

**ANALISA KLASSTER DATA PENDUDUK MISKIN
MENURUT PROVINSI DENGAN METODE K-MEANS**

SKRIPSI SARJANA SISTEM INFORMASI

Oleh :

Indah Tri Nur Azizah

207006516083



**PROGRAM STUDI INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
UNIVERSITAS NASIONAL
2023**

**ANALISA KLASSTER DATA PENDUDUK MISKIN
MENURUT PROVINSI DENGAN METODE K-MEANS**

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

Oleh:

Indah Tri Nur Azizah

207006516083



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
UNIVERSITAS NASIONAL
2023/2024**

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN
TUGAS AKHIR

ANALISIS PENGAMBILAN KEPUTUSAN PEMILIHAN KENDARAAN BERMOBIL



Dosen Pembimbing 1

(Panca Dewi Pamungkasari, S.T., M.T., Ph.D)

NIDN: 0305077903

Dosen Pembimbing 2

(Novi Dian Nathasia, S.Kom., MMSi.)

NIDN: 0714127701

HALAMAN KEASLIAN TUGAS AKHIR

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul:

ANALISA KLASTER DATA PENDUDUK MISKIN MENURUT PROVINSI DENGAN METODE K-MEANS

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, 27 Agustus 2024



Indah Tri Nur Azizah

NPM: 207006516083

LEMBAR PERSETUJUAN TUGAS AKHIR

LEMBAR PERSETUJUAN TUGAS AKHIR

Tugas Akhir dengan judul :

ANALISA KLASTER DATA PENDUDUK MISKIN MENURUT PROVINSI DENGAN METODE K-MEANS

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

Pembimbing 1


(Panca Dewi Pamungkasari), S.T, M.T, Ph.
NIDN 0305077903

Ketua Program Studi


(Dr. Andrianingsih, S. Kom., MMSI.)
NIDN. 0303097902

LEMBAR PERSETUJUAN JUDUL PEMBIMBING 1
LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Indah Tri Nur Azizah
 NPM : 207006516083
 Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
 Program Studi : Sistem Informasi
 Tanggal Sidang : 22 Agustus 2024

JUDUL DALAM BAHASA INDONESIA :

**ANALISA KLASTER DATA PENDUDUK MISKIN MENURUT
 PROVINSI DENGAN METODE K-MEANS**

JUDUL DALAM BAHASA INGGRIS :

**CLUSTER ANALYSIS OF POOR POPULATION DATA BY
 PROVINCE USING THE K-MEANS METHOD**

TANDA TANGAN DAN TANGGAL

Pembimbing 1	Ka. Prodi	Mahasiswa
TGL :	TGL :	TGL :
	 	

LEMBAR PERSETUJUAN JUDUL PEMBIMBING 2
LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Indah Tri Nur Azizah
 NPM : 207006516083
 Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
 Program Studi : Sistem Informasi
 Tanggal Sidang : 22 Agustus 2024

JUDUL DALAM BAHASA INDONESIA :

**ANALISA KLASTER DATA PENDUDUK MISKIN MENURUT
 PROVINSI DENGAN METODE K-MEANS**

JUDUL DALAM BAHASA INGGRIS :

**CLUSTER ANALYSIS OF POOR POPULATION DATA BY
 PROVINCE USING THE K-MEANS METHOD**

TANDA TANGAN DAN TANGGAL

Pembimbing 2	Ka-Prodi	Mahasiswa
TGL :	TGL :	TGL :
		

HALAMAN PENGESAHAN
TUGAS SARJANA
ANALISA KLASSTER DATA PENDUDUK MISKIN MENURUT
PROVINSI DENGAN METODE K-MEANS

Oleh

Indah Tri Nur Azizah

207006516083


Program Studi Sistem Informasi
Fakultas Teknologi Komunikasi dan Informatika

Disetujui pada Tanggal 27 Agustus 2024

Pembimbing 1

Pembimbing 2


(Panca Dewi Pamungkasari, S.T., M.T., Ph.D)
NIDN 0305077903


(Novi Dian Nathasia, S.Kom., MMSI)
NIDN. 0714127701

Penguji 1

Penguji 2


(Dr. Andrianingsih, S. Kom., MMSI.)
NIDN. 0303097902


(Ir. Endah Tri Esti Handayani, MMSI)
NIDN. 0314106802

KATA PENGANTAR

Dengan menyebut nama Allah SWT yang Maha Pengasih lagi Maha Penyayang, saya panjatkan puja dan puji syukur atas kehadiran-Nya, yang telah melimpahkan rahmat, hidayah, dan inayah-Nya kepada saya, sehingga saya dapat menyelesaikan Proposal Skripsi yang diberi judul dengan “ANALISA KLASSTER DATA PENDUDUK MISKIN MENURUT PROVINSI DENGAN METODE K-MEANS”

Proposal tersebut telah saya susun dengan maksimal dan mendapatkan bantuan dari berbagai pihak serta berbagai sumber bacaan yang ada di internet sehingga dapat memperlancar pembuatan makalah ini. Untuk itu saya menyampaikan banyak terima kasih kepada semua pihak yang telah berkontribusi dalam pembuatan proposal ini.

Semoga proposal ini dapat memberikan pengetahuan yang lebih luas kepada pembaca. Walaupun proposal ini memiliki kelebihan dan kekurangan, penulis membutuhkan kritik dan saran dari pembaca yang membangun.

Terima kasih.

Jakarta, 2 September 2024

ABSTRAK

Penerapan metode K-Means dalam analisis data kemiskinan di Indonesia bertujuan untuk mendukung distribusi bantuan sosial yang lebih efektif. Menggunakan data dari Badan Pusat Statistik (BPS) mengenai jumlah individu miskin dari tahun 2019 hingga 2023, metode K-means diterapkan dengan tambahan Rank Order Centroid (ROC) untuk meningkatkan penentuan centroid awal. Evaluasi dilakukan dengan menggunakan *elbow method* untuk menentukan jumlah cluster yang optimal dan *silhouette score* untuk menilai kualitas cluster. Hasilnya menunjukkan bahwa metode K-Means secara langsung menyarankan jumlah $k=2$ sebagai yang optimal. Proses clustering mengidentifikasi 2 cluster utama yaitu cluster 1 yang mencakup 31 provinsi dengan tingkat kemiskinan tinggi dan cluster 2 yang terdiri dari 3 provinsi dengan tingkat kemiskinan lebih rendah. *Silhouette score* yang mendekati 1 menunjukkan bahwa cluster tersebut terdefinisi baik dan berkualitas tinggi. Temuan ini memungkinkan pemerintah untuk mengalokasikan bantuan sosial dengan lebih akurat, membantu dalam perencanaan dan pelaksanaan kebijakan pengentasan yang lebih efektif, sesuai dengan target RPJMN 2020-2024.

Kata Kunci: Klusterisasi, Kemiskinan, K-Means, Rank Order Centroid (ROC).

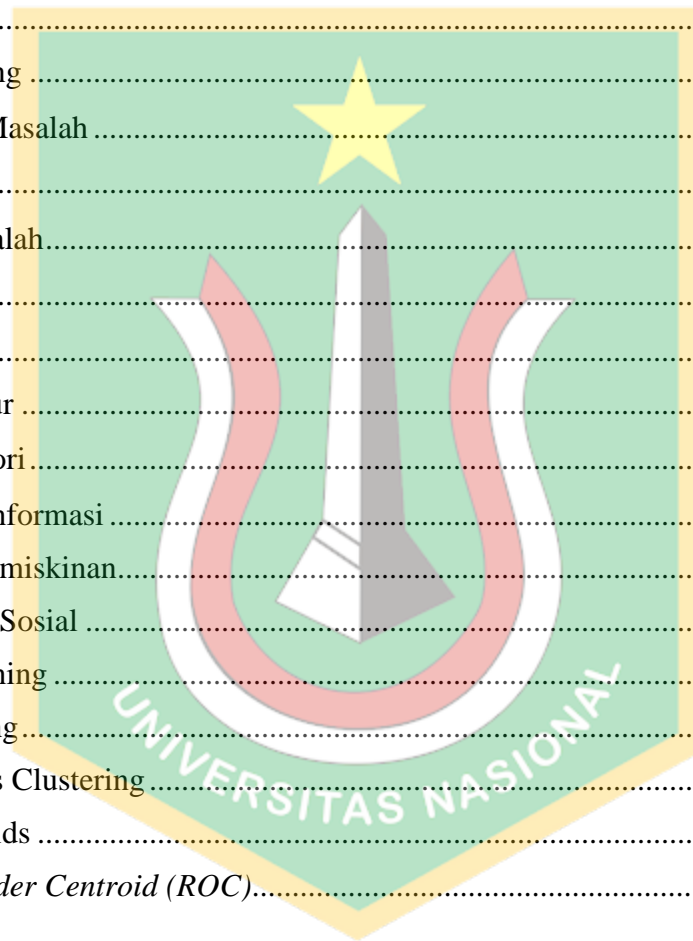
ABSTRACT

The application of the K-Means method in poverty data analysis in Indonesia aims to support more effective distribution of social assistance. Utilizing data from the Central Statistics Agency (BPS) on the number of impoverished individuals from 2019 to 2023, the K-Means method was applied with the addition of Rank Order Centroid (ROC) to improve the determination of initial centroids. Evaluation was carried out using the elbow method to determine the optimal number of clusters and silhouette score to assess cluster quality. The results indicate that the K-Means method directly suggested an optimal $k=2$. The clustering process identified two main clusters: Cluster 1, which includes 31 provinces with high poverty levels, and Cluster 2, consisting of 3 provinces with lower poverty levels. Silhouette scores approaching 1 indicate that the clusters are well-defined and of high quality. These findings enable the government to allocate social assistance more accurately, aiding in the planning and implementation of more effective poverty alleviation policies in line with the RPJMN 2020-2024 targets.

Keywords: *Clustering, Poverty, K-Means, Rank Order Centroid (ROC).*



HALAMAN PERNYATAAN ORISINALITAS	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR	iv
ABSTRAK.....	v
ABSTRACT.....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR.....	ix
DAFTAR TABEL.....	x
BAB I.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Tujuan.....	3
1.4 Batasan Masalah.....	3
1.5 Kontribusi.....	3
BAB II.....	4
2.1 Studi Literatur	4
2.2 Landasan Teori.....	8
2.2.1 Sistem Informasi	8
2.2.3 Garis Kemiskinan.....	9
2.2.4 Bantuan Sosial	9
2.2.5 Data Mining	9
2.2.6 Clustering.....	9
2.2.7 K-Means Clustering.....	10
2.2.8 K-Medoids	11
2.2.8 <i>Rank Order Centroid (ROC)</i>	11
2.2.9 Phyton	12
BAB III	13
3.1 Lokasi Penelitian.....	13
3.2 Waktu Penelitian	13
3.3 Penentuan Subjek Penelitian	13
3.4 Fokus Penelitian	14
3.5 Sumber data.....	14



3.6	Teknik Pengumpulan Data.....	14
3.7	Desain Penelitian.....	15
3.7.1	Tahap Penelitian.....	15
BAB IV	17
4.1	Business Understanding	17
4.2	Data Understanding.....	17
4.2.1	Scatter Data.....	18
4.3	Data Preparation.....	18
4.4	Modeling	20
4.4.1	Rank Order Centroid.....	20
4.5	Evaluation.....	21
4.5.1	Elbow Method.....	22
4.5.2	Silhouette Coefficient	23
4.6	Evaluation.....	24
4.6.1	Hasil Clustering	24
BAB V	27
5.1	Kesimpulan.....	27
5.2	Saran.....	27
DAFTAR PUSTAKA	29	



DAFTAR GAMBAR

Gambar 3.1 Crisp DM	15
Gambar 4.2.1 Scatter Data	18
Gambar 4.4.2 ROC	20
Gambar 4.5.1 Elbow Method	22
Gambar 4.5.2 Silhouette Coefficient	23

Gambar 4.6.1 Visualisasi 26



DAFTAR TABEL

Table 2 2 Jurnal Penelitian Terdahulu.....	4
Table 3 1 Waktu Penelitian	13
Tabel 4.3 1 Check Data Type	19

