

**PENGELOMPOKAN INDIKATOR PEMBANGUNAN MANUSIA
KABUPATEN/KOTA DI JAWA TIMUR MENGGUNAKAN K-MEANS++**

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Disusun Oleh :

ALFINATUZZAHRO

09030220045

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
2025**

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini.

Nama : Alfinatuzzahro
NIM : 09030220045
Program Studi : Matematika
Angkatan : 2020

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "PENGELOMPOKAN INDIKATOR PEMBANGUNAN MANUSIA KABUPATEN/KOTA DI JAWA TIMUR MENGGUNAKAN K-MEANS++". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang ditetapkan.

Sidoarjo, 28 Juli 2025

Yang menyatakan,



ALFINATUZZAHRO

NIM. 09030220045

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

Nama : Alfinatuzzahro

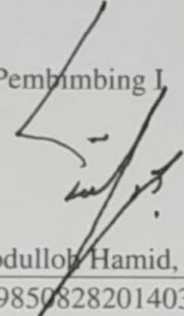
NIM : 09030220045

Judul skripsi : PENGELOMPOKAN INDIKATOR PEMBANGUNAN
MANUSIA KABUPATEN/KOTA DI JAWA TIMUR
MENGUNAKAN *K-MEANS++*

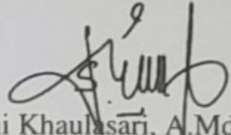
telah diperiksa dan disetujui untuk diujikan.

Surabaya, 3 Juni 2025

Pembimbing I


Dr. Abdulloh Hamid, M.Pd.
NIP. 198508282014031003

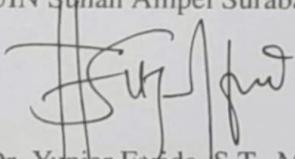
Pembimbing II


Hani Khaulasari, A.Md., S.Si., M.Si.
NIP. 199102092020122011

Mengetahui,

Ketua Program Studi Matematika

UIN Sunan Ampel Surabaya


Dr. Yuniar Farida, S.T., M.T.
NIP. 197905272014032002

LEMBAR PENGESAHAN TIM PENGUJI

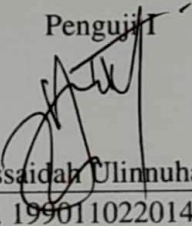
Skripsi oleh

Nama : Alfinatuzzahro
NIM : 09030220045
Judul Skripsi : PENGELOMPOKAN INDIKATOR PEMBANGUNAN
MANUSIA KABUPATEN/KOTA DI JAWA TIMUR
MENGUNAKAN *K-MEANS++*

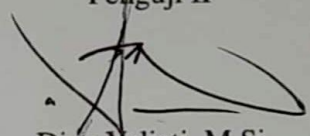
telah dipertahankan di depan tim penguji
pada tanggal 24 Juni 2025

Mengesahkan,
Tim Penguji

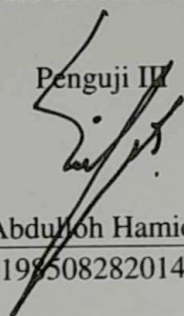
Penguji I


Nurissaidah Ulinnuha, M.Kom
NIP. 199011022014032004

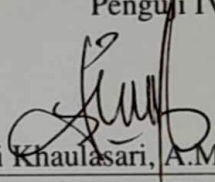
Penguji II


Dian Yuliati, M.Si
NIP. 198707142020122015

Penguji III


Dr. Abdulloh Hamid, M.Pd.
NIP. 198508282014031003

Penguji IV


Hani Khaulasari, A.Md., S.Si., M.Si.
NIP. 199102092020122011

Mengetahui,

Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya


Dr. A. Saepul Hamdani, M. Pd
NIP. 196507312000031002



UIN SUNAN AMPEL
SURABAYA

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Alfinatuzzahro
NIM : 09030220045
Fakultas/Jurusan : Sains dan teknologi / Matematika
E-mail address : alfinazahro2@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

☒ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :

Pengelompokan Indikator Pembangunan Manusia

Kabupaten/Kota di Jawa Timur Menggunakan K-Means++

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya,

Penulis

(Alfinatuzzahro)

nama terang dan tanda tangan

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN TIM PENGUJI	ii
PERNYATAAN KEASLIAN	iii
MOTTO	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
ABSTRAK	xii
ABSTRACT	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	11
1.3 Tujuan Penelitian	11
1.4 Manfaat Penelitian	12
1.4.1 Manfaat Teoritis	12
1.4.2 Manfaat Praktis	12
1.5 Batasan Masalah	12
1.6 Sistematika Penulisan	13
BAB II KAJIAN PUSTAKA	15
2.1 Indeks Pembangunan Manusia (IPM)	15
2.2 Umur Harapan Hidup (UHH)	15
2.3 Persentase Wanita Usia 15-49 Tahun Melahirkan Dibantu Dokter	16
2.4 Persentase Penduduk yang Memiliki Keluhan Kesehatan dan Rawat Jalan	16
2.5 Persentase Penduduk dengan Jaminan Kesehatan BPJS PBI	17

2.6	Harapan Lama Sekolah (HLS)	18
2.7	Rata-rata Lama Sekolah (RLS)	19
2.8	Angka Melek Huruf (AMH)	19
2.9	Indeks Pembangunan Literasi Masyarakat (IPLM)	20
2.10	Indeks Daya Beli (IDB)	21
2.11	Pengeluaran Per Kapita	22
2.12	Deteksi <i>Outlier</i>	22
2.13	Uji Kelayakan Analisis Faktor (KMO)	23
2.14	Uji Distribusi Normal Multivariat	24
2.15	Rumus <i>Variance Inflation Factor</i> (VIF)	25
2.16	Data Mining	26
2.17	Normalisasi Data	27
2.18	<i>Clustering</i>	28
2.19	Denormalisasi Data	29
2.20	<i>K-means++</i>	29
2.21	<i>Silhouette Coefficient</i>	32
2.22	Integrasi Keilmuan	33
2.22.1	Konsep Pembangunan dalam Islam	33
2.22.2	Integrasi Keilmuan dalam Pengembangan Sumber Daya Manusia	36
BAB III	METODE PENELITIAN	38
3.1	Jenis Penelitian	38
3.2	Variabel Penelitian	38
3.3	Langkah Analisis	39
BAB IV	HASIL DAN PEMBAHASAN	42
4.1	Uji Statistik Deskriptif	42
4.2	Deteksi <i>Outlier</i>	43
4.3	Uji Kelayakan Analisis Faktor (KMO)	44
4.4	Uji Normalitas Multivariat	44
4.5	Uji Multikolinearitas	45
4.6	Perhitungan Algoritma <i>K-Means++</i>	46
4.6.1	Normalisasi Data	46
4.6.2	<i>K-Means++</i>	48
4.6.3	Evaluasi <i>Silhouette Coefficient</i>	53
4.7	Hasil Pembentukan Klaster	56
4.8	Integrasi Keilmuan	58
BAB V	PENUTUP	61
5.1	Kesimpulan	61
5.2	Saran	61

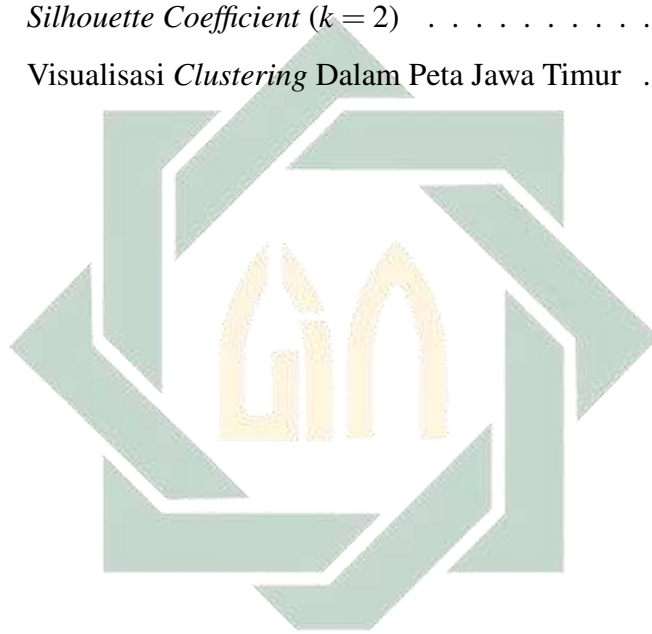
DAFTAR PUSTAKA	63
LAMPIRAN	71



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR GAMBAR

Gambar 3.1	Diagram Alir Kmeans++	41
Gambar 4.1	Deteksi Outlier Menggunakan Boxplot	43
Gambar 4.2	<i>Silhouette Coefficient</i> ($k = 2$)	54
Gambar 4.3	Visualisasi <i>Clustering</i> Dalam Peta Jawa Timur	57



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR TABEL

Tabel 2.1	Interpretasi Nilai KMO	24
Tabel 2.2	Interpretasi Nilai VIF	26
Tabel 2.3	Interpretasi Nilai <i>Silhouette Coefficient</i>	33
Tabel 3.1	Data Indikator Pembangunan Manusia 2023	38
Tabel 3.2	Variabel Data Penelitian Indikator Pembangunan Manusia	39
Tabel 4.1	Statistik Deskriptif	42
Tabel 4.2	Ringkasan Uji KMO	44
Tabel 4.3	Ringkasan Uji Normalitas Multivariat	44
Tabel 4.4	Hasil Uji Multikolinearitas (VIF)	45
Tabel 4.5	Hasil Normalisasi Data	47
Tabel 4.6	Centroid pertama (data ternormalisasi)	48
Tabel 4.7	Nilai Normalisasi Pacitan dan Centroid 1 (Kota Kediri)	48
Tabel 4.1	Probabilitas Pemilihan Centroid Kedua, Jarak terhadap K1, dan Probabilitas Kumulatif	50
Tabel 4.10	Centroid kedua (data ternormalisasi)	51
Tabel 4.11	Centroid setelah denormalisasi	53
Tabel 4.12	Nilai <i>Silhouette Coefficient</i> untuk Berbagai Jumlah Klaster	54
Tabel 4.13	Centroid Tiap Klaster (Denormalisasi)	55
Tabel 4.14	Distribusi Anggota Tiap Klaster, Jumlah, dan Persentasenya	56

ABSTRAK

PENGELOMPOKAN INDIKATOR PEMBANGUNAN MANUSIA KABUPATEN/KOTA DI JAWA TIMUR MENGGUNAKAN K-MEANS++

Ketimpangan capaian pembangunan manusia di Provinsi Jawa Timur menunjukkan adanya perbedaan yang signifikan antar wilayah, baik dari sisi pendidikan, kesehatan, maupun ekonomi. Kesenjangan ini menjadi tantangan dalam mewujudkan pembangunan yang merata dan berkeadilan sosial di tingkat daerah. Penelitian ini bertujuan untuk: (1) mengetahui hasil uji kelayakan analisis faktor, deteksi outlier, dan uji multikolinearitas pada indikator pembangunan manusia; (2) mengevaluasi kualitas hasil pengelompokan menggunakan algoritma K-Means++ untuk menentukan jumlah klaster yang optimal; serta (3) menginterpretasikan hasil klasterisasi guna memperoleh pemetaan wilayah berdasarkan kemiripan karakteristik pembangunan manusia. Data yang digunakan merupakan data tahun 2023 yang bersumber dari Badan Pusat Statistik. Variabel yang dianalisis mencakup dimensi kesehatan, pendidikan, dan ekonomi, antara lain: Umur Harapan Hidup (UHH), persentase wanita usia 15–49 tahun melahirkan dengan bantuan tenaga medis, persentase penduduk dengan keluhan kesehatan dan rawat jalan, kepemilikan BPJS PBI, Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS), Angka Melek Huruf (AMH), Indeks Pembangunan Literasi Masyarakat (IPLM), Indeks Daya Beli (IDB), dan Pengeluaran Per Kapita. Teknik yang digunakan adalah algoritma *K-Means++* yang mampu mengelompokkan wilayah-wilayah ke dalam klaster berdasarkan kemiripan karakteristik. Evaluasi struktur klaster dilakukan menggunakan metode *Silhouette Coefficient*, dan hasilnya menunjukkan bahwa pemilihan jumlah klaster terbaik adalah $k = 2$, dengan nilai *Silhouette Coefficient* sebesar 0,1997. Klaster pertama terdiri dari 13 daerah yang cenderung berlokasi di kota-kota besar dengan capaian pembangunan relatif tinggi, sedangkan klaster kedua mencakup 25 daerah yang terbesar di kawasan tapal kudan dan Pulau Madura dengan capaian yang lebih rendah. Hasil penelitian ini diharapkan dapat menjadi landasan dalam penyusunan strategi kebijakan pembangunan daerah yang lebih tepat sasaran, berbasis data, dan mampu mendorong pemerataan pembangunan secara berkelanjutan di Provinsi Jawa Timur.

Kata Kunci: Jawa Timur, *K-Means++*, Ketimpangan Wilayah, Klasterisasi, Pembangunan Manusia

ABSTRACT

GROUPING OF HUMAN DEVELOPMENT INDICATORS OF DISTRICT/CITY IN EAST JAVA USING K-MEANS++

Disparities in human development achievements in East Java Province highlight significant differences between regions in terms of education, health, and economy. These gaps present challenges in achieving equitable and socially just regional development. This study aims to: (1) examine the results of factor analysis feasibility tests, outlier detection, and multicollinearity tests on human development indicators; (2) evaluate the quality of clustering results using the *K-Means++* algorithm to determine the optimal number of clusters; and (3) interpret the clustering results to obtain a regional mapping based on the similarity of human development characteristics. The data used are from 2023 and sourced from Statistics Indonesia (BPS). The variables analyzed cover dimensions of health, education, and economy, including: Life expectancy (UHH), percentage of women aged 15 to 49 years of age giving birth assisted by medical personnel, percentage of the population with health complaints and outpatient services, ownership of BPJS PBI, expected years of schooling (HLS), mean years of schooling (RLS), literacy rate (AMH), community literacy development index (IPLM), purchasing power index (IDB), and expenditure per capita. The method applied is the *K-Means++* algorithm, which groups regions into clusters based on similarity in characteristics. The cluster structure was evaluated using the *Silhouette Coefficient* method, which showed that the optimal number of clusters is $k = 2$, with a *Silhouette Coefficient* value of 0.1997. The first group consists of 13 regions mostly located in major urban areas with relatively high development achievements, while the second cluster includes 25 regions primarily located in the Tapal Kuda area and Madura Island with lower development results. The findings of this study are expected to serve as a basis for designing more targeted and data-driven regional development policies that promote equitable and sustainable development across East Java Province.

Keywords: Clustering, East Java, Human Development, *K-Means++*, Regional Disparity

DAFTAR PUSTAKA

- Abdillah, M., Siregar, F., & Lubis, A. (2024). Analisis determinan kemiskinan di kabupaten/kota provinsi jawa timur. *Ekuilnomi: Jurnal Ekonomi dan Pembangunan*, 12(1), 45–56. <https://jurnal.usi.ac.id/index.php/ekuilnomi/article/download/237/988/2786>
- Adolph, R. (2016). *Renstra bappeda 2021–2026* (tech. rep.). Badan Perencanaan Pembangunan Daerah. Indonesia. <http://contoh-url-bappeda.go.id/renstra2021-2026.pdf>
- Al-Ghazali, A. H. (2005). *Al-mustashfa min 'ilm al-usul*. Dar al-Kutub al-'Ilmiyyah.
- Allorerung, P. P., Erna, A., Bagussahrir, M., & Alam, S. (2024). Analisis performa normalisasi data untuk klasifikasi k-nearest neighbor pada dataset penyakit. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 9(3), 178–191. <https://doi.org/10.14421/jiska.2024.9.3.178-191>
- Al-Qarafi, S. (2001). *Al-furuq*. Dar al-Kutub al-Ilmiyyah.
- Amalia, L., Darmawan, A., & Nurjannah, R. (2023). Pengaruh ketimpangan pendapatan terhadap kemiskinan di provinsi jawa timur. *Buletin Ekonomika Pembangunan*, 21(2), 101–110. <https://journal.trunojoyo.ac.id/bep/article/view/25948>
- Ambarwari, A., Adrian, Q. J., Herdiyeni, Y., dkk. (2020). Analysis of the effect of data scaling on the performance of the machine learning algorithm for plant identification [<https://doi.org/10.29207/resti.v4i1.1517>]. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4(1), 117–122.
- Anjani, M., Kurniawan, A., & Sari, D. (2021). Literacy and human capital development in indonesia: A micro data analysis. *Jurnal Pendidikan dan Pembangunan*, 11(2), 87–97. <https://doi.org/10.21009/jpp.112.07>
- Anwar, M. (2021). Pendekatan pembangunan manusia dalam perspektif ekonomi kesejahteraan. *Jurnal Ekonomi dan Pembangunan*, 29(1), 57–70.

- Arthur, D., & Vassilvitskii, S. (2006). *K-means++: The advantages of careful seeding* (tech. rep.). Stanford. <http://theory.stanford.edu/~sergei/papers/kMeansPP-soda.pdf>
- Arthur, D., & Vassilvitskii, S. (2007). K-means++: The advantages of careful seeding. *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, 1027–1035. <https://doi.org/10.1137/1.9781611972771.85>
- Ashari, I. F., Nugroho, E. D., Baraku, R., Yanda, I. N., & Liwardana, R. (2023). Analysis of elbow, silhouette, davies-bouldin, calinski-harabasz, and rand-index evaluation on k-means algorithm for classifying flood-affected areas in jakarta. *Journal of Applied Informatics and Computing*, 7(1), 45–54. <https://doi.org/10.30871/jaic.v7i1.4947>
- Bahmani, B., Moseley, B., Vattani, A., Kumar, R., & Vassilvitskii, S. (2012). Scalable k-means++. *arXiv preprint arXiv:1203.6402*.
- Bappenas. (2022). Strategi peningkatan indeks pembangunan manusia di indonesia [https://bappenas.go.id/berita/bappenas-sasaran-pembangunan-2023-untuk-transformasi-ekonomi-dan-bonus-demografi-5dSVW?utm_source=chatgpt.com].
- (BPS), B. P. S. (2023). *Indeks pembangunan manusia jawa timur 2023*. BPS, Jakarta.
- Dinas Kesehatan Provinsi Jawa Timur. (2024). Profil kesehatan provinsi jawa timur tahun 2023 [Diakses pada Mei 2025].
- Farida, Y., Firdaus, G. S., Wibowo, A. T., Sari, S. K., & Desinaini, L. N. (2022). Evaluation of food security area of east java province using fuzzy c-means (fcm) and technique for order preference by similarity to ideal solution (topsis). *Indonesian Journal of Computing and Cybernetics Systems*, 16(4), 303–320. <https://doi.org/10.22146/ijccs.82297>
- Han, J., Pei, J., & Kamber, M. (2022). *Data mining: Concepts and techniques* (4th). Morgan Kaufmann.

- Hartama, D., & Anjelita, M. (2022). Adnalysis of silhouette coefficient evaluation with euclidean distance in the clustering method (case study: Number of public schools in indonesia). *Jurnal Mantik*, 6(3), 3667–3677. <https://doi.org/10.35335/mantik.v6i3.3318>
- Hasanah, U., & Ahmadi, H. (2017). Pengaruh ketimpangan pendapatan, pendapatan per kapita, dan pengeluaran pemerintah di bidang kesehatan terhadap sektor kesehatan di indonesia. *Jurnal Ilmu Ekonomi Terapan*, 2(1), 31–48.
- Hepi, H., & Zakiah, W. (2018). The influence of life expectancy rate and old school rate to grdp percapita and economic growth in central kalimantan province 2011-2015. *Journal Magister Ilmu Ekonomi Universtas Palangka Raya: GROWTH*, 4(1), 56–68.
- Hermawan, H., & Hasugian, H. (2022). Penerapan data mining untuk clustering indeks pembangunan manusia berdasarkan provinsi di indonesia. *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, 1(1), 525–532. <https://senafti.budiluhur.ac.id/senafti/article/view/320>
- Hidayati, R., Fauzi, A., & Muthmainnah. (2021). Analisis perbandingan metode k-means menggunakan jarak euclidean, manhattan, dan minkowski pada data penerima bantuan sosial. *Jurnal Riset Informatika*, 3(2), 89–96. <https://doi.org/10.29313/jri.v3i2.7546>
- Jain, A. K. (2010). Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8), 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- JawaPos. (2024). *Indeks pembangunan manusia menurut kabupaten/kota di jawa timur tahun 2021–2024* [Diakses pada 23 Mei 2025]. Jawa Pos. <https://manadopost.jawapos.com/mpedia/286015724/>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36.
- Kementerian Kesehatan Republik Indonesia. (2021). *Profil kesehatan indonesia tahun 2020* (Diakses pada Mei 2025). Pusat Data dan Informasi, Kementerian Kesehatan RI. <https://pusdatin.kemkes.go.id/folder/view/01/structure-publikasi-pusat-data-dan-informasi.html>

- Kusnaldi, M. R., Gulo, T., & Aripin, S. (2022). Penerapan normalisasi data dalam mengelompokkan data mahasiswa dengan menggunakan metode k-means untuk menentukan prioritas bantuan uang kuliah tunggal [<https://doi.org/10.47065/josyc.v3i4.2112>]. *Journal of Computer System and Informatics (JoSYC)*, 3(4), 330–338.
- Kuzey, C., dkk. (2021). An investigation of the factors affecting erp implementation success. *Journal of Enterprise Information Management*.
- Lestari, P. A. P., Roesdiyanto, R., & Ulfah, N. H. (2020). Kebutuhan kesehatan dan akses pelayanan kesehatan dengan kepesertaan badan penyelenggara jaminan sosial (bpjs) kesehatan. *Jurnal Penelitian dan Pengembangan Kesehatan Masyarakat Indonesia*, 1(2), 138–156.
- Lina, I. R., & Wati, D. C. (2023). Klasifikasi pengeluaran per kapita di tiga provinsi sulawesi menggunakan k-nearest neighbor [<https://doi.org/10.36456/jstat.vol16.no1.a7193>]. *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, 16(1), 395–406.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519–530.
- Mauludi, A., Nur, F., dkk. (2023). Pengaruh pengaruh upah minimum, pertumbuhan ekonomi, indeks pembangunan manusia, dan belanja modal terhadap kemiskinan kabupaten/kota di jawa timur tahun 2017 s/d 2021. *IQTISHADIA Jurnal Ekonomi & Perbankan Syariah*, 10(1), 68–83.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis*. John Wiley & Sons.
- Mulyani, H., Setiawan, R. A., & Fathi, H. (2023). Optimization of k value in clustering using silhouette score (case study: Mall customers data). *Journal of Information Technology and Its Utilization*, 6(2), 123–130. <https://doi.org/10.56873/jitu.6.2.5243>
- Nizhamul, D. L., & Vedelya Istighfarah, N. D. D. (2023). Di hongkong dan singapura the influence of factors on the human development index (hdi). *Formosa Journal of Sustainable Research*, 99–106.

- Nugroho, N., & Adhinata, F. D. (2022). Penggunaan metode k-means dan k-means++ sebagai clustering data covid-19 di pulau jawa [<https://doi.org/10.34148/teknika.v11i3.502>]. *Teknika*, 11(3), 170–179.
- Oti, E. U., Olusola, M. O., Eze, F. C., & Enogwe, S. U. (2021). Comprehensive review of k-means clustering algorithms [<https://doi.org/10.31695/IJASRE.2021.34050>]. *criterion*, 12, 22–23.
- Pranoto, Y. (2021). *Analisis statistik multivariat dengan spss*. Deepublish.
- Putri, A. L., & Nugroho, D. (2023). Literacy development and non-formal education participation in indonesian provinces: A spatial analysis. *Asian Journal of Education and Social Studies*, 38(1), 24–33. <https://doi.org/10.9734/ajess/2023/v38i1855>
- Raga, P., Aksa, S. K., Herdian, T., Samosir, F., & Ponto, K. N. A. (2021). Pengukuran indeks daya beli jasa transportasi laut dan penyeberangan pada wilayah tertinggal. *Jurnal Sistem Transportasi & Logistik*, 1(1). <https://journal.itltrisakti.ac.id/index.php/jstl/article/view/627>
- Rahmatullah, A., & Putra, Y. (2023). Spatial disparities in human development across indonesian regions: A multilevel analysis. *Asian Development Review*, 40(2), 137–158. https://doi.org/10.1162/adev_a.00235
- Rahmawati, I., & Prabowo, A. (2022). Community-based education intervention and its impact on multidimensional poverty reduction in rural indonesia. *Journal of Human Development and Capabilities*, 23(4), 567–585. <https://doi.org/10.1080/19452829.2022.2048773>
- Rohali, A. (2025). Strategi pemerataan indeks pembangunan manusia di provinsi jawa timur: Kajian regional. *Prosiding Seminar Nasional Sains dan Teknologi Terapan*. <https://conference.ut.ac.id/index.php/saintek/article/download/5001/1894>
- Safitri, A., & Prakoso, B. (2022). Creative economy and its role in improving human development index in indonesian urban areas. *Journal of Regional Economic Studies*, 14(1), 45–58. <https://doi.org/10.52391/jres.v14i1.248>

- Salmerón-Gómez, R., García, C., & García, J. (2024). A redefined variance inflation factor: Overcoming the limitations of the variance inflation factor. *arXiv preprint arXiv:2401.12345*. <https://arxiv.org/abs/2401.12345>
- Sari, D. P., Nugroho, R. A., & Putri, M. D. (2021). Determinants of skilled birth attendance in indonesia: Evidence from the 2017 indonesia demographic and health survey. *BMC Public Health*, 21(1), 1–9. <https://doi.org/10.1186/s12889-021-11234-0>
- Sari, F. Y., Kuntari, M. S., Yati, W. A., Khaulasari, H., & Hafiyusholeh, M. (2025). Implementasi k-means clustering melalui pemanfaatan sampling kombinasi pada pengelompokan pola kesehatan mental mahasiswa sains dan teknologi. *Jurnal Nasional Teknologi dan Sistem Informasi*, 11(1), 9–16. <https://doi.org/10.25077/TEKNOSI.v11i1.2025.9-16>
- Sari, T. D. (2022). Pembangunan manusia berbasis keadilan sosial: Evaluasi indeks pembangunan manusia di indonesia. *Jurnal Sosial Humaniora*, 14(2), 120–133.
- Selatan, B. P. S. (K. H. (2015). *Indeks pembangunan manusia kabupaten halmahera selatan 2015*. BPS Kabupaten Halmahera Selatan, Labuha. <https://halmaheraselatankab.bps.go.id/id/publication/2016/10/07/22b3ae8b9d4e30bfaa4fdde0/indeks-pembangunan-manusia-kabupaten-halmahera-selatan-2015.html>
- Setiyawati, N., Bangkalang, D. H., & Purnomo, H. D. (2023). Comparison of k-means & k-means++ clustering models using singular value decomposition (svd) in menu engineering. *JOIV: International Journal on Informatics Visualization*, 7(3), 871–877. <https://doi.org/10.30630/joiv.7.3.1053>
- Shoimah, F. (2023). Mengukur indeks pembangunan literasi masyarakat (iplm) kabupaten probolinggo. *Pangripta: Jurnal Ilmiah Kajian Perencanaan Pembangunan*, 6(1), 1164–1175. <https://doi.org/10.58411/7d4rax37>
- Sinaga, T., & Damanik, D. (2024). Pentingnya melek huruf dalam pembangunan sosial ekonomi. *BPS Indonesia*.

- Sofia, A. D. (2023). Pengelompokan kabupaten/kota berdasarkan indikator indeks pembangunan manusia 2022 menggunakan k-harmonic means clustering. *Jurnal Riset Statistika*, 163–172. <https://doi.org/10.29313/jrs.v3i2.3130>
- Statistik, B. P. (2023). *Indikator pendidikan indonesia 2023*. <https://www.bps.go.id>
- Sudarno, R., & Lestari, W. (2022). The impact of basic education indicators on human development index in indonesia. *Jurnal Ekonomi dan Pendidikan*, 19(2), 58–68. <https://doi.org/10.21831/jep.v19i2.47563>
- Supandi, A., Saefuddin, A., & Sulvianti, I. D. (2020). Two step cluster application to classify villages in kabupaten madiun based on village potential data. *Xplore: Journal of Statistics*, 10(1), 12–26. <https://doi.org/10.29244/xplore.v10i1.272>
- Tarigan, P. M. S., Hardinata, J. T., Qurniawan, H., Safii, M., & Winanjaya, R. (2022). Implementasi data mining menggunakan algoritma apriori dalam menentukan persediaan barang: Studi kasus: Toko sinar harahap. *Jurnal Janitra Informatika dan Sistem Informasi*, 2(1), 9–19. <https://doi.org/10.25008/janitra.v2i1.142>
- Tirtana, D., Sepriyatno, Firmansyah, M. S., & Sopiandi, W. (2023). Perbandingan pengaruh indeks pembangunan manusia pada provinsi jawa tengah dan jawa timur. *JEMeS: Jurnal Ekonomi Manajemen dan Sosial*, 6(2), 22–31. <https://ojs.ejournalunigoro.com/index.php/JEMeS>
- Triola, M. F. (2014). *Elementary statistics* (12th). Pearson Education.
- Ulinnuha, N., & Sholihah, S. A. (2021). Analisis cluster untuk pemetaan data kasus covid-19 di indonesia menggunakan k-means. *Jurnal MSA (Matematika dan Statistika serta Aplikasinya)*, 9(2), 27–31. <https://doi.org/10.24252/msa.v9i2.19478>
- UNDP. (2023). Human development report 2023 [https://hdr.undp.org/data-center/human-development-index?utm_source=chatgpt.com].
- Wardhani, L., & Rakhmadi, R. (2021). Real per capita expenditure and subjective well-being in indonesia: Evidence from national socioeconomic survey.

Indonesian Journal of Development Planning, 5(3), 231–245. <https://doi.org/10.36574/jpp.v5i3.183>

- Wenny. (2024). Normalisasi data kependudukan dengan model min max dan algoritma k-means untuk pengelompokkan tingkat ekonomi masyarakat. *Bulletin of Information System Research*, 2(2), 53–63. <https://doi.org/10.62866/bios.v2i2.141>
- Widodo, A. (2021). Analisis spasial ketimpangan indeks pembangunan manusia di pulau jawa. *Jurnal Ekonomi dan Pembangunan Indonesia*, 22(1), 45–58.
- Widodo, A., & Sari, D. (2022). Strengthening primary health care to reduce maternal and infant mortality in remote areas: A community-based approach. *Indonesian Journal of Public Health*, 17(2), 101–112. <https://e-journal.unair.ac.id/IJPH/article/view/2024>
- Wijayanti, M., Sari, N., & Prabowo, A. (2021). Education, literacy, and human development: Regional analysis in java. *Jurnal Ekonomi dan Kebijakan Publik*, 12(3), 143–154. <https://doi.org/10.22212/jekp.v12i3.948>
- Wulandari, L., Farida, Y., Fanani, A., Ulinnuha, N., & Intan, P. K. (2020). Evaluation of disadvantaged regions in east java based-on the 33 indicators of the ministry of villages, development of disadvantaged regions, and transmigration using the ensemble rock (robust clustering using link) method. *Advances in Science, Technology and Engineering Systems Journal*, 5(5), 193–200. <https://doi.org/10.25046/aj050524>
- Yoliadi, D. N. (2023). Data mining dalam analisis tingkat penjualan barang elektronik menggunakan algoritma k-means. *Insearch Information System Research Journal*, 3(01).
- Yosalli, Y., Rusmitasari, H., Suyitno, S., Sariana, S., & Maretalinia, M. (2023). Factors related to health insurance bpjs pbi ownership among household members. *Jurnal EduHealth*, 14(2), 1015–1019.