



**PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS
PADA PENGELOMPOKAN KABUPATEN/KOTA ACUAN
PENGHITUNG INFLASI DI INDONESIA BERDASARKAN
KELOMPOK PENGELUARAN**

SKRIPSI

**untuk memenuhi persyaratan
dalam menyelesaikan program sarjana Strata-1 Statistika**

**Oleh
REGINA WATI NASYIKIN
NIM. 2111017320014**

**PROGRAM STUDI S-1 STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
JULI 2025**

HALAMAN JUDUL



**PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS
PADA PENGELOMPOKAN KABUPATEN/KOTA ACUAN
PENGHITUNG INFLASI DI INDONESIA BERDASARKAN
KELOMPOK PENGELUARAN**

SKRIPSI

**untuk memenuhi persyaratan
dalam menyelesaikan program sarjana Strata-1 Statistika**

**Oleh
REGINA WATI NASYIKIN
NIM. 2111017320014**

**PROGRAM STUDI S-1 STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
JULI 2025**

SKRIPSI

PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS PADA PENGELOMPOKAN KABUPATEN/KOTA ACUAN PENGHITUNG INFLASI DI INDONESIA BERDASARKAN KELOMPOK PENGELUARAN

Oleh
Regina Wati Nasyikin
NIM. 2111017320014

Telah dipertahankan pada hari Jum'at, tanggal 4-7-2025 dan disetujui oleh dosen pembimbing dan dosen penguji sebagai berikut:

Pembimbing I

Yeni Rahkmawati, S.Mat., M.Si
NIP. 199404032022032014

Penguji I

Selvi Annisa, S.Si., M.Si
NIP. 199212262022032016

Pembimbing II

Fuad Muhajirin Farid, S.Pd., M.Si
NIP. 198807112019031014

Penguji II

Irfan Azam Fikri, S.ST., M.Sc., M.S.E
NIP. 198903152012111001

Banjarbaru, 11 Juli 2025
Mengetahui,
Koordinator Program Studi Statistika
EMPA ULM



Prof. Dewi Anggraini, S.Si., M.App.Sci., Ph.D
NIP. 198303282005012001

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, 11 Juli 2025



REGINA WATI NASYIKIN

NIM. 2111017320014

ABSTRAK

Perbandingan Algoritma K-Means dan K-Medoids pada Pengelompokan Kabupaten/Kota Acuan Penghitung Inflasi di Indonesia Berdasarkan Kelompok Pengeluaran (Oleh: Regina Wati Nasyikin; Pembimbing: Yeni Rahkmawati dan Fuad Muhajirin Farid, 2025; 46 halaman)

Inflasi merupakan indikator penting dalam menilai kestabilan negara serta berperan krusial dalam arah kebijakan moneter dan fiskal. Inflasi dihitung menggunakan Indeks Harga Konsumen yang terbagi dalam sebelas kelompok pengeluaran. Tingkat inflasi *year on year* kelompok pengeluaran antar kabupaten/kota di Indonesia cenderung berbeda. Sehingga diperlukan analisis untuk mengelompokkan wilayah dengan pola inflasi yang mirip. Metode *clustering* digunakan karena dapat mengidentifikasi kelompok wilayah berdasarkan kemiripan data. Penelitian ini bertujuan menentukan algoritma *clustering* terbaik antara K-Means dan K-Medoids menggunakan jarak *Euclidean* dan *Squared Euclidean*. Pengklasteran dilakukan terhadap kabupaten/kota acuan penghitung inflasi *year on year* berdasarkan kelompok pengeluaran. Analisis dilakukan pada data bulan Maret, Juli dan Desember 2024. Penelitian ini juga menganalisis sebaran cluster secara geografis. Tahap penelitian meliputi *pre-processing* data, penerapan algoritma *clustering*, evaluasi menggunakan *silhouette coefficient*, dan visualisasi peta sebaran cluster. Hasil menunjukkan bahwa algoritma K-Means dengan jarak *Squared Euclidean* memberikan hasil pengklasteran terbaik. Nilai *silhouette coefficient* untuk data bulan Maret sebesar 0.375 (cluster lemah), Juli sebesar 0.258 (cluster lemah), dan Desember sebesar 0.239 (cluster tidak terstruktur dengan baik). Cluster yang terbentuk menggambarkan sebaran inflasi antar wilayah di Indonesia. Hasil ini dapat menjadi dasar pengambilan kebijakan pengendalian inflasi, khususnya pada cluster 2 yang memiliki inflasi *year on year* relatif tinggi.

Kata Kunci: K-Means, K-Medoids, *Euclidean*, *Squared Euclidean*, Inflasi

ABSTRACT

Comparison of K-Means and K-Medoids Algorithms in Clustering Regencies/Cities Used as Reference for Inflation Calculation in Indonesia Based on Expenditure Groups (By: Regina Wati Nasyikin; Supervisors: Yeni Rahkmawati and Fuad Muhajirin Farid, 2025; 46 pages)

Inflation is an important indicator of a country's economic stability and plays a crucial role in shaping monetary and fiscal policies. Inflation is measured using the Consumer Price Index (CPI), which is categorized into eleven expenditure groups. Year-on-year inflation rates vary across regencies/cities in Indonesia, requiring analysis to group regions with similar inflation patterns. Clustering methods are suitable as they identify regional groupings based on data similarity. This study aims to determine the best clustering algorithm between K-Means and K-Medoids using Euclidean and Squared Euclidean distances. Clustering is applied to regencies/cities that serve as references in calculating year-on-year inflation based on by expenditure group. The analysis uses data from March, July, and December 2024. This study also analyzes the geographical distribution of the clusters. The research stages include data preprocessing, implementation of clustering algorithms, silhouette coefficient evaluation, and visualization of cluster distribution maps. Results indicate that K-Means with Squared Euclidean distance provides the best clustering performance. The silhouette coefficient values are 0.375 for March (weak cluster), 0.258 for July (weak cluster), and 0.239 for December (poorly structured cluster). The resulting clusters illustrate the distribution of inflation across regions in Indonesia. These findings can support the formulating of inflation control policies, particularly for cluster 2, which show relatively high year-on-year inflation.

Keywords: K-Means, K-Medoids, Euclidean, Squared Euclidean, Inflation

PRAKATA

Puji dan syukur kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Perbandingan Algoritma K-Means dan K-Medoids pada Pengelompokan Kabupaten/Kota Acuan Penghitung Inflasi di Indonesia Berdasarkan Kelompok Pengeluaran”**. Skripsi ini disusun untuk memenuhi salah satu syarat dalam memperoleh gelar Sarjana Statistika di Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat. Selama proses penulisan, penulis mendapatkan banyak bimbingan, dukungan, dan bantuan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada:

1. Kedua orang tua, Asyikin dan Selowojowati, serta kakak dan adik, Mardiyawati Nasyikin, S.Pd; Sustinawati Nasyikin, S.E; Muhammad Sukma Nasyikin; dan Muhammad Ikhsan, S.Pd beserta keluarga besar yang senantiasa memberikan doa semangat, hingga dukungan secara moril maupun materiil.
2. Ibu Yeni Rahkmawati, S.Mat., M.Si dan Bapak Fuad Muhajirin Farid, S.Pd., M.Si selaku dosen pembimbing yang telah dengan sabar memberikan bimbingan, arahan, dan motivasi dari tahap pengusulan judul hingga penyelesaian skripsi.
3. Ibu Selvi Annisa, S.Si., M.Si dan Bapak Irfan Azam Fikri, S.ST., M.Sc., M.S.E selaku dewan penguji yang telah memberikan masukan, saran, dan bantuan sehingga skripsi ini menjadi lebih baik.
4. Prof. Dewi Anggraini, S.Si., M.App.Sci., Ph.D selaku Koordinator Program Studi Statistika FMIPA ULM, serta seluruh dosen dan staf yang telah memberikan ilmu, motivasi, dan dukungan selama perkuliahan.
5. Teman-teman SD hingga kuliah yang selalu mendukung, menemani, mendoakan, dan memberi semangat, khususnya Danah, Siti, Awa, Enor, Putjah, Asih, Ayu, Nita, Diva, Ajeng, Meyrlin, Pute, Eca, Fatia, Ian, dan Hani.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, akan tetapi penulis berharap penelitian ini bermanfaat bagi semua pihak.

Banjarbaru, 11 Juli 2025

Regina Wati Nasyikin

DAFTAR ISI

| | |
|--|-----|
| HALAMAN JUDUL | i |
| HALAMAN PENGESAHAN..... | ii |
| PERNYATAAN | iii |
| ABSTRAK..... | iv |
| <i>ABSTRACT</i> | v |
| PRAKATA | vi |
| DAFTAR ISI..... | vii |
| DAFTAR GAMBAR | ix |
| DAFTAR TABEL | x |
| DAFTAR LAMPIRAN | xi |
| DAFTAR ISTILAH, SINGKATAN, DAN LAMBANG | xii |
| BAB I PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang..... | 1 |
| 1.2 Rumusan Masalah | 4 |
| 1.3 Tujuan Penelitian..... | 4 |
| 1.4 Manfaat Penelitian | 4 |
| BAB II TINJAUAN PUSTAKA..... | 5 |
| 2.1 Penelitian Terdahulu | 5 |
| 2.2 Kajian Teori..... | 8 |
| 2.2.1 Inflasi | 8 |
| 2.2.2 Analisis Statistika Deskriptif | 9 |
| 2.2.3 <i>Clustering</i> | 9 |
| 2.2.4 Ukuran Kemiripan..... | 11 |
| 2.2.5 Algoritma K-Means <i>Clustering</i> | 12 |
| 2.2.6 Algoritma K-Medoids <i>Clustering</i> | 13 |
| 2.2.7 <i>Silhouette Coefficient</i> | 14 |
| BAB III METODE PENELITIAN | 16 |
| 3.1 Sumber Data | 16 |
| 3.2 Variabel Penelitian | 16 |
| 3.3 Prosedur Penelitian | 18 |
| BAB IV HASIL DAN PEMBAHASAN | 21 |
| 4.1 <i>Pre-Processing</i> Data..... | 21 |
| 4.2 Analisis Deskriptif..... | 22 |
| 4.3 Normalisasi Data..... | 23 |
| 4.4 Pemenuhan Asumsi <i>Clustering</i> | 25 |
| 4.5 Klasterisasi Menggunakan Algoritma K-Means..... | 26 |
| 4.6 Klasterisasi Menggunakan Algoritma K-Medoids | 28 |
| 4.7 Evaluasi <i>Silhouette Coefficient</i> | 30 |
| 4.8 Peta Sebaran Cluster | 32 |
| 4.8.1 Peta Sebaran Cluster Data Bulan Maret..... | 32 |
| 4.8.2 Peta Sebaran Cluster Data Bulan Juli | 36 |
| 4.8.3 Peta Sebaran Cluster Data Bulan Desember | 40 |
| BAB V PENUTUP | 45 |
| 5.1 Kesimpulan | 45 |

| | |
|----------------------|-----|
| 5.2 Saran..... | 46 |
| DAFTAR PUSTAKA | 47 |
| LAMPIRAN..... | 53 |
| RIWAYAT HIDUP | 119 |

PRODI STATISTIKA

DAFTAR GAMBAR

| | | |
|------------|--|----|
| Gambar 1.1 | Boxplot Inflasi <i>Year on Year</i> Bulan Desember Tahun 2024..... | 2 |
| Gambar 3.1 | Prosedur Penelitian..... | 20 |
| Gambar 4.1 | Ringkasan <i>Pre-processing</i> pada Data Inflasi <i>Year on Year</i> Kelompok Pengeluaran Bulan Maret, Juli, dan Desember..... | 21 |
| Gambar 4.2 | Boxplot Inflasi <i>Year on Year</i> Bulan Maret, Juli, dan Desember Tahun 2024..... | 22 |
| Gambar 4.3 | Matriks Korelasi <i>Pearson</i> Data Bulan Maret, Juli, dan Desember Tahun 2024..... | 26 |
| Gambar 4.4 | Peta Sebaran Cluster Data Bulan Maret Tahun 2024..... | 35 |
| Gambar 4.5 | Peta Sebaran Cluster Data Bulan Juli Tahun 2024..... | 39 |
| Gambar 4.6 | Peta Sebaran Cluster Data Bulan Desember Tahun 2024..... | 43 |

DAFTAR TABEL

| | | |
|------------|---|----|
| Tabel 2.1 | Kajian Penelitian Terdahulu Mengenai Pengklasteran Kabupaten/Kota pada Kasus Inflasi..... | 5 |
| Tabel 2.2 | Kajian Penelitian Terdahulu Mengenai Perbandingan Algoritma K-Means dan K-Medoids | 6 |
| Tabel 2.3 | Kriteria Korelasi <i>Pearson</i> | 11 |
| Tabel 2.4 | Kriteria <i>Silhouette Coefficient</i> | 15 |
| Tabel 3.1 | Variabel Penelitian..... | 16 |
| Tabel 4.1 | Hasil <i>Min-Max Normalization</i> pada Data Bulan Maret | 24 |
| Tabel 4.2 | Hasil <i>Min-Max Normalization</i> pada Data Bulan Juli..... | 24 |
| Tabel 4.3 | Hasil <i>Min-Max Normalization</i> pada Data Bulan Desember..... | 25 |
| Tabel 4.4 | <i>Centroid</i> Iterasi 1..... | 26 |
| Tabel 4.5 | Hasil Perhitungan K-Means Jarak <i>Squared Euclidean</i> Iterasi 1.... | 27 |
| Tabel 4.6 | <i>Medoids</i> Iterasi 1 | 28 |
| Tabel 4.7 | Hasil Perhitungan K-Medoids Jarak <i>Squared Euclidean</i> Iterasi 1 | 29 |
| Tabel 4.8 | <i>Silhouette Coefficient</i> Algoritma K-Means | 30 |
| Tabel 4.9 | <i>Silhouette Coefficient</i> Algoritma K-Medoids | 31 |
| Tabel 4.10 | Ringkasan <i>Silhouette Coefficient</i> Algoritma K-Means dan K-Medoids | 31 |
| Tabel 4.11 | Hasil Cluster Algoritma K-Means dengan Jarak <i>Squared Euclidean</i> Data Bulan Maret..... | 33 |
| Tabel 4.12 | Hasil Cluster Algoritma K-Means dengan Jarak <i>Squared Euclidean</i> Data Bulan Juli..... | 37 |
| Tabel 4.13 | Hasil Cluster Algoritma K-Means dengan Jarak <i>Squared Euclidean</i> Data Bulan Desember | 41 |

DAFTAR LAMPIRAN

| | |
|---|-----|
| Lampiran 1. Data Inflasi <i>Year on Year</i> Bulan Maret 2024..... | 53 |
| Lampiran 2. Data Inflasi <i>Year on Year</i> Bulan Juli 2024 | 57 |
| Lampiran 3. Data Inflasi <i>Year on Year</i> Bulan Desember 2024 | 61 |
| Lampiran 4. Hasil Normalisasi Data Bulan Maret 2024 | 65 |
| Lampiran 5. Hasil Normalisasi Data Bulan Juli 2024..... | 69 |
| Lampiran 6. Hasil Normalisasi Data Bulan Desember 2024..... | 73 |
| Lampiran 7. Ilustrasi Perhitungan K-Means Jarak <i>Euclidean</i> Data Bulan Maret | 78 |
| Lampiran 8. Ilustrasi Perhitungan K-Means Jarak <i>Squared Euclidean</i> Data Bulan Maret | 81 |
| Lampiran 9. Ilustrasi Perhitungan K-Means Jarak <i>Euclidean</i> Data Bulan Juli..... | 84 |
| Lampiran 10. Ilustrasi Perhitungan K-Means Jarak <i>Squared Euclidean</i> Data Bulan Juli..... | 87 |
| Lampiran 11. Ilustrasi Perhitungan K-Means Jarak <i>Euclidean</i> Data Bulan Desember..... | 90 |
| Lampiran 12. Ilustrasi Perhitungan K-Means Jarak <i>Squared Euclidean</i> Data Bulan Desember | 93 |
| Lampiran 13. Ilustrasi Perhitungan K-Medoids Jarak <i>Euclidean</i> Data Bulan Maret | 96 |
| Lampiran 14. Ilustrasi Perhitungan K-Medoids Jarak <i>Squared Euclidean</i> Data Bulan Maret..... | 99 |
| Lampiran 15. Ilustrasi Perhitungan K-Medoids Jarak <i>Euclidean</i> Data Bulan Juli..... | 102 |
| Lampiran 16. Ilustrasi Perhitungan K-Medoids Jarak <i>Squared Euclidean</i> Data Bulan Juli | 105 |
| Lampiran 17. Ilustrasi Perhitungan K-Medoids Jarak <i>Euclidean</i> Data Bulan Desember..... | 108 |
| Lampiran 18. Ilustrasi Perhitungan K-Medoids Jarak <i>Squared Euclidean</i> Data Bulan Desember | 111 |
| Lampiran 19. Rata-Rata Setiap Variabel di Masing-Masing Cluster K-Means Menggunakan <i>Squared Euclidean</i> Data Bulan Maret | 114 |
| Lampiran 20. Rata-Rata Setiap Variabel di Masing-Masing Cluster K-Means Menggunakan <i>Squared Euclidean</i> Data Bulan Juli..... | 114 |
| Lampiran 21. Rata-Rata Setiap Variabel di Masing-Masing Cluster K-Means Menggunakan <i>Squared Euclidean</i> Data Bulan Desember | 114 |
| Lampiran 22. <i>Syntax</i> RStudio Data Bulan Maret | 114 |

DAFTAR ISTILAH, SINGKATAN, DAN LAMBANG

| | | |
|-----------------------------|---|---|
| moneter | : | berhubungan dengan uang atau keuangan |
| fiskal | : | berkenaan dengan urusan pajak atau pendapatan negara |
| krusial | : | gawat: genting |
| <i>outlier</i> | : | nilai data yang berbeda dari nilai lain dalam satu set data |
| fluktuasi | : | gejala yang menunjukkan turun-naiknya harga |
| Indeks Harga Konsumen (IHK) | : | indeks yang menghitung rata-rata perubahan harga dari suatu kelompok barang dan jasa yang dikonsumsi oleh rumah tangga dalam kurun waktu tertentu |
| Survei Biaya Hidup (SBH) | : | survei pengeluaran konsumsi rumah tangga di daerah perkotaan (urban area) dan pedesaan (rural area) untuk menghasilkan pola konsumsi masyarakat |
| geopolitik | : | ilmu tentang pengaruh faktor geografi terhadap ketatanegaraan |
| inflasi <i>year on year</i> | : | perubahan harga barang dan jasa secara keseluruhan dalam satu tahun dibandingkan dengan tahun sebelumnya pada periode yang sama |
| <i>safe haven</i> | : | investasi yang dapat mempertahankan atau meningkatkan nilainya selama masa krisis ekonomi |
| cluster | : | kelompok |
| <i>centroid</i> | : | titik pusat |
| BPS | : | Badan Pusat Statistik |
| BMKG | : | Badan Meteorologi, Klimatologi, dan Geofisika |
| Staklim | : | Stasiun Klimatologi |
| d_{ij} | : | jarak objek data ke- i dengan <i>centroid</i> ke- j |
| X_{hi} | : | nilai dari variabel ke- h pada objek data ke- i |
| C_{hj} | : | nilai dari <i>centroid</i> ke- j pada variabel data ke- h |
| C_{hj}^* | : | <i>centroid</i> baru <i>cluster</i> ke- j pada variabel ke- h |
| n_j | : | banyaknya data pada <i>cluster</i> ke- j |
| $s(i)$ | : | <i>silhouette coefficient</i> |
| k | : | jumlah <i>cluster</i> |