



**PERBANDINGAN KUALITAS KLASTERISASI METODE K-MEANS
DAN X-MEANS PADA DATA KETAHANAN PANGAN
KABUPATEN/KOTA DI KALIMANTAN SELATAN**

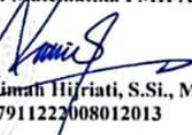
SKRIPSI

**Untuk memenuhi persyaratan dalam
menyelesaikan program sarjana Strata-1 Matematika**

**Oleh:
ARSELLA PUTRI NABILA
NIM. 2211011220002**

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
2026**

HALAMAN PENGESAHAN

SKRIPSI	
PERBANDINGAN KUALITAS KLASTERISASI METODE K-MEANS DAN X-MEANS PADA DATA KETAHANAN PANGAN KABUPATEN/KOTA DI KALIMANTAN SELATAN	
Oleh: Arsella Putri Nabila NIM 2211011220002	
telah dipertahankan di depan Dosen Penguji pada tanggal 28 April 2026	
Susunan Dosen Penguji:	
Pembimbing I  Thresye, S.Si., M.Si. NIP 197205042000122002	Dosen Penguji: 1. Oni Soesanto, S.Si., M.Si. () 2. Saman Abdurrahman, S.Si., M.Sc. ()
Pembimbing II  Hermei Lissa, S.Pd., M.Si. NIP 199005222022032012	
Banjarbaru, 25 Mei 2026 Jurusan Matematika FMIPA ULM Ketua,  Dr. Na'imah Hijriati, S.Si., M.Si. NIP 197911222008012013	

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, 9 Mei 2026



Arsellia Putri Nabila
NIM. 2211011220002

ABSTRAK

PERBANDINGAN KUALITAS KLASTERISASI METODE K-MEANS DAN X-MEANS PADA DATA KETAHANAN PANGAN KABUPATEN/KOTA DI KALIMANTAN SELATAN (Oleh: *Arsella Putri Nabila*; Pembimbing: *Thresye & Hermei Lissa*; 2026; 88 Halaman)

Ketahanan pangan merupakan isu strategis dalam mendukung pembangunan berkelanjutan, khususnya di Provinsi Kalimantan Selatan yang masih menghadapi tantangan dalam pemerataan distribusi pangan. Salah satu indikator yang digunakan untuk menilai kondisi tersebut adalah Indeks Ketahanan Pangan (IKP) yang mencakup aspek ketersediaan, akses, dan pemanfaatan pangan. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Kalimantan Selatan berdasarkan indikator ketahanan pangan serta membandingkan kualitas klasterisasi menggunakan metode K-Means dan X-Means. Data yang digunakan merupakan data sekunder dari Badan Perencanaan Pembangunan Nasional dan Badan Pusat Statistik Kalimantan Selatan. Tahapan penelitian meliputi preprocessing data dan klasterisasi data. Metode K-Means digunakan untuk melakukan pengelompokan data, sedangkan metode X-Means digunakan untuk melakukan pengelompokan data sekaligus menentukan jumlah klaster optimal menggunakan pendekatan *Bayesian Information Criterion* (BIC). Penelitian ini menghasilkan klaster optimal yaitu lima klaster. Evaluasi menggunakan Silhouette Coefficient menunjukkan nilai sebesar 0,1475 untuk K-Means dan 0,1513 untuk X-Means. Hasil tersebut menunjukkan bahwa metode X-Means memiliki kualitas klaster yang lebih baik dibandingkan metode K-Means dalam mengelompokkan kabupaten/kota berdasarkan indikator ketahanan pangan.

Kata Kunci: K-Means, X-Means, klasterisasi, ketahanan pangan

ABSTRACT

COMPARISON OF CLUSTERING QUALITY USING K-MEANS AND X-MEANS METHODS ON FOOD SECURITY DATA OF REGENCY/CITY IN SOUTH KALIMANTAN (Author: Arsella Putri Nabila; Supervisors: Thresye & Hermei Lissa; 2026; 88 pages)

Food security is a strategic issue in supporting sustainable development, particularly in South Kalimantan Province, which still faces challenges in the equitable distribution of food. One indicator used to assess this condition is the Food Security Index (IKP), which includes aspects of food availability, access, and utilization. This study aims to cluster regencies/cities in South Kalimantan based on food security indicators and to compare the clustering quality using the K-Means and X-Means methods. The data used are secondary data obtained from the National Development Planning Agency and the Central Bureau of Statistics of South Kalimantan. The research stages include data preprocessing and data clustering. The K-Means method is used to perform data clustering, while the X-Means method is used to perform data clustering as well as determine the optimal number of clusters using the Bayesian Information Criterion (BIC) approach. The results show that the optimal number of clusters is five clusters. Evaluation using the Silhouette Coefficient shows a value of 0.1475 for K-Means and 0.1513 for X-Means. These results indicate that the X-Means method has better clustering quality than the K-Means method in grouping regencies/cities based on food security indicators.

Keywords: K-Means, X-Means, clustering, food security

PRAKATA

Puji dan syukur penulis ucapkan kepada Tuhan Yang Maha Esa karena telah memberikan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Skripsi ini dengan judul “PERBANDINGAN KUALITAS KLASTERISASI METODE K-MEANS DAN X-MEANS PADA DATA KETAHANAN PANGAN KABUPATEN/KOTA DI KALIMANTAN SELATAN”. Penyusunan skripsi ini bertujuan untuk memenuhi salah satu persyaratan dalam rangka menyelesaikan program sarjana Strata-1 Matematika di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat.

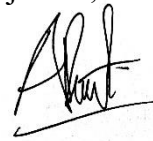
Penulis menyadari bahwa dalam penyusunan skripsi ini tidak terlepas dari bantuan, dukungan, kerja sama, maupun bimbingan dari berbagai pihak. Pada kesempatan ini, penulis menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat Banjarbaru.
2. Ibu Dr. Nai'mah Hijriati, S.Si., M.Si., selaku Ketua Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat Banjarbaru.
3. Bapak Akhmad Yusuf, S.Si., M.Kom., dan Bapak Muhammad Afief Balya, S.Mat., M.Si., selaku dosen pembimbing akademik yang selalu memberikan arahan selama masa perkuliahan.
4. Ibu Thresye, S.Si., M.Si. dan Ibu Hermei Lissa, S.Pd., M.Si., selaku dosen pembimbing utama dan pendamping skripsi yang telah memberikan bantuan, bimbingan, serta motivasi yang sangat berarti dalam pelaksanaan penelitian skripsi ini.
5. Bapak Oni Soesanto, S.Si., M.Si. dan Bapak Saman Abdurrahman, S.Si., M.Sc., selaku dosen penguji 1 dan penguji 2 yang telah memberikan koreksi dan saran sehingga penulisan penelitian skripsi ini dapat menjadi lebih baik.
6. Seluruh dosen dan staf Jurusan Matematika FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
7. Orang tua, adik, nenek, dan alm. kakek yang selalu memberikan dukungan, motivasi, kasih sayang, pengertian, serta doa.

8. Seluruh keluarga penulis dimanapun berada yang selalu mendukung dan mendoakan.
9. Sahabat penulis yaitu Annisa Kamilah dan Syintia Amanda yang selalu mengingatkan dalam berbagai hal serta selalu mendoakan selama masa perkuliahan.
10. Seluruh rekan mahasiswa Jurusan Matematika FMIPA ULM, khususnya teman-teman Matrix angkatan 2022 yang telah memberikan dukungan, bantuan dan motivasi kepada penulis selama penyusunan skripsi.

Semoga segala kebaikan yang diberikan mendapatkan balasan yang setimpal dari Allah Subhanahu Wa Ta'ala. Penulis menyadari bahwa masih terdapat kekurangan, ketidaktepatan, serta kesalahan dalam penulisan skripsi ini. Oleh karena itu, penulis dengan terbuka menerima kritik dan saran yang membangun. Semoga hasil penulisan tugas akhir ini dapat memberikan manfaat bagi para pembaca maupun pihak-pihak terkait.

Banjarbaru, 15 April 2026



Arsella Putri Nabila

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN	ii
PERNYATAAN.....	iii
ABSTRAK	iv
PRAKATA	vi
DAFTAR ISI.....	viii
ARTI LAMBANG DAN SINGKATAN.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Tujuan Penelitian	3
1.3 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Data Mining.....	5
2.1.1 Pengertian Data Mining.....	5
2.1.2 Teknik Data Mining	5
2.2 Metode Klasterisasi	7
2.2.1 Klasterisasi (<i>Clustering</i>)	7
2.2.2 Jenis-Jenis Klasterisasi (<i>Clustering</i>)	8
2.2.3 Metode K-Means	10
2.3 Metode X-Means	11
2.4 Metode Z-Score Scaling	12
2.5 Metode Interquartile Range (IQR)	13
2.6 Metode Winsorisasi	14
2.7 Silhouette Coefficient.....	15
2.8 Ketahanan Pangan	17
BAB III METODE PENELITIAN	21
3.1 Bahan Penelitian	21
3.2 Prosedur Penelitian	23

BAB IV HASIL DAN PEMBAHASAN	26
4.1 Proses Klasterisasi Menggunakan Metode K-Means dan X-Means.....	26
4.1.1 Tahap <i>Preprocessing</i> Data	26
4.1.2 Penerapan Metode K-Means	35
4.1.3 Penerapan Metode X-Means	39
4.2 Penentuan Jumlah Klaster Optimal Menggunakan X-Means (BIC) ...	41
4.3 Perbandingan Kualitas Klasterisasi K-Means dan X-Means	
Berdasarkan Nilai <i>Silhouette Coefficient</i>.....	48
4.3.1 Perhitungan Nilai <i>Silhouette Coefficient</i> pada Metode K-Means...	48
4.3.2 Perhitungan Nilai <i>Silhouette Coefficient</i> pada Metode X-Means..	51
4.3.3 Perbandingan Hasil <i>Silhouette Coefficient</i> Metode K-Means dan X-Means	53
4.3.4 Deskripsi Karakteristik Klaster	54
BAB V PENUTUP.....	57
5.1 Kesimpulan.....	57
5.2 Saran	57
DAFTAR PUSTAKA	59
LAMPIRAN	

ARTI LAMBANG DAN SINGKATAN

D_{ij}	: Jarak antara i dan j
x_i	: Koordinat x objek
x_j	: Koordinat x pusat
y_i	: Koordinat y objek
y_j	: Koordinat y pusat
x_1	: Nilai <i>record</i> data ke 1
x_2	: Nilai <i>record</i> data ke 2
$\sum x$	: Jumlah <i>record</i> data
n	: Jumlah data yang diamati (ukuran sampel)
L	: Estimasi likelihood
k	: Jumlah <i>cluster</i>
p	: Jumlah parameter
d	: Jumlah atribut
c_j	: <i>Centroid</i> ke- j yang sudah terpilih
SSE	: <i>Sum of Square Error</i> untuk semua objek dalam kumpulan data
$x \in C_i$	: Nilai keanggotaan titik data x ke pusat kelompok
C_i	: Pusat <i>cluster</i> ke- i
$D(x, C_i)$	: Jarak dari titik x ke kelompok C_i yang diikuti
$a(i)$	: Jarak rata-rata antara sampel i dan semua sampel lainnya dalam klaster yang sama
$b(i)$	: Jarak rata-rata minimum dari sampel i ke sampel-sampel dalam semua klaster lainnya
Q1	: Nilai di bawah 25% data
Q3	: Nilai di bawah 75% data
IQR	: <i>Interquartile Range</i>
BIC	: <i>Bayesian Information Criterion</i>

x : nilai data asli
 μ rata-rata (*mean*) dari seluruh data
 σ : simpangan baku (standar deviasi)

DAFTAR TABEL

Tabel 3.1 Indikator Ketahanan Pangan.....	21
Tabel 3.2 Dataset 8 Indikator Ketahanan Pangan Tahun 2023	22
Tabel 4.1 Data indikator 2	27
Tabel 4.2 Data Outlier	28
Tabel 4.3 Data setelah penanganan outlier	30
Tabel 4.4 Data indikator 1	31
Tabel 4.5 Data setelah normalisasi	32
Tabel 4.6 Hasil Perhitungan Seluruh Data.....	34
Tabel 4.7 Nilai Mean dan Standar Deviasi	35
Tabel 4.8 Nilai centroid awal K-Means untuk $k=5$	36
Tabel 4.9 Hasil perhitungan jarak euclidean K-Means.....	37
Tabel 4.10 Hasil klasterisasi K-Means	38
Tabel 4.11 Nilai centroid awal X-Means untuk $k=2$	39
Tabel 4.12 Hasil perhitungan jarak euclidean X-Means.....	39
Tabel 4.13 Hasil klasterisasi X-Means $K=2$	40
Tabel 4.14 Anggota klaster 1	41
Tabel 4.15 Nilai centroid awal split klaster 1	42
Tabel 4.16 Anggota klaster 1.1	42
Tabel 4.17 Anggota klaster 1.2	43
Tabel 4.18 Nilai centroid awal split klaster 1.1	43
Tabel 4.19 Anggota klaster 1.1.1	44
Tabel 4.20 Anggota klaster 1.1.2	44
Tabel 4.21 Nilai centroid awal split klaster 1.1.2	45
Tabel 4.22 Anggota klaster 1.1.2.1	45
Tabel 4.23 Anggota klaster 1.1.2.2	45
Tabel 4.24 Nilai centroid awal split klaster 1.2	46
Tabel 4.25 Anggota klaster 1.2.1	46
Tabel 4.26 Anggota klaster 1.2.2	46
Tabel 4.27 Anggota klaster 2	47
Tabel 4. 28 Nilai centroid awal split klaster 2	47
Tabel 4. 29 Anggota klaster 2.1	47
Tabel 4. 30 Anggota klaster 2.2	48
Tabel 4.31 Nilai Silhouette Coefficient K-Means untuk 5 Klaster.....	50
Tabel 4.32 Nilai Silhouette Coefficient X-Means untuk 5 Klaster.....	52
Tabel 4.33 Perbandingan Nilai Silhouette Coefficient	53
Tabel 4.34 Rata-Rata Indikator Data Awal	55
Tabel 4.35 Nilai Centroid Klaster Optimal.....	55
Tabel 4.36 Jarak Euclidean.....	55
Tabel 4.37 Permisalan Kategori Berdasarkan Urutan Klaster	56

DAFTAR GAMBAR

Gambar 3.1 Diagram Alir Penelitian	25
Gambar 4.1 Visualisasi Boxplot Data Sebelum Penanganan Outlier	29
Gambar 4.2 Visualisasi Boxplot Data Setelah Penanganan Outlier	31