

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PENERAPAN *REGRESSION-BASED CLUSTERING* DAN ANALISIS DATA PANEL UNTUK MENGIDENTIFIKASI FAKTOR DEMAM BERDARAH DI JAWA BARAT

ADISTI SUCI RAHMAH



**PROGRAM STUDI STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Penerapan *Regression-based Clustering* dan Analisis Data Panel untuk Mengidentifikasi Faktor Demam Berdarah di Jawa Barat” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juni 2025

Adisti Suci Rahmah
G1401211027

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRAK

ADISTI SUCI RAHMAH. Penerapan *Regression-based Clustering* dan Analisis Data Panel Untuk Mengidentifikasi Faktor Demam Berdarah di Jawa Barat. Dibimbing oleh I MADE SUMERTAJAYA dan LA ODE ABDUL RAHMAN.

Demam berdarah dengue (DBD) merupakan salah satu penyakit endemis yang menjadi masalah negara tropis dan subtropis. Provinsi Jawa Barat secara konsisten mencatat jumlah kasus DBD tertinggi di Indonesia sejak tahun 2017 hingga tahun 2023 sehingga perlu perhatian dan pemahaman mengenai faktor yang berkontribusi dalam penyebaran DBD. Penelitian ini bertujuan untuk mengidentifikasi pola pertumbuhan kasus DBD di Jawa Barat melalui pengelompokan wilayah berdasarkan parameter pertumbuhan kasus menggunakan algoritma *partitioning around medoid* (PAM) serta menganalisis faktor-faktor yang berpengaruh terhadap DBD menggunakan analisis data panel. Data yang digunakan merupakan data sekunder dari 27 kabupaten/kota di Jawa Barat selama periode 2016–2023. Hasil pengelompokan membentuk empat gerombol, yaitu Gerombol 1 wilayah dengan jumlah kasus awal tinggi dan cenderung menurun perlahan, Gerombol 2 wilayah dengan jumlah kasus awal rendah hingga sedang mengalami lonjakan bertahap, Gerombol 3 wilayah dengan jumlah kasus awal rendah namun mengalami lonjakan signifikan dalam waktu singkat, dan Gerombol 4 wilayah dengan jumlah kasus awal tinggi yang berhasil menurunkan angka kasus secara tajam. Analisis data panel menunjukkan bahwa model terbaik adalah model *random effect* spesifik waktu, dengan nilai *R-squared* sebesar 57,24%. Faktor yang berpengaruh signifikan terhadap jumlah kasus DBD adalah persentase rumah tangga yang memiliki akses terhadap sanitasi layak dan rata-rata lama sekolah.

Kata kunci: analisis data panel, demam berdarah, *incidence rate*, *partitioning around medoid* (PAM), *regression-based clustering*.

ABSTRACT

ADISTI SUCI RAHMAH. Application of Regression-based Clustering and Panel Data Analysis to Identify Factors of Dengue Fever in West Java. Supervised by I MADE SUMERTA JAYA and LA ODE ABDUL RAHMAN.

Dengue hemorrhagic fever is one of the endemic diseases that poses a significant problem in tropical and subtropical countries. West Java Province has consistently recorded the highest number of DHF cases in Indonesia from 2017 to 2023, necessitating a deeper understanding of its contributing factors. This study aims to identify the growth patterns of DHF cases in West Java by clustering regions based on case growth parameters using the Partitioning Around Medoid (PAM) and analyzing the factors affecting DHF incidence through panel data analysis. The data used in this study are secondary data from 27 districts and cities in West Java over the period of 2016–2023. The clustering analysis resulted in four clusters, Cluster 1 consists of regions with a high initial case that gradually decreased, Cluster 2 includes regions with low to moderate initial cases experiencing a gradual increase, Cluster 3 covers regions with initially low cases that experienced a sharp surge, and Cluster 4 consists of regions with high initial cases that rapidly decreased. The best model had random effects with a time-specific model, with R-squared 57,24%. Significant factors were the percentage of households with access to proper sanitation and the average length of schooling.

Keywords: dengue hemorrhagic fever, incidence rate, panel data analysis partitioning around medoid (PAM), regression-based clustering.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025¹
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

PENERAPAN *REGRESSION-BASED CLUSTERING* DAN ANALISIS DATA PANEL UNTUK MENGIDENTIFIKASI FAKTOR DEMAM BERDARAH DI JAWA BARAT

ADISTI SUCI RAHMAH

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Judul Skripsi : Penerapan *Regression-based Clustering* dan Analisis Data Panel
untuk Mengidentifikasi Faktor Demam Berdarah di Jawa Barat

Nama : Adisti Suci Rahmah
NIM : G1401211027

Disetujui oleh

Pembimbing 1:
Prof. Dr. Ir. I Made Sumertajaya, M.Si.



Pembimbing 2:
La Ode Abdul Rahman, S.Si., M.Si.



Diketahui oleh

Ketua Program Studi Statistika dan Sains Data:
Dr. Bagus Sartono, S.Si., M.Si.
NIP 19780411200501 1 002



Tanggal Ujian:
30 Juni 2025

Tanggal Lulus:



- 

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Januari 2025 sampai bulan Juni 2025 ini ialah penerapan statistika dalam kasus kesehatan, dengan judul “Penerapan *Regression-based Clustering* dan Analisis Data Panel Untuk Mengidentifikasi Faktor Demam Berdarah di Jawa Barat”.

Terima kasih penulis ucapkan kepada semua pihak yang telah memberikan doa, dukungan, dan kontribusi dalam penyelesaian karya ilmiah ini, diantaranya:

1. Bapak Prof. Dr. Ir. I Made Sumertajaya, M.Si. dan Bapak La Ode Abdul Rahman, S.Si., M.Si. selaku dosen pembimbing yang telah memberikan arahan dan bimbingan selama penulisan karya ilmiah ini.
2. Ibu Pika Silvianti, S.Si., M.Si selaku dosen penguji yang telah memberikan saran dan masukan dalam penulisan karya ilmiah ini.
3. Seluruh dosen dan staff Departemen Statistika SSMI IPB yang membantu penulis selama perkuliahan sehingga dapat menyelesaikan pendidikan dan administrasi sampai lulus.
4. Bapak Syahrial, Ibu Desri Sulistyawati, Diana Fadhilah, Haya Luthfiyah, dan Keluarga Besar Arifin RB yang senantiasa memberikan dukungan, cinta, kasih sayang, dan doa selama penulis melaksanakan studi.
5. Rafiqah, Nurul, Syifa, dan Fristy selaku sahabat penulis yang telah kebersamai, memberikan motivasi, dan mendengarkan curhatan penulis selama perkuliahan dan penulisan karya ilmiah ini.
6. Zulfa, Naswa, Oktavia, Vita, Farras, Cahyani, Alfikri, dan Farrel selaku sahabat penulis yang telah menemani penulis selama berkuliah di Bogor, yang selalu ada untuk kebersamai serta memberikan dukungan, nasihat, dan arahan selama perkuliahan.
7. Nalla, Faishal, dan Atif sebagai sahabat penulis sesama mahasiswa rantau dari Padang yang selalu ada dan selalu mendukung penulis untuk menyelesaikan karya ilmiah ini.
8. Hakim, Reynd, Nadila, Fazrin, Dindana, Arfiah, Fuadiyah, Natasha, dan Luthfi sebagai teman satu bimbingan penulis yang telah berjuang bersama mulai dari persiapan, pengerjaan, bimbingan, sampai karya ilmiah ini selesai.
9. Teman-teman Statistika 58 dan seluruh pihak yang telah memberikan kontribusi dan kebaikan, meskipun tidak dapat disebutkan satu per satu, sehingga penulis berhasil menyelesaikan karya ilmiah ini.

Penulis menyadari masih banyak kekurangan dalam penyusunan, kritik dan saran membangun sangat diharapkan untuk kesempurnaan karya ilmiah ini. Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Juni 2025

Adisti Suci Rahmah



DAFTAR ISI

DAFTAR TABEL	xi
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xi
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	2
II TINJAUAN PUSTAKA	3
2.1 Demam Berdarah Dengue (DBD)	3
2.2 <i>Incidence Rate</i>	3
2.3 Analisis Gerombol	3
2.4 Evaluasi Hasil Penggerombolan	5
2.5 Regresi Data Panel	6
2.6 Model Pendugaan Parameter Data Panel	6
2.7 Spesifikasi Model Regresi Data Panel	8
2.8 Uji Asumsi Klasik	10
III METODE	12
3.1 Data	12
3.2 Prosedur Analisis	13
IV HASIL DAN PEMBAHASAN	15
4.1 Eksplorasi Data	15
4.2 Analisis Gerombol	18
4.3 Individu Baru Melalui Agregasi	21
4.4 Pemeriksaan Multikolinearitas	21
4.5 Spesifikasi Model	22
4.6 Pengujian Asumsi Sisaan	27
4.7 Interpretasi Model Terbaik	28
V SIMPULAN DAN SARAN	31
5.1 Simpulan	31
5.2 Saran	31
DAFTAR PUSTAKA	32
LAMPIRAN	36
RIWAYAT HIDUP	41

DAFTAR TABEL

1	Tabel 1 Rentang nilai <i>silhouette</i> dan interpretasi	5
2	Tabel 2 Daftar peubah penjelas	12
3	Tabel 3 Hasil evaluasi gerombol	19
4	Tabel 4 Hasil penggerombolan	20
5	Tabel 5 Karakteristik gerombol	20
6	Tabel 6 Pengujian multikolinearitas	22
7	Tabel 7 Hasil pendugaan parameter model gabungan	22
8	Tabel 8 Hasil pendugaan parameter model pengaruh tetap	23
9	Tabel 9 Hasil pendugaan parameter model pengaruh acak	25
10	Tabel 10 Hasil analisis uji Chow	26
11	Tabel 11 Hasil uji asumsi klasik	27
12	Tabel 12 Hasil pendugaan parameter model setelah transformasi	28
13	Tabel 13 Hasil uji asumsi klasik baru	28
14	Tabel 14 Nilai pengaruh acak waktu model terbaik	29

DAFTAR GAMBAR

1	Gambar 1 Tren <i>incidence rate</i> DBD Jawa Barat tahun 2016-2023	15
2	Gambar 2 Tren <i>incidence rate</i> masing-masing kabupaten/kota	15
3	Gambar 3 <i>Boxplot</i> masing-masing peubah	16
4	Gambar 4 <i>Scatterplot</i> peubah penjelas dengan peubah respon	17
5	Gambar 5 Matriks korelasi antar peubah	17
6	Gambar 6 Grafik <i>k</i> optimal	18
7	Gambar 7 Profil (a) Gerombol 1 (b) Gerombol 2 (c) Gerombol 3 (d) Gerombol 4	19
8	Gambar 8 Tren nilai rata-rata <i>incidence rate</i> gerombol	20
9	Gambar 9 Pemetaan hasil gerombol	21

DAFTAR LAMPIRAN

1	Lampiran 1 Hasil pendugaan parameter <i>incidence rate</i>	37
2	Lampiran 2 Individu baru yang setelah proses penggerombolan	38
3	Lampiran 3 Nilai pengaruh tetap individu	38
4	Lampiran 4 Nilai pengaruh tetap waktu	39
5	Lampiran 5 Nilai pengaruh tetap dua arah	39
6	Lampiran 6 Nilai pengaruh acak individu	39
7	Lampiran 7 Nilai pengaruh acak waktu	40
8	Lampiran 8 Nilai pengaruh acak dua arah	40

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.