

**PEMODELAN INDEKS PEMBANGUNAN MANUSIA  
BERDASARKAN PENGELOMPOKAN TINGKAT  
KESEJAHTERAAN WILAYAH DI PULAU JAWA  
MENGUNAKAN METODE HYBRID K-MEANS++ DAN MARS**

**SKRIPSI**



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

Disusun Oleh  
**HENDRIK SUBAEKTI**  
**09020222031**

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

**2026**

## PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : Hendrik Subaekti  
NIM : 09020222031  
Program Studi : Matematika  
Angkatan : 2022

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "Pemodelan Indeks Pembangunan Manusia Berdasarkan Pengelompokan Tingkat Kesejahteraan Wilayah di Pulau Jawa Menggunakan Metode Hybrid K-Means++ dan MARS". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 15 Juni 2026

Yang menyatakan,



Hendrik Subaekti  
NIM. 09020222031

# LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

Nama : Hendrik Subaekti  
NIM : 09020222031  
Judul Skripsi : **Pemodelan Indeks Pembangunan Manusia Berdasarkan Pengelompokan Tingkat Kesejahteraan Wilayah di Pulau Jawa Menggunakan Metode Hybrid K-Means++ dan MARS**

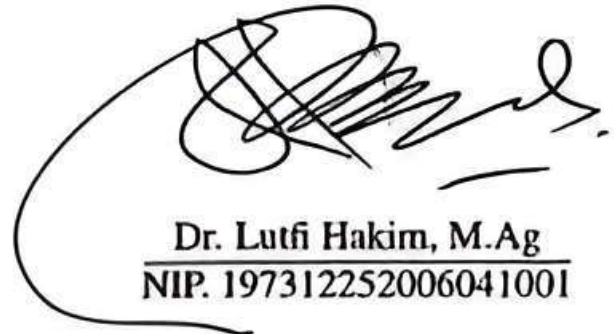
Telah diperiksa dan disetujui untuk diujikan.

Pembimbing I



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

Pembimbing II



Dr. Lutfi Hakim, M.Ag  
NIP. 197312252006041001

Mengetahui,  
Ketua Program Studi Matematika  
UIN Sunan Ampel Surabaya



Dr. Yuniar Farida, M.T  
NIP. 197905272014032002

## PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh:

Nama : Hendrik Subackti  
NIM : 09020222031  
Judul Skripsi : **Pemodelan Indeks Pembangunan Manusia Berdasarkan Pengelompokan Tingkat Kesejahteraan Wilayah di Pulau Jawa Menggunakan Metode Hybrid K-Means++ dan MARS**

Telah dipertahankan di depan Tim Penguji pada tanggal 15 Juni 2026.

**Mengesahkan,  
Tim Penguji**

**Penguji I**



Dr. Dian Candra Rini Novitasari, M.Kom  
NIP. 198511242014032001

**Penguji II**



Dr. Ahmad Hanif Asyhar, M.Si  
NIP. 198601232014031001

**Penguji III**



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

**Penguji IV**



Dr. Lutfi Hakim, M.Ag  
NIP. 197312252006041001

**Mengetahui,**

Dekan Fakultas Sains dan Teknologi  
UIN Sunan Ampel Surabaya



Dr. A. Saopul 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: perpustakaan@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 : Hendrik Sabakti  
NIM : 09020222031  
Fakultas/Jurusan : FST / Matematika  
E-mail address : HendrikSabaktiP2@gmail.com.

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

Skripsi  Tesis  Disertasi  Lain-lain (.....)

yang berjudul :

Pemodelan Indeks Pembangunan manusia Berdasarkan  
Pengelompokan Tingkat Kesejahteraan Wilayah di paku Jawa  
menggunakan metode hybrid K-Means tt dan Max.

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 02 Juni 2026.

nama terang dan tanda tangan  
Hendrik Sabakti.

## ABSTRAK

### **Pemodelan Indeks Pembangunan Manusia Berdasarkan Pengelompokan Tingkat Kesejahteraan Wilayah di Pulau Jawa Menggunakan Metode Hybrid K-Means++ dan MARS**

Indeks Pembangunan Manusia (IPM) merupakan indikator penting untuk menilai kesejahteraan wilayah, namun disparitas antar daerah di Pulau Jawa masih menjadi tantangan dalam perumusan kebijakan yang tepat sasaran. Penelitian ini bertujuan mengelompokkan kabupaten/kota di Pulau Jawa berdasarkan karakteristik kesejahteraan, memodelkan faktor-faktor yang memengaruhi IPM pada setiap kelompok, serta mengevaluasi akurasi pendekatan *hybrid* K-Means++ dan MARS. Data bersumber dari Badan Pusat Statistik tahun 2024, mencakup 119 kabupaten/kota dengan sepuluh variabel prediktor. Metode yang digunakan adalah K-Means++ dengan reduksi PCA sebanyak 3 komponen (total variansi 60%), dilanjutkan pemodelan MARS pada masing-masing klaster. Hasil penelitian menunjukkan terbentuk dua klaster, yaitu Klaster 1 (kesejahteraan rendah–menengah) dengan 64 wilayah yang didominasi daerah pedesaan dan agraris, serta Klaster 2 (kesejahteraan tinggi) dengan 55 wilayah yang didominasi daerah perkotaan dan industri seperti Jabodetabek, Surabaya Kota, dan Bandung Kota, dengan nilai *silhouette score* sebesar 0.337 yang termasuk dalam kategori cukup baik. Model MARS pada Klaster 1 menghasilkan nilai  $R^2$  sebesar 82.85% dengan variabel dominan Jumlah Sanitasi Layak ( $X_3$ ), Persentase Penduduk Miskin ( $X_7$ ), dan Tingkat Kepadatan Penduduk ( $X_8$ ), sedangkan pada Klaster 2 diperoleh nilai  $R^2$  sebesar 86.34% dengan variabel dominan Jumlah Sanitasi Layak ( $X_3$ ), Persentase Penduduk Miskin ( $X_7$ ), serta Jumlah Fasilitas Kesehatan ( $X_2$ ) dan Indeks Literasi Minimum ( $X_9$ ). Pendekatan *hybrid* K-Means++ MARS terbukti lebih unggul dalam akurasi prediksi dengan nilai RMSE rata-rata sebesar 1.8282 dibandingkan MARS sebesar 2.7530, sehingga lebih direkomendasikan untuk perumusan kebijakan pembangunan manusia yang tepat sasaran per wilayah, sedangkan MARS lebih sesuai untuk pemahaman faktor secara global pada tingkat provinsi atau regional.

**Kata kunci:** Indeks Pembangunan Manusia, *K-Means++*, Kesejahteraan Wilayah, Klasterisasi, *MARS*.

## ABSTRACT

### Human Development Index Modeling Based on Regional Welfare Level Grouping in Java Island Using the Hybrid K-Means++ and MARS Methods

Human Development Index (HDI) is an important indicator used to measure regional welfare; however, disparities among regions in Java Island remain a challenge in formulating appropriate development policies. This study aims to classify regencies/cities in Java based on welfare characteristics, model the factors affecting HDI in each group, and evaluate the accuracy of the hybrid K-Means++ and Multivariate Adaptive Regression Splines (MARS) approach. The data were obtained from the Central Bureau of Statistics (BPS) in 2024, covering 119 regencies/cities with ten predictor variables. The method applied was K-Means++ clustering with Principal Component Analysis (PCA) reduction into 3 components (total variance of 60%), followed by MARS modeling for each cluster. The results showed that two clusters were formed: Cluster 1 (low–middle welfare) consisting of 64 regions dominated by rural and agricultural areas, and Cluster 2 (high welfare) consisting of 55 regions dominated by urban and industrial areas such as Greater Jakarta, Surabaya City, and Bandung City, with a silhouette score of 0.337 categorized as fairly good. The MARS model in Cluster 1 produced an  $R^2$  value of 82.85% with dominant variables including Proper Sanitation ( $X_3$ ), Percentage of Poor Population ( $X_7$ ), and Population Density Level ( $X_8$ ). Meanwhile, Cluster 2 obtained an  $R^2$  value of 86.34% with dominant variables including Proper Sanitation ( $X_3$ ), Percentage of Poor Population ( $X_7$ ), Health Facilities ( $X_2$ ), and Minimum Literacy Index ( $X_9$ ). The hybrid K-Means++ MARS approach proved to provide better predictive accuracy with an average Root Mean Square Error (RMSE) value of 1.8282 compared to MARS of 2.7530. Therefore, the hybrid approach is more recommended for region-specific human development policy formulation, while MARS alone is more suitable for understanding global factor patterns at provincial or regional levels.

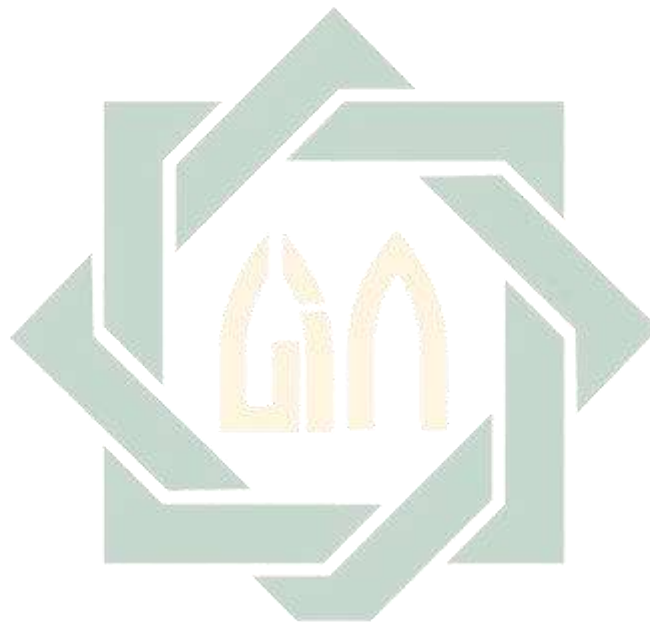
**Keywords:** Human Development Index, K-Means++, Regional Welfare, Clustering, MARS.

# DAFTAR ISI

<b>HALAMAN JUDUL</b>	<b>i</b>
<b>LEMBAR PERSETUJUAN PEMBIMBING</b>	<b>ii</b>
<b>PENGESAHAN TIM PENGUJI SKRIPSI</b>	<b>iii</b>
<b>HALAMAN PERNYATAAN KEASLIAN</b>	<b>iv</b>
<b>MOTTO</b>	<b>v</b>
<b>PERSEMBAHAN</b>	<b>vi</b>
<b>KATA PENGANTAR</b>	<b>vii</b>
<b>DAFTAR ISI</b>	<b>ix</b>
<b>DAFTAR TABEL</b>	<b>xiii</b>
<b>DAFTAR GAMBAR</b>	<b>xv</b>
<b>ABSTRAK</b>	<b>xvii</b>
<b>ABSTRACT</b>	<b>xviii</b>
<b>I PENDAHULUAN</b>	<b>1</b>
1.1. Latar Belakang Masalah . . . . .	1
1.2. Rumusan Masalah . . . . .	11
1.3. Tujuan Penelitian . . . . .	11
1.4. Manfaat Penelitian . . . . .	11
1.5. Batasan Masalah . . . . .	12
1.6. Sistematika Penulisan . . . . .	13
<b>II TINJAUAN PUSTAKA</b>	<b>15</b>
2.1. Kesejahteraan Masyarakat di Pulau Jawa . . . . .	15
2.2. Indeks Pembangunan Manusia (IPM) . . . . .	16
2.3. Faktor Pembentuk Klasterisasi Kesejahteraan dan Penentu IPM . . . . .	16
2.4. Preprocessing Data . . . . .	20
2.4.1. Pengecekan Data Kosong ( <i>Missing Value</i> ) . . . . .	21
2.4.2. Deteksi <i>Outlier</i> dengan Metode IQR . . . . .	21

2.4.3. Penanganan <i>Outlier</i> dengan Imputasi Median . . . . .	22
2.5. Normalisasi Data dengan <i>Z-Score</i> . . . . .	23
2.6. Uji Normalitas Multivariat Mardia . . . . .	24
2.7. Uji <i>Kaiser-Meyer-Olkin</i> (KMO) . . . . .	26
2.8. Uji Bartlett . . . . .	27
2.9. Uji Multikolinearitas . . . . .	28
2.10. Reduksi Dimensi dengan <i>Principal Component Analysis</i> (PCA) . . . . .	29
2.11. Metode <i>Elbow</i> . . . . .	32
2.12. <i>K-Means++ Clustering</i> . . . . .	33
2.13. <i>Silhouette Coefficient</i> . . . . .	35
2.14. Regresi Non Parametrik . . . . .	38
2.15. Multivariate Adaptive Regression Splines (MARS) . . . . .	38
2.16. Pemilihan Model MARS Terbaik . . . . .	43
2.17. Integrasi Keislaman . . . . .	44
<b>III METODE PENELITIAN</b>	<b>49</b>
3.1. Jenis Penelitian . . . . .	49
3.2. Sumber Data . . . . .	49
3.3. Variabel Penelitian . . . . .	50
3.4. Struktur Data Penelitian . . . . .	50
3.5. Metode Analisis Data . . . . .	51
3.6. Tahapan Penelitian . . . . .	52
3.6.1. Input Dataset . . . . .	52
3.6.2. <i>Pre-Processing</i> & Uji Asumsi Data . . . . .	53
3.6.3. Reduksi Dimensi PCA . . . . .	53
3.6.4. Tahapan <i>K-Means++ Clustering</i> . . . . .	55
3.6.5. Tahapan <i>Multivariate Adaptive Regression Spline</i> (MARS) . . . . .	59
3.7. Skema Uji Coba <i>Tuning</i> Parameter . . . . .	63
<b>IV HASIL DAN PEMBAHASAN</b>	<b>67</b>
<b>V PENUTUP</b>	<b>147</b>
5.1. Kesimpulan . . . . .	147
5.2. Saran . . . . .	148
<b>DAFTAR PUSTAKA</b>	<b>151</b>

<b>LAMPIRAN</b>	<b>163</b>
<b>Kode Program Python Klasterisasi K-Means++</b>	<b>177</b>
<b>Kode Program R Pemodelan K-MEANS++ MARS</b>	<b>196</b>



UIN SUNAN AMPEL  
S U R A B A Y A

# DAFTAR TABEL

2.1	Kriteria Nilai KMO . . . . .	27
2.2	Kriteria Nilai <i>Variance Inflation Factor</i> (VIF) . . . . .	29
2.3	Kriteria Nilai <i>Silhouette Coefficient</i> . . . . .	37
3.1	Data Penelitian . . . . .	49
3.2	Variabel Penelitian . . . . .	50
3.3	Struktur Data <i>Clustering</i> . . . . .	50
3.4	Struktur Data Pemodelan MARS . . . . .	51
3.5	Rentang Parameter dalam Tuning K-Means++ . . . . .	63
3.6	Rentang Parameter dalam Tuning MARS . . . . .	65
4.1	Dataset Penelitian . . . . .	67
4.2	Jumlah Missing Value pada Setiap Variabel . . . . .	68
4.3	Hasil Pemeriksaan <i>Outlier</i> . . . . .	69
4.4	Hasil Imputasi Median . . . . .	70
4.5	Hasil Uji Normalisasi Data . . . . .	71
4.6	Perhitungan Normalisasi Variabel $X_1$ . . . . .	71
4.7	Hasil Jarak Mahalanobis . . . . .	72
4.8	Hasil Uji Skewness & Kurtosis Multivariat . . . . .	75
4.9	Contoh Perhitungan $M_{1j}^3$ untuk Kabupaten Pacitan . . . . .	79
4.10	Perhitungan Kurtosis Multivariat . . . . .	80
4.11	Hasil Uji KMO . . . . .	81
4.12	Matriks Korelasi Antar Variabel . . . . .	82
4.13	Perhitungan Nilai $\sum r_{6j}^2$ pada Variabel $X_6$ . . . . .	83
4.14	Perhitungan Korelasi Parsial dan Kuadrat Korelasi Parsial untuk Variabel $X_6$ . . . . .	84
4.15	Hasil Uji Barlett . . . . .	85
4.16	Hasil Uji Multikolinearitas . . . . .	88

4.17 Hasil Uji PCA . . . . .	90
4.18 Hasil Loading Matrix . . . . .	93
4.19 Hasil Skor Komponen Utama . . . . .	95
4.20 Hasil Uji <i>Elbow Method</i> . . . . .	97
4.21 Hasil Inisialisasi <i>K-Means++</i> . . . . .	99
4.22 Hasil Pelacakan <i>WCSS</i> Terkecil . . . . .	100
4.23 Hasil Iterasi Konvergen <i>K-Means++</i> . . . . .	101
4.24 Hasil Perhitungan Jarak terhadap Centroid . . . . .	103
4.25 Nilai Centroid Akhir Tiap Klaster . . . . .	104
4.26 Hasil Pembentukan Klaster . . . . .	110
4.27 Waktu Komputasi <i>K-Means++</i> . . . . .	112
4.28 Distribusi Data per Klaster . . . . .	113
4.29 Hasil Uji Nonlinearitas Klaster 1 . . . . .	117
4.30 Hasil Uji Nonlinearitas Klaster 2 . . . . .	120
4.31 Hasil Uji Nonlinearitas Klaster Jawa . . . . .	120
4.32 Hasil Estimasi BF, MI, MO Klaster 1 . . . . .	122
4.33 Hasil Estimasi BF, MI, MO Klaster 2 . . . . .	124
4.34 Hasil Estimasi BF, MI, MO Klaster Jawa . . . . .	125
4.35 Hasil Parameter MARS Klaster 1 . . . . .	129
4.36 Hasil Estimasi Parameter MARS Klaster 2 . . . . .	131
4.37 Hasil Estimasi Parameter MARS Klaster jawa . . . . .	133
4.38 Hasil Variabel <i>Importance</i> Klaster 1 . . . . .	135
4.39 Nilai GCV pada Setiap Subset Model . . . . .	136
4.40 Perhitungan Nilai $\Delta GCV$ . . . . .	137
4.41 Keberadaan Variabel $X_7$ pada Setiap Subset . . . . .	137
4.42 Hasil Variabel <i>Importance</i> Klaster 2 . . . . .	139
4.43 Hasil Variabel <i>Importance</i> Klaster Jawa . . . . .	140
4.44 Perbandingan Nilai Akurasi . . . . .	141
4.45 Perbandingan Waktu Komputasi . . . . .	144

# DAFTAR GAMBAR

2.1	Peta Pulau Jawa . . . . .	15
3.1	Flowchart Penelitian . . . . .	52
3.2	Flowchart Penelitian . . . . .	53
3.3	Flowchart Penelitian . . . . .	55
3.4	Flowchart Penelitian . . . . .	59
4.1	Visualisasi <i>Heatmap</i> Korelasi . . . . .	82
4.2	Pemilihan K Terbaik <i>Elbow Method</i> . . . . .	96
4.3	Hasil <i>Silhouette coefficient</i> . . . . .	107
4.4	Peta Pembagian Wilayah Klaster . . . . .	111
4.5	Hasil <i>Scatterplot</i> Klaster 1 . . . . .	114
4.6	Hasil <i>Scatterplot</i> Klaster 2 . . . . .	115
4.7	Hasil <i>Scatterplot</i> Klaster Jawa . . . . .	116

UIN SUNAN AMPEL  
S U R A B A Y A

## DAFTAR PUSTAKA

- Adhitya, B., Prabawa, A., and Kencana, H. (2022). Analisis pengaruh pendidikan, kesehatan, sanitasi dan rata-rata jumlah anggota keluarga per rumah tangga terhadap kemiskinan di indonesia. *Ekonomis: Journal of Economics and Business*, 6(1):288.
- Adriansah and U.S. (2022). Komparasi hasil analisis beda rata-rata menggunakan metode statistik parametrik dan nonparametrik. *Jurnal Riset Matematika dan Sains Terapan (JRMST)*, 2(2).
- Aggarwal, C. C. (2017). *Outlier Analysis*. Springer, Cham, 2 edition.
- Akbar, M. I. (2023). *Ekospiritualisme Al-Qur'an (Studi atas Tanggungjawab Manusia sebagai Khalifah Fi Al-Ardh dalam Penyelamatan Alam)*. PhD thesis, Institut PTIQ Jakarta.
- Alisjahbana, A. S. and Murniningtyas, E. (2020). Tujuan pembangunan berkelanjutan (sdgs) di indonesia. [https://sdgcenter.unpad.ac.id/wp-content/uploads/2020/04/Tujuan-Pembangunan-Berkelanjutan-SDGs-di-Indonesia\\_compressed.pdf](https://sdgcenter.unpad.ac.id/wp-content/uploads/2020/04/Tujuan-Pembangunan-Berkelanjutan-SDGs-di-Indonesia_compressed.pdf). SDG Center Universitas Padjadjaran. Accessed: 2025-11-21.
- Ampulembang, M. (2017). Pemodelan mars pada data kemiskinan di sulawesi. *Jurnal Matematika, Statistika dan Komputasi*.
- Anggraini, D., Sudharyati, N., Putra, R. A., Ramdhan, N., Putra, M. I. N., and Putra, H. H. (2023). Pengaruh pertumbuhan ekonomi, indeks pembangunan manusia (ipm), dan kemiskinan terhadap tingkat pengangguran terbuka di provinsi jambi selama tahun 2017–2021. 7:672–676.
- Ardian, R., Syahputra, M., and Dermawan, D. (2022). Pengaruh pertumbuhan ekonomi terhadap tingkat pengangguran terbuka di indonesia. *Jurnal Ekonomi, Bisnis Dan Manajemen*, 1(3):190–198.

- Arthur, D. and Vassilvitskii, S. (2007). K-means++: The advantages of careful seeding. pages 1027–1035.
- Aulia, M. K., Utaminingsih, E., and Prihatin, N. (2025). Model prediksi risiko kesehatan perkotaan berbasis lingkungan dengan xgboost. *Computer Science (CO-SCIENCE)*, 5(2):95–102.
- Azmi, F. (2025). Kebijakan desentralisasi kesehatan di indonesia: Dampak dan tantangan dalam pengelolaan kesehatan daerah. *Jurnal Kesehatan Tambusai*, 6(2):5735.
- Badan Perencanaan Pembangunan Nasional (2022). Strategi peningkatan indeks pembangunan manusia di Indonesia. <https://bappenas.go.id/berita/bappenas-sasaran-pembangunan-2023-untuk-transformasi-ekonomi>. Accessed: 2025-11-21.
- Badan Perencanaan Pembangunan Nasional (2023). Laporan tahunan sdgs 2023. <https://sdgs.bappenas.go.id/website/wp-content/uploads/2023/11/Laporan-tahunan-SDGs-2023.pdf>. Accessed: 2025-11-21.
- Badan Pusat Statistik (2023). Indeks pembangunan manusia. <https://www.bps.go.id/id/publication/2023/05/16/ef80bec78ab91cb5b703b943/indeks-pembangunan-manusia-2022.html>. Accessed: 2025-10-21.
- Badan Pusat Statistik (2025). Jumlah penduduk miskin masih didominasi di pulau jawa. <https://nasional.kontan.co.id/news/bps-jumlah-penduduk-miskin-masih-didominasi-di-pulau-jawa>. Accessed: 2025-11-21.
- Bartlett, M. S. (1954). A note on the multiplying factors for various  $\chi^2$  approximations. *Journal of the Royal Statistical Society, Series B*, 16(2):296–298. Dapat diakses melalui: <https://www.jstor.org/stable/2984057>.

- Boehmke, B. and Greenwell, B. (2019). *Hands-On Machine Learning with R*. Chapman and Hall/CRC, New York.
- Diani, M., Flora, R., and Syakurah, R. (2023). Optimalisasi pemerataan sdm kesehatan di indonesia. *Journal of Nursing and Public Health*, 11(1):234–245.
- Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing*. Marcel Dekker, 2 edition.
- Fatmawati, E. (2023). Perencanaan dasar dalam mengukur kajian indeks pembangunan literasi masyarakat (iplm) dan indeks literasi masyarakat (ilm). *IQRA': Jurnal Perpustakaan dan Informasi*.
- Friedman, J. H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics*, 19(1):1–67.
- García, S., Luengo, J., and Herrera, F. (2015). *Data Preprocessing in Data Mining*, volume 72 of *Intelligent Systems Reference Library*. Springer, Cham.
- Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2019). *Multivariate Data Analysis*. Cengage Learning, Hampshire, United Kingdom, 8 edition.
- Han, J., Kamber, M., and Pei, J. (2012a). *Data Mining: Concepts and Techniques*. Morgan Kaufmann, Waltham, MA, 3 edition.
- Han, J., Kamber, M., and Pei, J. (2012b). *Data Mining: Concepts and Techniques*. Morgan Kaufmann, Waltham, Massachusetts, 3 edition. Dapat diakses melalui: <https://www.sciencedirect.com/book/9780123814791>.
- Handayani, S. and Mutiara, N. (2025). Transformasi perpustakaan berbasis inklusi sosial upt perpustakaan proklamator bung karno dalam meningkatkan kesejahteraan masyarakat. *Otonomi*, 25(1):22–30.
- Handoyo, U. (2023). Pengaruh kepadatan penduduk terhadap indeks pembangunan manusia di indonesia. *Journal of Nusantara Economic Science*, 1(1):19–23.

- Harahap et al. (2023). Regresi data panel dalam analisis faktor-faktor yang mempengaruhi ipm di kalimantan barat. *Statistika*, 23(2).
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. Springer, 2 edition.
- Hidayanti, H. (2018). Pemerataan tenaga kesehatan di kabupaten lamongan (distribution of health workers in lamongan district). *Cakrawala*, 12(2).
- Indiati, P. S., Taufiq, M., Oktafia, R., and Sandi, P. (2024). Pengaruh kemiskinan dan pengangguran terhadap ipm di kabupaten jombang. *Jurnal Ilmiah Wahana Pendidikan*, 10(19):734–743.
- Irham, A. R. and Putri, R. M. (2023). Kepadatan penduduk terhadap indeks pembangunan manusia di provinsi lampung. 24(1):91–100.
- Johnson, R. A. and Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*. Pearson Prentice Hall, Upper Saddle River, New Jersey, 6 edition. Dapat diakses melalui: <https://www.pearson.com/en-us/subject-catalog/p/applied-multivariate-statistical-analysis/P200000006161>.
- Jolliffe, I. T. (2002). *Principal Component Analysis*. Springer Series in Statistics. Springer, New York, 2 edition. Dapat diakses melalui: <https://link.springer.com/book/10.1007/b98835>.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1):141–151. Dapat diakses melalui: <https://doi.org/10.1177/001316446002000116>.
- Kaiser, H. F. and Rice, J. (1974). Little jiffy, mark IV. *Educational and Psychological Measurement*, 34(1):111–117.
- Kartini, A. Y. and Cahyani, N. (2023). Hybrid k-means–multivariate adaptive regression splines (mars) for distribution of dengue fever risk mapping in bojonegoro district. *Barekeng: Journal of Mathematics and Its Applications*, 17(1):313–322.

Kartini, A. Y. and Ummah, L. N. (2022). Pemodelan kejadian balita stunting di kabupaten bojonegoro dengan metode Geographically Weighted Regression dan Multivariate Adaptive Regression Splines. *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, 15(1):127–136.

Kaufman, L. and Rousseeuw, P. J. (2005). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley Series in Probability and Statistics. Wiley-Interscience, Hoboken, New Jersey. Dapat diakses melalui: <https://www.wiley.com/en-us/Finding+Groups+in+Data-p-9780471735786>.

Kementerian Dalam Negeri Republik Indonesia (2024). Portal tanah air indonesia.

Kementerian Kesehatan Republik Indonesia (2021). Profil kesehatan Indonesia tahun 2020. [https://kemkes.go.id/app\\_asset/file\\_content\\_download/Profil-Kesehatan-Indonesia-2021.pdf](https://kemkes.go.id/app_asset/file_content_download/Profil-Kesehatan-Indonesia-2021.pdf). Accessed: 2025-11-21.

Kementerian Pendayagunaan Aparatur Negara dan Reformasi Birokrasi (2025). IPM Indonesia 2025 naik jadi 75,90: harapan hidup dan pendidikan terus meningkat. <https://www.menpan.go.id/site/berita-terkini/berita-daerah/ipm-indonesia-2025-naik-jadi-75-90-harapan-hidup-dan-pendidik>. Accessed: 2025-11-21.

Khasanah (2023). Peran umkm (usaha mikro kecil menengah) guna meningkatkan kesejahteraan masyarakat. *Demagogi: Journal of Social Sciences, Economics and Education*, 1(1):11–18.

Kodinariya, T. M. and Makwana, P. R. (2013). Review on determining number of cluster in K-Means clustering. *International Journal of Advance Research in Computer Science and Management Studies*, 1(6):90–95. Dapat diakses melalui: <http://ijarcsms.com/docs/paper/volume1/issue6/V1I6-0015.pdf>.

Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W. (2005). *Applied Linear Statistical Models*. McGraw-Hill, 5 edition.

- Mar'ah, Z., Ruliana, R., and Septiana, M. (2024). Pemodelan multivariate adaptive regression spline (MARS) pada indeks harga saham gabungan (IHSG) tahun 2018–2023. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 6(1):1–10.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3):519–530.
- Mattalunru, M. R., Annas, S., and Aidid, M. K. (2022). Aplikasi multivariate adaptive regression splines (MARS) untuk mengetahui faktor yang mempengaruhi curah hujan di kota makassar. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 4(1):9–19.
- Maulana, I., Salsabila, Z., and Dermawan, D. (2022). Analisis indeks pembangunan manusia. *Jurnal Madani Ilmiah Multidisipliner*.
- Milborrow, S. (2021). Notes on the earth package.
- Minu, I. W., Abubakar, A., and Sohrah, S. (2025). Analisis ayat al-qur'an tentang kemiskinan: Kajian tafsir tematik qs al-Zāriyāt ayat 19. *AL-QIBLAH: Jurnal Studi Islam dan Bahasa Arab*, 4(2):200–216.
- Mirah, A. et al. (2020). Pertumbuhan ekonomi dan kemiskinan di provinsi sulawesi utara. *Jurnal Pembangunan Ekonomi dan Keuangan Daerah*, 21(1):85–100.
- Montgomery, D. C., Peck, E. A., and Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. John Wiley & Sons, 5 edition.
- Muammar (2021). Analisis sektor unggulan di kota banda aceh berdasarkan metode tipologi klassen dan shift-share.
- Mukhlisin, M. and Solihudin, E. N. (2020). Kepemilikan jamban sehat pada masyarakat. *Faletehan Health Journal*, 7(3):119–123.
- Nabilah, N. A., Perdana, H., and Sulistianingsih, E. (2024). Pengelompokan provinsi di indonesia berdasarkan indikator kesejahteraan masyarakat dengan algoritma K-Means++. *Buletin Ilmiah Matematika Statistika dan Terapannya*, 13(3):419–426.

- Nasikh, M. A. (2024). Tantangan jawa-sentrisme menuju keseimbangan pembangunan. <https://fib.unair.ac.id/fib/2024/01/26/tantangan-jawa-sentrisme-menuju-keseimbangan-pembangunan/>. Fakultas Ilmu Budaya, Universitas Airlangga. Accessed: 2025-11-21.
- Nugroho, N. and Adhinata, F. D. (2022). Penggunaan metode k-means dan k-means++ sebagai clustering data covid-19 di pulau jawa. Institut Teknologi Telkom Purwokerto.
- Oktora, S. R. (2015). *Pemodelan MARS untuk Respon Kategorik*. Institut Pertanian Bogor.
- Pertiwi, Y., Permana, D., Amalita, N., and Salma, A. (2023). Modeling human development index in papua and west sumatera with multivariate adaptive regression spline. *UNP Journal of Statistics and Data Science*, 1(3):188–195.
- Pintowati, D. and Otok, B. W. (2012). Pemodelan mars pada data kemiskinan. *Jurnal Sains dan Seni ITS*.
- Pranata et al. (2022). Perbandingan metode klastering k-means dan dbscan. *Teorema: Teori dan Riset Matematika*, 9(2):247–260.
- Prastika, N. L. P. et al. (2021). Pemodelan mars dengan pendekatan gcv. *Jurnal Statistika*.
- Prosiding SainTek (2025). Pemetaan kabupaten/kota di provinsi jawa timur tahun 2023 berdasarkan karakteristik ipm dengan analisis biplot dan klasterisasi k-means. In *Prosiding Seminar Nasional Sains dan Teknologi SainTek*.
- Purnama, Y. (2024). *15 Prinsip Mencari Rezeki*. Self-published, Yogyakarta.
- Purwoko, A. and S., R. (2023). Pengaruh persentase penduduk miskin dan pertumbuhan ekonomi terhadap indeks pembangunan manusia (ipm) pada kabupaten berau. *Eco-Build Journal*, 7:16.

- Puspitasari, W. (2022). Muhadjir: Kemenko pmk kawal dua prioritas nasional. <https://www.antaraneews.com/berita/2922913/muhadjir-kemenko-pmk-kawal-dua-prioritas-nasional>. Published: June 6, 2022, Accessed: 2025-11-21.
- Rendi, R. I., Prasetiawati, E., and Amin, M. N. (2025). Ayat-ayat kepemimpinan dalam tafsir al-munir perspektif tafsir maqasidi abdul mustaqim. *Journal of Islamic Scriptures in Non-Arabic Societies*, 2(1):54–68.
- Restu, Z. D., Yulyani, V., and Perdana, A. A. (2022). Faktor-faktor yang berhubungan dengan kepemilikan jamban sehat di kelurahan pesawahan kota bandar lampung tahun 2021. *Jurnal Kesmas Khatulistiwa*, 9(4):209.
- Riyadi, S., Firdaus, A. S., and Anwar, K. (2025). Dasar al-qur'an dan hadits tentang kepemimpinan. *Integrative Perspectives of Social and Science Journal*, 2(2):2526–2540.
- Rousseuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20(1):53–65. Dapat diakses melalui: [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- Sabrina, F. A. (2023). Uji asumsi klasik untuk menghindari pelanggaran asumsi klasik pada regresi linier ordinary least squares (ols) dalam ekonometrika. *Jurnal Ilmiah Manajemen Ekonomi dan Akuntansi (JIMEA)*, 1(1):195–203.
- Salsabila, Hafiyusholeh, M., Khaulasari, H., Novitasari, D. C. R., and Yuniarti D.N (2024). Implementasi K-Means *Clustering* dalam pemetaan wilayah rawan penyakit leptospirosis di jawa timur. *Jurnal Riset dan Aplikasi Matematika*, 8(2):189–200.
- Shafana, N. R. and Gunawan, G. (2022). Analisis faktor yang mempengaruhi kesadaran penduduk dalam vaksin Covid-19 menggunakan metode Multivariate Adaptive Regression Spline. *Journal Riset Matematika*, 1(2).

- Shoimah, F. (2023). Mengukur indeks pembangunan literasi masyarakat (iplm) kabupaten probolinggo. *Pangripta: Jurnal Ilmiah Kajian Perencanaan Pembangunan*, 6(1):1164–1175.
- Sihombing, P. R. et al. (2024). Pemodelan ipm berdasarkan kelompok wilayah provinsi di indonesia. *Science Statistics Journal*, 2(1):35–47.
- Simanjuntak, J. V., Muchtar, M., and Sihombing, P. R. (2024). Pengaruh sanitasi, air minum, dan dana alokasi khusus (dak) fisik reguler terhadap indeks pembangunan manusia. *Journal of Law, Administration, and Social Science*, 4(2):317–327.
- Sinaga et al. (2025). Analisis faktor-faktor yang mempengaruhi indeks pembangunan manusia di provinsi jawa barat menggunakan regresi linier berganda. *Socius: Jurnal Penelitian Ilmu-Ilmu Sosial*, 2(12):568–582.
- Sriningsih, R., Otok, B. W., and Sutikno (2021). Factors affecting the number of dengue fever cases in West Sumatra province using the Multivariate Adaptive Regression Splines (MARS) approach. In *Journal of Physics: Conference Series*, volume 1722, page 012094. IOP Publishing.
- Subaekti, H., Hakim, L., Khaulasari, H., and Yuliati, D. (2026). An integrated K-Means++–Davies–Bouldin index approach for educational resource-based district clustering: A case study of districts in surabaya. *Jambura Journal of Mathematics*, 8(1).
- Suciani, A., Ruhiat, D., and Rahayu, S. D. (2022). Komparasi hasil analisis beda rata-rata menggunakan metode statistik parametrik dan nonparametrik. *JRMST: Jurnal Riset Matematika dan Sains Terapan*, 2(2).
- Sugiharti (2023). Analisis regresi kuantil terhadap faktor-faktor sosial ekonomi. In *Proceedings Saintek*. Universitas Terbuka.
- Sugiyono (2020). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta, Bandung, cetakan ke-2 edition.

- Sumayow, A. G., Pangemanan, P. A., and Tangkere, E. G. (2018). Analisis sektor basis dan non basis di kabupaten bolaang mongondow timur. *Agri-Sosioekonomi*, 14(2):279.
- Suraya et al. (2023). Tinjauan naratif tentang mengeksplorasi menyeluruh kelebihan k-means algoritma klasterisasi. *Jurnal Ilmu Data*, 3(1).
- Suryadi, A. (2021). *Menapak Indonesia: Menelusuri Setiap Wilayah Provinsi, Kabupaten dan Kota Seluruh Indonesia Jilid 2 (Pulau Jawa)*, volume 2. CV Jejak (Jejak Publisher).
- Susanti, I. and Saumi, F. (2022). Penerapan metode analisis regresi linear berganda untuk mengatasi masalah multikolinearitas pada kasus indeks pembangunan manusia (ipm) di kabupaten aceh tamiang. *Gamma-Pi: Jurnal Matematika dan Terapan*.
- Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., and Satoto, B. D. (2018). Integration k-means clustering method and elbow method for identification of the best customer profile cluster. *IOP Conference Series: Materials Science and Engineering*, 336(1):012017.
- Syarifah et al. (2025). Analisis pengaruh pendidikan, kesehatan, dan pengangguran terhadap ipm di indonesia. *Jurnal Kompetensi Ilmu Sosial*, 3(2).
- Tan, P.-N., Steinbach, M., Karpatne, A., and Kumar, V. (2019). *Introduction to Data Mining*. Pearson Education, New York, NY, 2 edition.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.
- Ulinuh, N. et al. (2025). Pengelompokan pembangunan provinsi di indonesia dengan analisis komponen utama dan cluster hierarki: Berdasarkan data kemiskinan, pengangguran, pendidikan, dan ipm. In *Prosiding Seminar Nasional Statistika Aktuaria*, volume 4, pages 1–16.
- United Nations Development Programme (2023). Human development report 2023. <https://hdr.undp.org/data-center/human-development-index>. Accessed: 2025-11-21.

- United Nations Development Programme (2024). Human development report 2023/2024: Breaking the gridlock – reimagining cooperation in a polarized world. Technical report, United Nations Development Programme, New York.
- Wahba, G. (1990). *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics.
- Wicaksono, D. S., Basith, A., Yusuf, A., Chaerunisa, Sanjaya, N., and Machfud, S. (2025). Analisis kluster dan prediksi indikator ekonomi makro indonesia menggunakan algoritma K-Means dan model regresi. *PESHUM: Jurnal Pendidikan, Sosial dan Humaniora*.
- Witari, R. A. and Astuti, E. T. (2024). Urgensi pemberantasan stunting perspektif al-qur'an surat an-nisa' ayat 9. *Jurnal Inovasi Pendidikan*, 2(2):79–87.
- Wu, H. and Zhang, J.-T. (2006). *Nonparametric Regression Methods for Longitudinal Data Analysis*. John Wiley & Sons.
- Zai, A. et al. (2025). Analisis k-means cluster kabupaten/kota di provinsi kalimantan selatan berdasarkan indikator ipm. *Komputika: Jurnal Sistem Komputer*, 14(2).
- Zannah, M., Rizki, S. W., and Aprizkiyandari, S. (2022). Pemodelan tingkat pengangguran terbuka di kalimantan barat dengan pendekatan linear mixed model. *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)*, 11:677–686.