

SKRIPSI

**PENGELOMPOKKAN KABUPATEN/KOTA DI JAWA
TIMUR BERDASARKAN INDIKATOR TENAGA KERJA
DENGAN METODE K-MEANS DAN K-MEDOIDS
CLUSTERING**



**DIKA MHARDIKA SARI
NIM. 212409005**

**PROGRAM STUDI STATISTIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS PGRI ADI BUANA SURABAYA
2023**

**PENGELOMPOKKAN KABUPATEN/KOTA DI
JAWA TIMUR BERDASARKAN INDIKATOR
TENAGA KERJA DENGAN METODE K-MEANS
DAN K-MEDOIDS CLUSTERING**

SKRIPSI

Diajukan Kepada Fakultas Sains dan Teknologi Universitas PGRI
Adi Buana Surabaya Untuk Memenuhi Sebagian Persyaratan
Dalam Memperoleh Gelar Sarjana Statistika Pada Program Studi
Statistika

**DIKA MHARDIKA SARI
NIM. 212409005**

**PROGRAM STUDI STATISTIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS PGRI ADI BUANA SURABAYA
2023**

LEMBAR PERSETUJUAN

Yang bertanda tangan di bawah ini menyatakan bahwa :

Nama : Dika Mhardika Sari
NIM : 212409005
Pembimbing : Artanti Indrasetianingsih, S. Si., M. Si
Judul Proposal : Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Jumlah Penduduk Berumur 15 Tahun Ke Atas Yang Bekerja dan Rata-Rata Pendapatan dengan Metode K-Means dan K-Medoids Clustering

Disetujui untuk diseminarkan pada seminar proposal skripsi di Program Studi Statistika Universitas PGRI Adi Buana Surabaya.

Surabaya, 10 Januari 2023

Pembimbing,



Artanti Indrasetianingsih, S. Si., M. Si.
NIP/NPP : 0609466 / DY


LEMBAR PENGESAHAN

Yang bertanda tangan di bawah ini menyatakan bahwa :

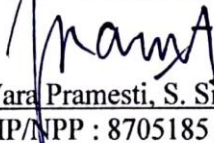
Nama : Dika Mhardika Sari
NIM : 212409005
Pembimbing : Artanti Indrasetianingsih, S. Si., M. Si
Judul Proposal : Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Jumlah Penduduk Berumur 15 Tahun Ke Atas Yang Bekerja dan Rata-Rata Pendapatan dengan Metode K-Means dan K-Medoids Clustering
Tanggal Ujian : 25 Januari 2023

Telah dinyatakan lulus ujian proposal dan proposal tersebut telah diperiksa, diperbaiki dan disetujui oleh tim penguji.

Disetujui Oleh Tim Penguji Skripsi:


Artanti Indrasetianingsih, S. Si., (Pembimbing)
M. Si
NIP/NPP : 0609466 / DY


Alfisyahrina Hapsery, M. Si (Pembimbing)
NPP/NIP. 1804856 / DY


Wara Pramesti, S. Si., M. Si (Penguji)
NIP/NPP : 8705185 / DY


Mengetahui,
Ketua Program Studi Statistika,
Alfisyahrina Hapsery, M. Si
NPP/NIP. 1804856 / DY


BERITA ACARA UJIAN SKRIPSI

Skripsi ini diterima dan disetujui oleh panitia ujian skripsi Sarjana Statistika Program Studi Statistika Fakultas Sains dan Teknologi Universitas PGRI Adi Buana Surabaya.

Pada Hari : Senin
Tanggal : 26 Januari
Tahun : 2023

Panitia Ujian Skripsi

1. Ketua Dekan :



Dr. Dian Karunia Binawati, M.Si.

2. Sekretaris :

Dr. Arif Yachya, S.Si., M.Si.

3. Anggota :

Artanti Indrasejaningsih, M.Si.

Alfisyahrina Hapsery, M.Si

Wara Pramesti, S. Si., M. Si

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan rahmat dan hidayat Nya sehingga Penulis dapat menyelesaikan proposal penelitian yang berjudul “Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Tenaga Kerja Dengan Metode K-Means dan K-Medoids Clustering”.

Proposal ini disusun dalam rangka memenuhi prasyarat memperoleh gelar Sarjana (S1) pada Program Studi Statistika Fakultas Sains dan Teknologi Universitas PGRI Adi Buana Surabaya.

Penulis berterima kasih kepada Ibu Artanti Indrasetianingsih, S. Si., M.Si selaku dosen pembimbing yang tanpa lelah memberikan dukungan dan bantuan sehingga Penulis dapat menyelesaikan proposal ini dengan baik pada waktu yang ditentukan.

Jika terdapat kesalahan pada proposal ini, Penulis terbuka untuk menerima kritik dan saran yang membangun demi kebaikan bersama. Harapannya semoga proposal penelitian ini dapat memberikan manfaat bagi semua. Terima kasih.

Surabaya, 12 Juli 2023



Dika Mhardika Sari

NIM. 212409005

HALAMAN PERSEMBAHAN

MOTTO

“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya”

(Q.S Al-Baqarah Ayat 286)

PERSEMBAHAN

Skripsi ini dipersembahkan untuk orang-orang yang telah memberikan bantuan, dukungan, do'a serta harapan dalam menyelesaikan skripsi ini:

1. Kedua orang tua tercinta ibu Irdina Larasanti, S.E dan bapak Teddy Achmadi, S.E
2. Saudara-saudara kandung Ninta Pininta Sari dan Muhammad Khundini
3. Dosen pembimbing 1 ibu Artanti Indrasetianingsih
4. Dosen penguji 1 Alfisyahrina Hapsery, M. Si
5. Dosen penguji 2 Wara Pramesti, S. Si., M. Si
6. Teman-teman Statistika angkatan 2021

SURAT PERNYATAAN

Yang bertanda tangan di bawah ini menyatakan bahwa :

Nama : Dika Mhardika Sari
NIM : 212409005
Pembimbing 1 : Artanti Indrasetyaningasih, M.Si
Judul Skripsi : Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Tenaga Kerja Dengan Metode K-Means dan K-Medoids Clustering

Menyatakan bahwa skripsi ini adalah hasil karya saya dan bukan merupakan duplikasi sebagian atau seluruhnya dari karya orang lain, kecuali bagian sumber informasinya dicantumkan.

Pernyataan ini dibuat sebenar – benarnya secara sadar dan bertanggung jawab dan saya bersedia menerima sanksi pembatalan skripsi apabila terbukti melakukan duplikasi terhadap skripsi atau karya ilmiah lain yang sudah ada.

Surabaya, 08 September 2023



Dika Mhardika Sari

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
BERITA ACARA UJIAN SKRIPSI.....	iv
KATA PENGANTAR.....	v
HALAMAN PERSEMBAHAN	vi
SURAT PERNYATAAN	vii
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiii
ABSTRAK	xiv
<i>ABSTRACT</i>	xv
BAB 1 PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Tujuan Penelitian	6
1.4 Manfaat Penelitian	6
1.5 Batasan Masalah	7
BAB 2 TINJAUAN PUSTAKA	
2.1 <i>Outlier</i>	8
2.2 Statistika Deskriptif secara Numerik	8
2.3 Statistika Deskriptif secara Visual	10
2.4 Analisis Faktor.....	11
2.5 Analisis <i>Cluster</i>	15
2.6 <i>Silhouette Index</i> dan <i>Silhouette Coefficient</i>	18
2.7 Uji Normalitas Univariat dan Multivariat	20
2.8 Indikator Tenaga Kerja	21
2.9 Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja.....	22
2.10 Rata-rata Pendapatan di Jawa Timur	23
BAB 3 METODOLOGI PENELITIAN	
3.1 Sumber Data	24
3.2 Variabel Penelitian	24
3.3 Langkah Analisis	27

DAFTAR ISI

	Halaman
BAB 4 ANALISIS DAN PEMBAHASAN	
4.1 Perkembangan Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Provinsi Jawa Timur pada Tahun 2020.....	30
4.2 Analisis Faktor dengan Data <i>Outlier</i> Terhadap Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur.....	45
4.3 Hasil <i>Clustering</i> dengan Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan 12 Variabel dengan Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur Tahun 2020 Menggunakan <i>K-Means</i> dan <i>K-Medoids Clustering</i>	48
4.4 Hasil <i>Clustering</i> Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Data Faktor 1,2,3 Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur Tahun 2020 Menggunakan <i>K-Means</i> and <i>K-Medoids Clustering</i>	54
4.5 Perbandingan Hasil <i>Clustering</i> Menggunakan <i>K-Means</i> dan <i>K-Medoids</i>	71
BAB 5 KESIMPULAN DAN SARAN	
5.1 Kesimpulan.....	76
5.2 Saran.....	77
DAFTAR PUSTAKA.....	78
LAMPIRAN	81

DAFTAR TABEL

	Halaman
Tabel 2.1 Interpretasi nilai MSA	12
Tabel 2.2 Interpretasi Nilai <i>Silhouette Coefficient</i> (SC)	20
Tabel 3.1 Variabel Penelitian	24
Tabel 3.2 Pembagian Wilayah Kab/Kota di Provinsi Jawa Timur	26
Tabel 3.3 Struktur Data Tahun 2020	26
Tabel 4.1 Deteksi <i>Outlier</i> Multivariat Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur Tahun 2020	31
Tabel 4.2 Deteksi <i>Outlier</i> Multivariat Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur Tahun 2020	45
Tabel 4.3 MSA Awal per Variabel Data Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur Tahun 2020 Dengan <i>Outlier</i>	46
Tabel 4.4 Total Varians per Jumlah Komponen (Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur Tahun 2020 Tanpa <i>Outlier</i>)	47
Tabel 4.5 <i>Rotated Component Matrix</i> (Jumlah Penduduk Beumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur Tahun 2020 Dengan <i>Outlier</i>)	47
Tabel 4.6 Hasil <i>Clustering K-Means</i> 12 Variabel	50
Tabel 4.7 Hasil <i>Clustering K-Medoids</i> 12 Variabel	52
Tabel 4.8 Perbandingan <i>Silhouette</i> 12 Variabel	53
Tabel 4.9 Hasil <i>Clustering K-Means</i> Faktor 1	56
Tabel 4.10 Hasil <i>Clustering K-Medoids</i> Faktor 1	58
Tabel 4.11 Perbandingan <i>Silhouette</i> Faktor 1	59
Tabel 4.12 Hasil <i>Clustering K-Means</i> Faktor 2	61
Tabel 4.13 Hasil <i>Clustering K-Medoids</i> Faktor 2	64
Tabel 4.14 Perbandingan <i>Silhouette</i> Faktor 2	65
Tabel 4.15 Hasil <i>Clustering K-Means</i> Faktor 3	67
Tabel 4.16 Hasil <i>Clustering K-Medoids</i> Faktor 3	69
Tabel 4.17 Perbandingan <i>Silhouette</i> Faktor 3	70
Tabel 4.18 Perbandingan <i>Silhouette</i> Keseluruhan	71
Tabel 4.19 Hasil Uji Distribusi Normal	73
Tabel 4.20 Deskriptif per <i>Cluster Optimum</i>	73
Tabel 4.21 Rangkuman Karakteristik per <i>Cluster Optimum</i>	74

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Diagram Alir Penelitian	29
Gambar 4.1 Diagram Profil Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan tahun 2020	32
Gambar 4.2 Peta Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja Selama Seminggu yang Lalu Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama 1	33
Gambar 4.3 Peta Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja Selama Seminggu yang Lalu Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama 2	34
Gambar 4.4 Peta Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja Selama Seminggu yang Lalu Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama 3	35
Gambar 4.5 Peta Rata-rata Upah/Gaji Bersih Sebulan Pekerja Formal Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama 2	36
Gambar 4.6 Peta Rata-rata Upah/Gaji Bersih Sebulan Pekerja Formal Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama 3	37
Gambar 4.7 Peta Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama 1	38
Gambar 4.8 Peta Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama 2	39
Gambar 4.9 Peta Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama 3	40
Gambar 4.10 Peta Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Pendidikan Tertinggi yang Ditamatkan (Tidak Pernah Sekolah/Belum Tamat SD)	41
Gambar 4.11 Peta Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Pendidikan Tertinggi yang Ditamatkan (Tamat SD/MI)	42
Gambar 4.12 Peta Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Pendidikan Tertinggi yang Ditamatkan (Tamat SMP/MTs).....	43
Gambar 4.13 Peta Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Pendidikan Tertinggi yang Ditamatkan (Tamat SMA ke Atas).....	44
Gambar 4.14 <i>Average Silhouette Width K-Means</i> 12 Variabel.....	49
Gambar 4.15 <i>Cluster Plot K-Means</i> 12 Variabel.....	50

Gambar 4.16 <i>Silhouette K-Medoids</i> 12 Variabel	51
Gambar 4.17 <i>Cluster Plot K-Medoids</i> 12 Variabel.....	52
Gambar 4.18 <i>Average Silhouette Width K-Means</i> Faktor 1	55
Gambar 4.19 <i>Cluster Plot K-Means</i> Faktor 1	56
Gambar 4.20 <i>Silhouette K-Medoids</i> Faktor 1	57
Gambar 4.21 <i>Cluster Plot K-Medoids</i> Faktor 1	58
Gambar 4.22 <i>Average Silhouette Width K-Means</i> Faktor 2.....	61
Gambar 4.23 <i>Cluster Plot K-Means</i> Faktor 2	61
Gambar 4.24 <i>Silhouette K-Medoids</i> Faktor 2	63
Gambar 4.25 <i>Cluster Plot K-Medoids</i> Faktor 2	63
Gambar 4.26 <i>Silhouette K-Means</i> Faktor 3.....	66
Gambar 4.27 <i>Cluster Plot K-Means</i> Faktor 2	66
Gambar 4.28 <i>Silhouette K-Medoids</i> Faktor 3	68
Gambar 4.29 <i>Silhouette K-Medoids</i> Faktor 3	69
Gambar 4.30 Pemetaan Hasil <i>Clustering</i> Terbaik	72

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Hasil <i>Output Uji Bartlett</i>	81
Lampiran 2. Hasil <i>Output Total Variance Explained</i>	81
Lampiran 3. Hasil <i>Output Rotated Component Matrix^a</i>	82
Lampiran 4. Syntax R Studio <i>Average Silhouette Width Method “K-Means”</i>	82
Lampiran 5. Syntax R Studio <i>Nilai Silhouette 1-10 Method “K-Means”</i>	85
Lampiran 6. Syntax R Studio <i>Method “K-Medoids”</i>	90
Lampiran 7. Hasil <i>Output Uji Distribusi Normal Kolmogorov-Smirnov</i>	93
Lampiran 8. Hasil <i>Output Deskriptif Mean dan Median</i>	94