

**PENERAPAN ALGORITMA *K-MEANS CLUSTERING*  
DALAM PENGELOMPOKAN KABUPATEN/KOTA  
BERDASARKAN INDIKATOR KEMISKINAN DI  
PROVINSI PAPUA**

**SKRIPSI SARJANA SISTEM INFORMASI**

Oleh :

Yusriana Chusna Fadilah

207006516056



**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNOLOGI KOMUNIKASI DAN  
INFORMATIKA  
UNIVERSITAS NASIONAL**

**2024**

**PENERAPAN ALGORITMA *K-MEANS CLUSTERING*  
DALAM PENGELOMPOKAN KABUPATEN/KOTA  
BERDASARKAN INDIKATOR KEMISKINAN DI  
PROVINSI PAPUA**

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar  
Sarjana Sistem Informasi dari Fakultas Teknologi Komunikasi dan Informatika

Oleh :

Yusriana Chusna Fadilah

207006516056



**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNOLOGI KOMUNIKASI DAN  
INFORMATIKA  
UNIVERSITAS NASIONAL**

**2024**

# HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

TUGAS AKHIR

**PENERAPAN ALGORITMA *K-MEANS CLUSTERING*  
DALAM PENGELOMPOKKAN KABUPATEN/KOTA  
BERDASARKAN INDIKATOR KEMISKINAN DI PROVINSI**



Dosen Pembimbing I

Ir. Asrul Sani, M.Kom., M.T., Ph.D

NIDN. 0303067003

## HALAMAN KEASLIAN TUGAS AKHIR

### PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

**PENERAPAN ALGORITMA *K-MEANS CLUSTERING* DALAM  
PENGELOMPOKKAN KABUPATEN/KOTA BERDASARKAN INDIKATOR  
KEMISKINAN DI PROVINSI PAPUA**

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 26 Agustus 2024



Yusrina Chusna Fadilah

NPM : 207006516056

## LEMBAR PERSETUJUAN TUGAS AKHIR

### LEMBAR PERSETUJUAN TUGAS AKHIR

Tugas Akhir dengan judul :

**PENERAPAN ALGORITMA *K-MEANS CLUSTERING* DALAM  
PENGELOMPOKKAN KABUPATEN/KOTA BERDASARKAN INDIKATOR  
KEMISKINAN DI PROVINSI PAPUA**

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Akhir Semester Genap 2023-2024 pada tanggal 22 Agustus 2024

**Dosen Pembimbing I**



Ir. Asrul Sani., M.Kom., M.T., Ph.D

NIDN. 0303067003

**Ketua Program Studi**



Dr. Andrianingsih., S.Kom., MMSI.

NIDN. 0303097902

**HALAMAN PENGESAHAN**  
**TUGAS SARJANA**  
**PENERAPAN ALGORITMA *K-MEANS CLUSTERING* DALAM**  
**PENGELOMPOKAN KABUPATEN/KOTA BERDASARKAN**  
**INDIKATOR KEMISKINAN DI PROVINSI PAPUA**



Oleh:  
**Yusriana Chusna Fadilah**  
207006516056

Program Studi Sistem Informasi  
Fakultas Teknologi Komunikasi dan Informatika

Disetujui pada tanggal: 26 Agustus 2024

Pembimbing 1

**Ir. Asrul Sani, M.Kom., M.T., Ph.D**  
NIDN. 0303067003

Penguji 1



**Dr. Andrianingsih, S.Kom., M.M.S.I**  
NIDN. 0303097902

Penguji 2



**Agus Iskandar, S.Kom., M.Kom.**  
NIDN. 0310087503

**LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI**

Nama : Yusriana Chusna Fadilah  
NPM : 207006516056  
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika  
Program Studi : Sistem Informasi  
Tanggal Sidang : 22 Agustus 2024

JUDUL DALAM BAHASA INDONESIA :

**PENERAPAN ALGORITMA *K-MEANS CLUSTERING* DALAM  
PENGELOMPOKAN KABUPATEN/KOTA BERDASARKAN  
INDIKATOR KEMISKINAN DI PROVINSI PAPUA**

JUDUL DALAM BAHASA INGGRIS :

**APPLICATION OF *K-MEANS CLUSTERING* FOR GROUPING  
DISTRICT/CITIES BASED ON POVERTY INDICATORS IN PAPUA  
PROVINCE**

<b>TANDA TANGAN DAN TANGGAL</b>		
<b>Pembimbing 1</b>	<b>Ko. Prodi</b>	<b>Mahasiswa</b>
TGL : 26-08-2024	TGL : 26-08-2024	TGL : 26-08-2024
		

## KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT, yang senantiasa memberikan rahmat dan petunjuk-Nya sehingga penulis dapat menyelesaikan perjalanan akademis hingga tahap penyusunan skripsi ini yang berjudul **"PENERAPAN ALGORITMA K-MEANS CLUSTERING DALAM PENGELOMPOKAN KABUPATEN/KOTA BERDASARKAN INDIKATOR KEMISKINAN DI PROVINSI PAPUA"**. Skripsi ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional.

Dalam proses penyusunan skripsi ini, saya merasa sangat bersyukur atas dukungan, bimbingan, dan motivasi yang saya terima selama perjalanan akademis saya. Tanpa bantuan dan arahan yang diberikan oleh dosen pembimbing, serta dorongan yang saya terima dari keluarga, teman-teman, dan rekan-rekan sejawat, pencapaian ini tidak akan menjadi mungkin. Oleh karena itu, izinkanlah saya untuk mengucapkan terima kasih sebesar-besarnya kepada:

1. Allah SWT yang telah memberikan kesehatan serta umur yang panjang sehingga penulis dapat mengemban ilmu di Universitas Nasional.
2. Dr. Andrianingsih, S.Kom., M.M.S.I selaku Ketua Program Studi Sistem Informasi Universitas Nasional
3. Ir. Asrul Sani, M.Kom., M.T., Ph.D selaku dosen pembimbing tunggal skripsi penulis.
4. Kedua orang tua penulis yang tercinta, Ayahanda yang telah tiada dan Ibu yang selalu ada. Terima kasih atas semua perjuangan, kasih sayang, perhatian, dan kepedulian yang tiada henti. Ayahanda, meskipun sudah berpulang, setiap doa dan pengorbananmu tetap hidup dalam ingatan dan hati penulis. Ibu, dukungan dan cinta Ibu selama ini telah menjadi sumber kekuatan dan semangat penulis dalam menyelesaikan skripsi ini.
5. Para kakak dan adik selaku keluarga yang selalu memberi dukungan, motivasi, arahan, nasihat, dan bantuan yang tiada henti. Kalian telah



memenuhi berbagai kebutuhan penulis dan memberikan semangat yang luar biasa selama proses penyusunan skripsi ini.

6. Saya juga ingin mengapresiasi diri sendiri atas kemampuan dan usaha yang telah saya berikan dalam menyelesaikan skripsi ini. Saya sangat bersyukur karena mampu mempertahankan semangat yang tinggi, berpikir positif, tidak pernah menyerah, dan selalu percaya pada diri sendiri. Saya juga menjaga **keteraturan dalam ibadah dan doa. Semua ini** merupakan bukti nyata bahwa saya berhak meraih kesuksesan di masa depan.
7. Widi, Raka dan Adit sebagai **sahabat seperjuangan** dari awal perkuliahan hingga akhir dan selalu berusaha bersama dengan bertukar pikiran, memberi semangat dan motivasi serta masukan dan saran agar dapat menyelesaikan skripsi.
8. Semua pihak yang tidak dapat saya sebutkan satu per satu, saya mengucapkan terima kasih atas bantuan yang diberikan, baik secara langsung maupun tidak langsung, sejak awal perkuliahan hingga penyelesaian skripsi ini.

Penulis berharap skripsi ini dapat memberikan nilai tambah bagi pembaca. Penulis juga menyadari bahwa masih ada ruang untuk perbaikan. Oleh karena itu, masukan dan saran yang konstruktif akan sangat dihargai.

Jakarta, 13 Agustus 2024



Yusriana Chusna Fadilah

## ABSTRAK

Kemiskinan adalah kondisi di mana individu atau wilayah tidak mampu mencapai tingkat hidup yang layak atau meningkatkan standar kehidupan mereka. Berdasarkan data Badan Pusat Statistik tahun 2023, Provinsi Papua memiliki tingkat kemiskinan mencapai 26,03% dan populasi miskin sebanyak 915,15 ribu jiwa, menghadapi tantangan serius dalam mengatasi masalah ini. Penelitian ini bertujuan untuk mengidentifikasi pola kemiskinan di Kabupaten/Kota Provinsi Papua menggunakan algoritma K-Means Clustering untuk mengelompokkan wilayah berdasarkan karakteristik kemiskinan yang serupa. Hasil penelitian menunjukkan bahwa algoritma K-Means Clustering berhasil mengelompokkan Kabupaten/Kota ke dalam dua cluster, yaitu C0 dan C1, dengan menggunakan metode Elbow dan Silhouette Score untuk menentukan jumlah cluster optimal. Cluster C0, yang termasuk dalam kategori tingkat kemiskinan tinggi, terdiri dari 17 Kabupaten/Kota memiliki karakteristik persentase penduduk miskin tinggi, garis kemiskinan tinggi, rata-rata lama sekolah rendah, Indeks Pembangunan Manusia (IPM) rendah, serta Indeks Kedalaman Kemiskinan (P1) dan Indeks Keparahan Kemiskinan (P2) tinggi, dengan jumlah pengangguran dan pengeluaran per kapita yang rendah. Sebaliknya, cluster C1, termasuk dalam kategori tingkat kemiskinan rendah, terdiri dari 12 Kabupaten/Kota memiliki karakteristik persentase penduduk miskin rendah, garis kemiskinan rendah, rata-rata lama sekolah tinggi, IPM tinggi, serta P1 dan P2 rendah, dengan jumlah pengangguran dan pengeluaran per kapita yang tinggi. Rekomendasi strategis untuk C0 mencakup implementasi Bantuan Tunai Langsung (BLT), pembangunan infrastruktur dasar, perbaikan pendidikan, dan peningkatan IPM melalui Pusat Kesehatan dan klinik bergerak. Untuk C1, strategi meliputi penguatan ekonomi lokal, insentif pajak dan subsidi, pengembangan kurikulum keterampilan, dan inovasi dalam kesehatan dan pendidikan. Penelitian ini bertujuan memberikan landasan strategis bagi pemerintah dalam merancang kebijakan pembangunan yang lebih fokus di setiap Kabupaten/Kota di Provinsi Papua, diharapkan dapat membantu mengatasi kemiskinan secara lebih efektif dan mendukung pembangunan berkelanjutan di wilayah tersebut.

**Kata Kunci:** *K-Means Clustering, Kemiskinan, Provinsi Papua, Metode Elbow, Metode Silhouette.*

## ABSTRACT

Poverty is a condition where individuals or regions are unable to achieve a decent standard of living or improve their quality of life. According to data from the Central Statistics Agency (BPS) in 2023, Papua Province has a poverty rate of 26.03% and a poor population of 915,150 people, facing serious challenges in addressing this issue. This study aims to identify poverty patterns in districts/cities in Papua Province using the K-Means Clustering algorithm to group areas based on similar poverty characteristics. The study finds that the K-Means Clustering algorithm successfully grouped districts/cities into two clusters, C0 and C1, using the Elbow and Silhouette Score methods to determine the optimal number of clusters. Cluster C0, categorized as high poverty, consists of 17 districts/cities with characteristics of high poverty percentage, high poverty line, low average years of schooling, low Human Development Index (HDI), and high Poverty Depth Index (P1) and Poverty Severity Index (P2), with low unemployment and per capita expenditure. In contrast, Cluster C1, categorized as low poverty, consists of 12 districts/cities with characteristics of low poverty percentage, low poverty line, high average years of schooling, high HDI, and low P1 and P2, with high unemployment and per capita expenditure. Strategic recommendations for C0 include implementing Direct Cash Assistance (BLT), developing basic infrastructure, improving education, and enhancing HDI through health centers and mobile clinics. For C1, strategies include strengthening the local economy, providing tax incentives and subsidies, developing skill curricula, and innovating in health and education. This study aims to provide a strategic foundation for the government in designing more focused development policies for each district/city in Papua Province, with the hope of effectively addressing poverty and supporting sustainable development in the region.

**Keyword: K-Means Clustering, Poverty, Papua Province, Elbow Method, Silhouette Method.**

## DAFTAR ISI

HALAMAN PENGESAHAN.....	ii
KATA PENGANTAR.....	vii
ABSTRAK.....	ix
ABSTRACT.....	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah.....	6
1.3 Tujuan Penelitian.....	7
1.4 Batasan Masalah.....	7
1.5 Manfaat Penelitian.....	7
1.6 Sistematika Penulisan.....	8
BAB II TINJAUAN PUSTAKA.....	9
2.1 Penelitian Terdahulu.....	9
2.2 Studi Literatur.....	16
2.3 Landasan Teori.....	16
2.3.1 Kemiskinan.....	16
2.3.2 <i>Data Mining</i> .....	16
2.3.3 <i>Clustering</i> .....	17
2.3.4 Algoritma K-Means.....	17
2.3.5 Metode <i>Elbow</i> .....	18
2.3.6 Metode <i>Silhouette</i> .....	18
2.3.7 <i>Python</i> .....	18
2.3.8 <i>Google Colab</i> .....	19
2.3.9 <i>Streamlit</i> .....	19
BAB III METODOLOGI PENELITIAN.....	20

3.1 <i>View Data</i> .....	21
3.2 <i>Data Understanding</i> .....	21
3.3 <i>Data Preparation</i> .....	22
3.3.1 <i>Data Preprocessing</i> .....	22
3.3.2 <i>Data Transformation</i> .....	22
3.4 <i>Modeling</i> .....	22
3.5 <i>Evaluation</i> .....	23
3.6 <i>Deployment</i> .....	24
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	<b>25</b>
4.1 Implementasi Tahapan Penelitian .....	25
4.1.1 <i>View Data</i> .....	25
4.1.2 <i>Data Understanding</i> .....	26
4.1.3 <i>Data Preparation</i> .....	28
4.1.3.1 <i>Data Preprocessing</i> .....	28
4.1.3.2 <i>Data Transformation</i> .....	30
4.1.4 <i>Modeling</i> .....	32
4.1.5 <i>Evaluation</i> .....	32
4.1.6 <i>Deployment</i> .....	34
4.2 Perhitungan Manual <i>K-Means</i> .....	36
4.3 Analisa Hasil .....	42
4.3.1 Visualisasi .....	42
4.3.2 Karakteristik .....	54
4.3.3 Strategi .....	56
<b>BAB V KESIMPULAN DAN SARAN</b> .....	<b>62</b>
5.1 Kesimpulan .....	62
5.2 Saran .....	64
<b>DAFTAR PUSTAKA</b> .....	<b>xi</b>
<b>LAMPIRAN</b> .....	<b>xiv</b>

## DAFTAR GAMBAR

Gambar 1. 1 Peta Kemiskinan Papua .....	3
Gambar 3. 1 Tahapan Penelitian .....	20
Gambar 4.1 1 Data Kemiskinan BPS 2022 .....	25
Gambar 4.1 2 Data Kemiskinan BPS 2023 .....	26
Gambar 4.1 3 Dataset dalam File CSV .....	26
Gambar 4.1 4 Import Dataset Kemiskinan .....	27
Gambar 4.1 5 Sebelum Label Encoding .....	29
Gambar 4.1 6 Elbow Method .....	32
Gambar 4.1 7 Silhouette Score .....	33
Gambar 4.1 8 Halaman Data Overview .....	34
Gambar 4.1 9 Halaman Clustering .....	35
Gambar 4.1 10 Halaman Visualisasi .....	35
Gambar 4.1 11 Halaman Strategi .....	36
Gambar 4.3 1 Histogram Distribusi Kemiskinan .....	43
Gambar 4.3 2 Scatter Plot: Persentase Penduduk Miskin .....	47
Gambar 4.3 3 Scatter Plot: Rata-Rata Lama Sekolah .....	49
Gambar 4.3 4 Scatter Plot: Indeks Kedalaman Kemiskinan .....	50
Gambar 4.3 5 Scatter Plot: Jumlah Pengangguran .....	51



## DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu .....	9
Tabel 3. 1 Indikator Kemiskinan .....	21
Tabel 4.1 1 Tipe data dalam Dataset .....	27
Tabel 4.1 2 Missing Value .....	28
Tabel 4.1 3 Data Duplicate .....	29
Tabel 4.1 4 Setelah Label Encoding .....	30
Tabel 4.1 5 Data Transformation .....	31
Tabel 4.1 6 Silhouette Score .....	33
Tabel 4.2 1 Centroid Iterasi 1 .....	36
Tabel 4.2 2 Jarak Terdekat Iterasi 1 .....	37
Tabel 4.2 3 Euclidean Distance Iterasi 1 .....	38
Tabel 4.2 4 Centroid Iterasi 2 .....	40
Tabel 4.2 5 Euclidean Distance Iterasi 2 .....	40
Tabel 4.2 6 Anggota Cluster .....	42
Tabel 4.3 1 Karakteristik Cluster .....	55
Tabel 4.3 2 Detail Karakteristik Cluster .....	55
Tabel 4.3 3 Rekomendasi Strategi Tingkat Kemiskinan Tinggi .....	57
Tabel 4.3 4 Rekomendasi Strategi Tingkat Kemiskinan Rendah .....	59

