pp. 46–56, Mar. 2024, doi: 10.58602/jima-ilkom.v3i1.26.
- [13] Bakhtiar Robi Tegar, Herlambang Agus Bambang, and Anam Khoiril Ahmad, "Penggunaan SIG untuk Pemetaan dan Analisis Jumlah Sebaran mahasiswa dan Perguruan Tinggi di Kabupaten Brebes," *Jurnal Ilmiah Research Student*, vol. 1, no. 3, pp. 641–650, Jan. 2024.
- [14] F. Rolansa, "Pengembangan interaktif dashboard kemahasiswaan di program studi teknik informatika dengan teknologi big data," *Jurnal Pendidikan Informatika dan Sains*, vol. 10, no. 2, pp. 110–118, Dec. 2021, doi: 10.31571/saintek.v10i2.2190.
- [15] A. Ramadhan, F. Achmad, I. Zulkarnain, and M. Aritsugi, "Evaluation of K-Means, DBSCAN, and Hierarchical Clustering for Strategic Segmentation of Tourism SMEs in Rembang, Indonesia," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 3, pp. 1605–1630, Jul. 2025, doi: 10.52436/1.jutif.2025.6.3.4602.
- [16] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means Algorithm: A Comprehensive Survey and Performance Evaluation," *Electronics (Basel)*, vol. 9, no. 8, p. 1295, Aug. 2020, doi: 10.3390/electronics9081295.
- [17] B. Syaeful Hadi, "Urgensi Dimensi Faktor Geografis Dalam Penentuan Kebijakan Komunikasi Dan Informasi," *Geomedia: Majalah Ilmiah dan Informasi Kegeografian*, vol. 5, no. 1, pp.129-136, May 2017, doi: 10.21831/gm.v5i1.14209.
- [18] T. T. D. Susanto, L. Syafruddin, and S. F. Abdullah, "Peningkatan Kualitas Pengambilan Keputusan Melalui Manajemen Berbasis Data Di Sekolah," *JURNAL MADINASIKA Manajemen Pendidikan dan Keguruan*, vol. 6, no. 2, pp. 235–250, Jun. 2025, doi: 10.31949/madinaska.v6i2.13723.

- [19] A. M. Bagirov, R. M. Aliguliyev, and N. Sultanova, "Finding compact and well-separated clusters: Clustering using silhouette coefficients," *Pattern Recognition*, vol. 135, pp.109-144. 2023, doi: 10.1016/j.patcog.2022.109144.
- [20] Kusaeri and N. Wahyudi, "The Relative Influence of Age, Gender, Regional Differences and Their Relationship with Students' Numeracy Levels on Madrasah Competency Assessment in Indonesia," *JPI (Jurnal Pendidikan Indonesia)*, vol. 13, no. 2, pp. 288–297, 2024, doi: 10.23887/jpiundiksha.v13i2.63570.
- [21] R. Lembani, A. Gunter, M. Breines, and M. T. B. Dalu, "The same course, different access: the digital divide between urban and rural distance education students in South Africa," *Journal of Geography in Higher Education*, vol. 44, no. 1, pp. 70–84, 2020, doi: 10.1080/03098265.2019.1694876.
- [22] S.B.H. Sakur, M. Silangen, & D. Tuwohingide, "Penerapan Algoritma K-Means Cluster dan Metode TOPSIS pada Pemilihan Mahasiswa kunjungan Industri. *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 11, no. 3, pp. 851-860, 2022, doi:http://dx.doi.org/10.35889/jutisi.v11i3.1045.