

**PENERAPAN *K-MEANS CLUSTERING* DENGAN EVALUASI
V-MEASURE UNTUK PENGELOMPOKAN WILAYAH
BERDASARKAN DATA POTENSI SUMBER
KESEJAHTERAAN SOSIAL**

SKRIPSI

**OLEH
RIZA MAR'ATUS SHOLIHAH
NIM. 200601110054**



**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2025**

**PENERAPAN *K-MEANS CLUSTERING* DENGAN EVALUASI
V-MEASURE UNTUK PENGELOMPOKAN WILAYAH
BERDASARKAN DATA POTENSI SUMBER
KESEJAHTERAAN SOSIAL**

SKRIPSI

**Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh
RIZA MAR'ATUS SHOLIHAH
NIM. 200601110054**

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2025**

**PENERAPAN *K-MEANS CLUSTERING* DENGAN EVALUASI
V-MEASURE UNTUK PENGELOMPOKAN WILAYAH
BERDASARKAN DATA POTENSI SUMBER
KESEJAHTERAAN SOSIAL**

SKRIPSI

Oleh
Riza Mar'atus Sholihah
NIM. 200601110054

Telah Disetujui untuk Diuji

Malang, 13 Juni 2025

Dosen Pembimbing I



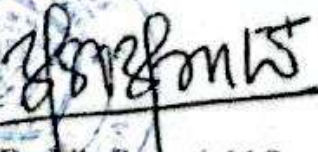
Dr. Fachrud Rozi, M.Si
NIP. 19800527 200801 1 012

Dosen Pembimbing II



Dr. Abdussakir, M.Pd
NIP. 19751006 200312 1 001

Mengetahui,
Ketua Program Studi Matematika


Dr. Elly Susanti, M.Sc
NIP. 19741129 200012 2 005

**PENERAPAN K-MEANS CLUSTERING DENGAN EVALUASI
V-MEASURE UNTUK PENGELOMPOKAN WILAYAH
BERDASARKAN DATA POTENSI SUMBER
KESEJAHTERAAN SOSIAL**

SKRIPSI

**Oleh
Riza Mar'atus Sholihah
NIM. 200601110054**

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)

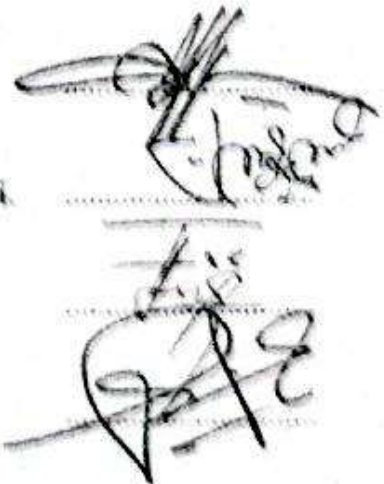
Tanggal 18 Juni 2025

Ketua Penguji : Abdul Aziz, M.Si

Anggota Penguji 1 : Ria Dhea Layla Nur Karisma, M.Si

Anggota Penguji 2 : Dr. Fachrur Rozi, M.Si

Anggota Penguji 3 : Dr. Abdussakir, M.Pd



Mengetahui,
Ketua Program Studi Matematika



Dr. Elly Susanti, M.Sc.
NIP. 19741129 200012 2 005



PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Riza Mar'atus Sholihah
NIM : 200601110054
Program Studi : Matematika
Fakultas : Sains dan Teknologi
Judul Skripsi : Penerapan *K-Means Clustering* dengan Evaluasi *V-Measure*
untuk Pengelompokan Wilayah Berdasarkan Data Potensi
Sumber Kesejahteraan Sosial

menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan dan pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar rujukan. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 18 Juni 2025
Yang membuat pernyataan,



Riza Mar'atus Sholihah
NIM. 200601110054

MOTO

"Hiduplah seakan kamu mati besok, belajarlh seakan kamu hidup selamanya."

-Mahatma Gandhi-

PERSEMBAHAN

Dengan segenap ketulusan hati, skripsi ini peneliti persembahkan kepada:
Keluarga tercinta, khususnya kepada kedua orang tua peneliti yang senantiasa memberikan segala doa, dukungan, kasih sayang, dan semangat yang tak pernah surut. Nasihat-nasihat Bapak dan Ibu selalu menjadi penerang dalam setiap langkah dan penguat peneliti.

KATA PENGANTAR

Assalamu 'alaikum Warahmatullahi Wabarakatuh

Alhamdulillah, puji syukur kehadiran Allah swt., yang telah memberikan rahmat dan pertolongan-Nya kepada peneliti, sehingga peneliti dapat menyelesaikan penelitian proposal skripsi ini sebagai salah satu syarat kelulusan pendidikan di Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Peneliti juga mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan selama proses penyusunan proposal skripsi ini, secara khusus kepada:

1. Prof. Dr. H. M. Zainuddin, M.A., selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Hj. Sri Harini, M.Si, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang
3. Dr. Elly Susanti, S.Pd., M.Sc, selaku Ketua Program Studi Matematika, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Fachrur Rozi, M.Si, selaku Dosen Wali dan Dosen Pembimbing I, yang telah memberikan pengetahuan, bimbingan yang bijaksana, dan motivasi, serta nasihat yang tiada henti kepada peneliti.
5. Dr. Abdussakir, M.Pd, selaku Dosen Pembimbing II, atas dukungan dan nasihat berharga kepada peneliti.
6. Abdul Aziz, M.Si., selaku Ketua Penguji dalam ujian skripsi yang telah memberikan saran yang bermanfaat bagi peneliti.
7. Ria Dhea Layla Nur Karisma, M.Si selaku Anggota Penguji I dalam ujian skripsi yang telah memberikan saran yang bermanfaat bagi peneliti.
8. Seluruh dosen Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang, yang telah mencurahkan ilmu dan dedikasi mereka untuk kemajuan kami.
9. Ibu Nurul Hidayah dan Bapak Kusnadi Ridlo, orang tua tercinta, serta Fadlilatul Ilmillah, kakak tersayang peneliti, dengan penuh cinta dan rasa syukur, peneliti menghaturkan terima kasih yang mendalam. Kasih sayang, doa tanpa henti, dan

dukungan tulus yang mereka berikan adalah cahaya yang menerangi setiap langkah, sumber kekuatan di kala lelah, dan semangat yang tak pernah padam bagi peneliti dalam menyelesaikan tugas akhir ini. Semoga setiap untaian doa mereka berbalas kebahagiaan yang tiada tara.

10. Seluruh mahasiswa angkatan 2020 Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang, yang telah bersama dalam perjalanan ini, berbagi tantangan dan kebahagiaan, serta membangun kenangan yang tak terlupakan.

Semoga segala bantuan dan dukungan yang telah diberikan mendapatkan balasan yang setimpal dan penuh berkah dari Allah swt. Dengan ridho Allah swt., peneliti berharap dapat memberikan manfaat bagi diri peneliti dan para pembaca. Amin.

Wassalamu 'alaikum Warahmatullahi Wabarakatuh

Malang, 18 Juni 2025

Peneliti

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGAJUAN.....	ii
HALAMAN PERSETUJUAN	iii
PERNYATAAN KEASLIAN	
TULISAN	Error
! Bookmark not defined.	
MOTO.....	v
PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR SIMBOL	xiv
DAFTAR LAMPIRAN	xv
ABSTRAK	xvi
ABSTRACT	xvii
مستخلص البحث.....	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	7
1.3 Tujuan Penelitian.....	8
1.4 Manfaat Penelitian.....	8
1.5 Batasan Masalah.....	9
BAB II KAJIAN TEORI	10
2.1 Teori Pendukung	10
2.1.1 <i>K-Means Clustering</i>	10
2.1.2 Standardisasi Data	13
2.1.3 Deteksi <i>Outlier</i>	13
2.1.4 Uji Multikolinieritas	14
2.1.5 <i>Elbow Method Criterion</i>	16
2.1.6 <i>V-Measure</i>	17
2.1.7 Data Potensi Sumber Kesejahteraan Sosial	23
2.2 Kajian Integrasi Topik dengan al-Quran/Hadits.....	25
2.3 Kajian Topik dengan Teori Pendukung.....	28
BAB III METODE PENELITIAN	29
3.1 Jenis Penelitian	29
3.2 Data dan Sumber Data.....	29
3.3 Variabel Penelitian	29
3.4 Teknik Analisis Data	30
3.5 Diagram Alir Analisis Data	33
BAB IV HASIL DAN PEMBAHASAN	35
4.1 Analisis Statistika Deskriptif	35
4.2 Standardisasi Data	49

4.3 Deteksi <i>Outlier</i>	53
4.4 Uji Multikolinieritas	55
4.5 Penerapan <i>Elbow Method Criterion</i>	56
4.6 Penerapan Metode <i>K-Means Clustering</i>	58
4.7 Evaluasi Hasil Pengelompokan dengan Metode <i>V-Measure</i>	73
4.8 Kajian Penelitian dalam Perspektif Islam.....	84
BAB V PENUTUP	87
5.1 Kesimpulan.....	87
5.2 Saran	88
DAFTAR RUJUKAN	90
LAMPIRAN	94
RIWAYAT HIDUP	99

DAFTAR TABEL

Tabel 2.1	Kriteria Nilai Homogenitas dan Kelengkapan	19
Tabel 2.2	Kriteria Nilai <i>V-Measure</i>	22
Tabel 3.1	Indikator Data Potensi Sumber Kesejahteraan Sosial.....	30
Tabel 4.1	Hasil Statistika Deskriptif Variabel Penelitian	35
Tabel 4.2	Tabel Data Standardisasi.....	49
Tabel 4.3	Hasil Nilai <i>Variance Inflation Factor</i> (VIF).....	56
Tabel 4.4	Nilai <i>Sum of Square Error</i> (SSE).....	57
Tabel 4.5	Nilai <i>Centroid</i> Awal untuk Iterasi ke-1.....	59
Tabel 4.6	Hasil Perhitungan Jarak <i>Euclidean</i> Iterasi ke-1	60
Tabel 4.7	Nilai <i>Centroid</i> Baru untuk Iterasi ke-2.....	62
Tabel 4.8	Hasil Perhitungan Jarak <i>Euclidean</i> Iterasi ke-2	64
Tabel 4.9	Hasil Perhitungan Jarak <i>Euclidean</i> Iterasi ke-3	65
Tabel 4.10	Hasil Perhitungan Jarak <i>Euclidean</i> Iterasi ke-4	66
Tabel 4.11	Hasil Metode <i>K-Means Clustering</i>	67
Tabel 4.12	Hasil Pengelompokan Wilayah dengan <i>K-Means Clustering</i>	68
Tabel 4.13	Tabel Hasil Pengelompokan Data Metode <i>Centroid Linkage</i>	74
Tabel 4.14	Tabel Hasil Pengelompokan Data.....	75
Tabel 4.15	Distribusi Data Setiap Kelompok	76
Tabel 4.16	Tabel Hubungan Antara Anggota Kelompok dan <i>Cluster</i>	78
Tabel 4.17	Distribusi Data dari Setiap <i>Cluster</i>	80
Tabel 4.18	Tabel Hubungan Anggota <i>Cluster</i> dengan Kelompok.....	81

DAFTAR GAMBAR

Gambar 3.1	Diagram Alir Analisis Data.....	33
Gambar 4.1	Diagram Batang Variabel Pekerja Sosial Profesional (x_1)	37
Gambar 4.2	Diagram Batang Variabel Pekerja Sosial Masyarakat (x_2).....	38
Gambar 4.3	Diagram Batang Variabel Taruna Siaga Bencana (x_3)	39
Gambar 4.4	Diagram Batang Variabel Lembaga Kesejahteraan Sosial (x_4)....	40
Gambar 4.5	Diagram Batang Variabel Karang Taruna (x_5)	41
Gambar 4.6	Diagram Batang Variabel Lembaga Konsultasi Kesejahteraan Keluarga (x_6).....	42
Gambar 4.7	Diagram Batang Variabel Keluarga Pioner (x_7)	43
Gambar 4.8	Diagram Batang Variabel Wahana Kesejahteraan Sosial Berbasis Masyarakat (x_8).....	44
Gambar 4.9	Diagram Batang Variabel Wanita Pemimpin Kesejahteraan Sosial (x_9)	45
Gambar 4.10	Diagram Batang Variabel Penyuluh Sosial (x_{10}).....	46
Gambar 4.11	Diagram Batang Variabel Tenaga Kesejahteraan Sosial Berbasis Masyarakat (x_{11})	47
Gambar 4.12	Diagram Batang Variabel Dunia Usaha (x_{12}).....	48
Gambar 4.13	Dotplot Variabel x_1 Setelah Standardisasi	50
Gambar 4.14	Dotplot Variabel x_2 Setelah Standardisasi	51
Gambar 4.15	Dotplot Variabel x_3 Setelah Standardisasi	51
Gambar 4.16	Dotplot Variabel x_4 Setelah Standardisasi	51
Gambar 4.17	Dotplot Variabel x_5 Setelah Standardisasi	51
Gambar 4.18	Dotplot Variabel x_6 Setelah Standardisasi	51
Gambar 4.19	Dotplot Variabel x_7 Setelah Standardisasi	52
Gambar 4.20	Dotplot Variabel x_8 Setelah Standardisasi	52
Gambar 4.21	Dotplot Variabel x_9 Setelah Standardisasi	52
Gambar 4.22	Dotplot Variabel x_{10} Setelah Standardisasi	52
Gambar 4.23	Dotplot Variabel x_{11} Setelah Standardisasi	52
Gambar 4.24	Dotplot Variabel x_{12} Setelah Standardisasi	53
Gambar 4.25	Grafik Metode <i>Elbow</i> pada Nilai SSE.....	58
Gambar 4.26	Visualisasi Hasil Pengelompokkan <i>K-Means Clustering</i>	69

DAFTAR SIMBOL

n	: Banyak data
m	: Banyak variabel dalam data
$\vec{v}_h$	: Vektor titik <i>centroid</i> ke- h
$\vec{x}_i$	: Vektor data ke- i dengan $\vec{x}_i = (x_{i1}, x_{i2}, \dots, x_{im})^t$
n_h	: Banyak data dalam setiap kelompok ke- h
d_{ih}	: Jarak dari data ke- i terhadap <i>centroid</i> kelompok ke- h
x_{ij}	: Data ke- i pada variabel ke- j
v_{hj}	: Nilai <i>centroid</i> kelompok ke- h pada variabel ke- j
z_{ij}	: Nilai <i>z-score</i> data ke- i pada variabel ke- j
x_{ij}	: Data ke- i pada variabel ke- j
$\bar{x}_j$	: Nilai rata-rata data pada variabel ke- j
s_j	: Nilai standar deviasi pada variabel ke- j
VIF_j	: <i>Variance Inflation Factor</i> untuk variabel x_j
c_h	: Banyak data yang masuk dalam kelompok ke- h
$H(C)$	: Nilai entropi bersama hasil pengelompokan C
$H(C Y)$	: Nilai entropi bersyarat dari C diberikan hasil pengelompokan acuan Y
a_{ch}	: Data ke- c dalam kelompok h
H	: Nilai homogenitas
$H(Y)$	: Nilai entropi bersama dari hasil pengelompokan acuan Y
$H(Y C)$	: Nilai entropi bersyarat hasil pengelompokan acuan Y dan kelompok C
K	: Nilai kelengkapan
v	: Nilai <i>V-Measure</i>

DAFTAR LAMPIRAN

Lampiran 1	Data Potensi Sumber Kesejahteraan Sosial Per Kecamatan di Kabupaten dan Kota Mojokerto	94
Lampiran 2	Hasil Standardisasi Data Potensi Sumber Kesejahteraan Sosial Per Kecamatan di Kabupaten dan Kota Mojokerto	95
Lampiran 3	Hasil Perhitungan <i>Centroid</i> Baru dari Setiap Iterasi	96
Lampiran 4	<i>Syntax Code</i> RStudio untuk Pengolahan Data Penelitian.....	97

ABSTRAK

Sholihah, Riza Mar'atus. 2025. **Penerapan *K-Means Clustering* dengan Evaluasi *V-Measure* untuk Pengelompokan Wilayah Berdasarkan Data Potensi Sumber Kesejahteraan Sosial**. Skripsi. Program Studi Matematika. Fakultas Sains dan Teknologi. Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Fachrur Rozi, M.Si. (II) Dr. Abdussakir, M.Pd.

Kata kunci: *K-Means Clustering*, *V-Measure*, Pengelompokan Wilayah, Kesejahteraan Sosial, Mojokerto.

Kesejahteraan sosial merupakan bagian dari dimensi pembangunan nasional suatu negara sebagai upaya mencapai inklusi sosial melalui peningkatan ketahanan hidup dan sistem penghidupan, khususnya wilayah Kota dan Kabupaten Mojokerto. Penelitian ini bertujuan untuk mengelompokkan wilayah Kabupaten dan Kota Mojokerto berdasarkan data Potensi Sumber Kesejahteraan Sosial (PSKS) dengan menggunakan metode *K-Means Clustering*. Metode ini dipilih karena efisien dalam menangani data berdimensi tinggi dan mampu menghasilkan banyak kelompok yang dapat ditentukan sesuai kebutuhan dan terstruktur. Data sekunder yang digunakan terdiri atas 12 indikator data potensi sumber kesejahteraan sosial dari Badan Pusat Statistik Dinas Sosial wilayah Mojokerto, yang mencerminkan keberadaan lembaga dan kondisi sosial masyarakat di 18 kecamatan. Tahapan penelitian meliputi standardisasi data dengan *z-score*, deteksi *outlier*, penentuan banyak kelompok optimal menggunakan metode Elbow, penerapan *K-Means Clustering*, serta evaluasi hasil pengelompokan menggunakan metode *V-Measure*. Hasil penelitian menunjukkan bahwa wilayah Mojokerto dapat dikelompokkan dalam lima kelompok dengan karakteristik spesifik, yaitu kawasan perkotaan, perbatasan, bersejarah, industry, dan pegunungan. Evaluasi hasil pengelompokan menggunakan *V-Measure* menghasilkan nilai sebesar 0.756, dengan nilai homogenitas 0.859 dan kelengkapan 0.672, yang mengindikasikan kualitas hasil pengelompokan yang akurat.

ABSTRACT

Sholihah, Riza Mar'atus. 2025. **Implementation of K-Means Clustering with V-Measure Evaluation for Regional Grouping Based on Social Welfare Resource Potential Data.** Undergraduate Thesis. Mathematics Study Program. Faculty of Science and Technology. Universitas Islam Negeri Maulana Malik Ibrahim Malang. Supervisors: (I) Dr. Fachrur Rozi, M.Si. (II) Dr. Abdussakir, M.Pd.

Keywords: K-Means Clustering, V-Measure, Regional Grouping, Social Welfare, Mojokerto.

Social welfare is a key dimension of national development, aimed at achieving social inclusion by enhancing livelihood resilience and systems, particularly in Mojokerto City and Regency. This study aims to group regions in Mojokerto Regency and City based on Social Welfare Resource Potential (PSKS) data using the K-Means Clustering method. This method was chosen for its efficiency in handling high-dimensional data and its ability to produce a predetermined number of structured clusters. The secondary data comprised 12 indicators of social welfare resource potential from the Central Berau of Statistics and Social Office of Mojokerto, reflecting the presence of institutions and community social conditions across 18 sub-districts. Research stages included data standardization using z-score, outlier detection, optimal cluster determination using the Elbow method, K-Means Clustering implementation, and cluster evaluation using the V-Measure. The results indicate that Mojokerto regions can be grouped into five clusters with specific characteristics: urban, border, historical, industrial, and mountainous areas. The V-Measure evaluation yielded a score of 0.756, with homogeneity at 0.859 and completeness at 0.672, indicating accurate clustering quality.

مستخلص البحث

الصالحة، ريزا مرآة. ٢٠٢٥. تطبيق تجميع ك مينس مع تقييم في ميسور لتجميع المناطق بناء على بيانات إمكانات مصادر الرعاية الاجتماعية. البحث العلمي. قسم الرياضيات. كلية العلوم والتكنولوجيا. جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف: (١) الدكتور فخر الرازي، الماجستير في العلوم (٢) الدكتور عبد الشاكر، الماجستير في العلوم

الكلمات الأساسية: تجميع ك مينس ، مقياس في ميسور، تجميع المناطق، الرعاية الاجتماعية، موجوكرطا.

تعد الرعاية الاجتماعية بُعداً أساسياً في التنمية الوطنية لأي دولة، كجزء من الجهود لتحقيق الإدماج الاجتماعي من خلال تعزيز صمود الحياة ونظم المعيشة، خاصة في مدينة ومقاطعة موجوكرطا. هدفت هذه الدراسة إلى تجميع مناطق مقاطعة ومدينة موجوكرطا بناء على بيانات إمكانات مصادر الرعاية الاجتماعية باستخدام طريقة تجميع ك مينس. تم اختيار هذه الطريقة لكفاءتها في التعامل مع البيانات عالية الأبعاد وقدرتها على إنتاج عدد محدد مسبقاً من المجموعات المنظمة. تكونت البيانات الثانوية من ١٢ مؤشراً لبيانات إمكانات مصادر الرعاية الاجتماعية من الوكالة المركزية للإحصاء ودائرة الخدمات الاجتماعية في منطقة موجوكرطا، والتي تعكس وجود المؤسسات والظروف الاجتماعية للمجتمع في ١٨ منطقة فرعية. شملت مراحل البحث توحيد البيانات باستخدام، كشف القيم الشاذة، تحديد العدد الأمثل للمجموعات باستخدام طريقة البو، تجميع ك مينس ، وتقييم النتائج باستخدام في ميسور. أظهرت النتائج أن مناطق موجوكرطا يمكن تقسيمها إلى خمس مجموعات ذات خصائص محددة: المناطق الحضرية، الحدودية، التاريخية، الصناعية، والجبلية. حقق التقييم باستخدام مقياس في ميسور قيمة ٠,٧٥٦، مع تجانس ٠,٨٥٩، واكتمال ٠,٦٧٢، مما يشير إلى دقة جودة التجميع.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Teknik analisis statistika multivariat yang mengkaji struktur hubungan antar variabel dan objek dinamakan *clustering*, yang bertujuan untuk mengelompokkan objek ke dalam beberapa kelompok dengan memperhatikan sifat atau karakteristik yang sama (Putri dkk., 2022). Secara umum terdapat dua metode *clustering*: *clustering* hierarki dan *clustering* non hierarki. *Clustering* hierarki adalah metode *clustering* yang membentuk tingkatan seperti pohon yang dihasilkan oleh tahapan pada proses pengelompokannya, namun hasil yang didapatkan kurang optimal (Kusuma & Agani, 2015). Sedangkan *clustering* non hierarki adalah metode *clustering* yang membagi setiap kelompok tanpa adanya tingkatan antar kelompok dan memberikan hasil yang stabil dan konsisten dalam setiap iterasi. *Clustering* non hierarki juga memiliki fleksibilitas dalam menentukan banyak kelompok yang paling sesuai dengan data dan tujuan yang akan dicapai. Selain itu, *clustering* non hierarki dapat menangani data dengan dimensi tinggi karena perhitungan jarak antar titik dapat dilakukan dalam ruang berdimensi rendah (Orleans, 2022).

Algoritma *clustering* non hierarki yang paling umum dan sederhana adalah *K-Means Clustering*. Algoritma ini membagi data ke dalam beberapa kelompok (sebanyak k) yang berbeda, dimana setiap data tergabung ke dalam kelompok yang letaknya berdekatan dengan pusat kelompok (Orleans, 2022). Hasil pengelompokan tersebut akan dievaluasi dengan metode *V-Measure* untuk menunjukkan kualitas atau kebenaran struktur dalam setiap kelompok dalam hubungannya dengan

kelompok pembandingnya. *V-Measure* adalah salah satu ukuran atau nilai numerik yang digunakan untuk mengevaluasi kualitas hasil pengelompokan yang dihitung sebagai rata-rata harmonik. *V-Measure* melakukan evaluasi dengan menggabungkan konsep homogenitas dan kelengkapan untuk memberikan pengukuran yang seimbang (Rosenberg & Hirschberg, 2007). Konsep homogenitas mencerminkan sejauh mana setiap kelompok bersifat homogen, dimana setiap data dalam satu kelompok memiliki kemiripan satu sama lain. Sedangkan konsep kelengkapan mencerminkan sejauh mana semua data dalam satu kelompok ditempatkan dalam satu kelompok oleh algoritma *clustering* (Rosenberg & Hirschberg, 2007).

Metode yang digunakan dalam penelitian ini adalah metode *K-Means Clustering* dengan menggunakan data potensi sumber kesejahteraan sosial atau yang dapat disingkat PSKS dalam lingkup wilayah Kabupaten Mojokerto dan Kota Mojokerto. Permasalahan sosial dari wilayah Mojokerto diakibatkan karena adanya penyakit sosial yang memunculkan kelompok masyarakat miskin. Penyakit sosial yang dimaksud seperti para tuka karya, tuna wisma, dan tuna susila yang belum bisa diterima oleh masyarakat sekitar (Hikmat, 2012). Kelompok ini terbentuk karena adanya urbanisasi yang kemudian mendirikan pemukiman sendiri yang memiliki bangunan cenderung kurang layak, salah satunya di Desa Mojopilang Kecamatan Mojoanyar Kabupaten Mojokerto. Selain itu juga terdapat kawasan yang padat penduduk di wilayah Balongrawe Baru, Kelurahan Kedundung, Kecamatan Magersari, Kota Mojokerto. Menurut Bupati Mojokerto, Ikfina Rahmawati, dilansir dari *website* Dinas Komunikasi dan Informasi yang diunggah pada tanggal 11 Juli 2024, menjelaskan bahwa pengkajian tentang pemukiman

kumuh pada umumnya mencakup tiga hal, diantaranya keadaan fisik, sosial ekonomi yang bermukim, dan dampak yang diakibatkan dari kondisi tersebut. Selain itu, wilayah Mojokerto memiliki banyak potensi dalam sektor wisata namun dalam pengelolaannya kurang maksimal sehingga menyebabkan tekanan pada sumber daya lokal dan mengurangi kualitas pengalaman wisatawan.

Permasalahan yang terjadi saling berkaitan dengan data Potensi Sumber Kesejahteraan Sosial (PSKS) yang mencakup 12 indikator utama. Lingkungan kumuh yang ada dapat diperbaiki dengan peran Tenaga Kesejahteraan Sosial Kecamatan (TKSK) dan Lembaga Kesejahteraan Sosial (LKS), yang mampu mengidentifikasi serta menangani masalah tersebut melalui program rehabilitasi dan pemberdayaan masyarakat. Sementara itu, pengelolaan tempat wisata yang kurang optimal dapat ditingkatkan dengan melibatkan Dunia Usaha dan Karang Taruna yang dapat menciptakan peluang ekonomi dan mendukung pengembangan pariwisata lokal. Dengan meningkatnya urbanisasi, peran pekerja sosial profesional dan pekerja sosial masyarakat menjadi krusial untuk mendukung masyarakat yang baru pindah dan memberikan layanan yang diperlukan. Selain itu, penyuluh sosial dapat memberikan edukasi tentang adaptasi terhadap perubahan yang terjadi. Keluarga juga berperan penting dalam kesejahteraan, yakni wahana kesejahteraan sosial berbasis masyarakat dan keluarga pioner dapat menjadi contoh praktik baik dalam mendukung kesejahteraan komunitas dibersamai dengan lembaga konsultasi kesejahteraan keluarga. Terlebih lagi, keberadaan taruna siaga bencana sangat penting untuk meningkatkan kesiapsiagaan masyarakat terhadap bencana, yang dapat berdampak pada kesejahteraan. Keterlibatan wanita pemimpin kesejahteraan

sosial juga dapat memperkuat partisipasi perempuan dalam program-program kesejahteraan, berkontribusi pada peningkatan kualitas hidup keluarga.

Kesejahteraan sosial merupakan bagian dari dimensi pembangunan nasional suatu negara sebagai upaya mencapai inklusi sosial melalui peningkatan ketahanan hidup dan sistem penghidupan. Pembangunan kesejahteraan sosial pada hakikatnya guna meningkatkan kualitas individu maupun kelompok dalam masyarakat. Setiap individu memiliki harkat dan martabat untuk berperan dan menjalankan fungsinya dalam kehidupan (Maspaitella & Rahakbauwi, 2014). Melalui Kementrian Sosial Republik Indonesia sebagai sektor utama telah mencanangkan berbagai program yang berupaya menanggulangi kemiskinan antara lain berupa subsidi jaminan kesehatan, Program Keluarga Harapan (PKH), tunjangan prestasi akademik (BBP-PPA), dan program biaya pendidikan dan bantuan lainnya (Suhartini & Yuliani, 2021). Namun pada kenyataannya, permasalahan ini masih belum dapat teratasi karena adanya kerancuan data mengenai masyarakat miskin. Banyaknya masyarakat miskin yang harusnya tercantum dalam data kemiskinan tidak tercatat, sebaliknya banyak masyarakat yang tergolong mampu dan kaya tercatat miskin. Hal ini berdampak signifikan terhadap ketepatan sasaran dari program-program pemerintah (Khakim, 2022).

Menurut Yuhenny (2024) berdasarkan data dari Badan Pusat Statistik (BPS), banyak penduduk miskin di Indonesia meningkat sebesar 2,07% pada bulan Maret tahun 2020, dari 26,42 juta jiwa menjadi 27,54 juta jiwa pada bulan Maret tahun 2021. Hal ini mempengaruhi aktivitas ekonomi penduduk dan juga tingkat kemiskinan. Mojokerto sebagai salah wilayah di Indonesia juga tidak luput dari masalah kemiskinan ini. Pada tahun 2023, menurut BPS Kabupaten Mojokerto

2024, wilayah Kabupaten Mojokerto memiliki angka kemiskinan mencapai 9,80% atau naik sebesar 0,09% dengan jumlah penduduk miskin sekitar 112,86 juta jiwa. Angka kemiskinan tersebut mengalami peningkatan dibandingkan pada tahun sebelumnya, yakni 111,03 juta jiwa yang disebabkan karena adanya peningkatan aktifitas ekonomi masyarakat. Menurut Berita Resmi Statistik No. 04/07/3516/Th. V, 17 Juli 2024 faktor yang turut mempengaruhi kemiskinan karena adanya kenaikan harga beras dibandingkan pada tahun sebelumnya, namun telah dilakukan berbagai program yang cukup efektif dalam memberikan intervensi kepada kelompok sasaran yang dikombinasikan dengan ketepatan pilihan paket komoditas harga bergejolak untuk pengurangan beban pengeluaran.

Berkenaan dengan konsep penerapan metode matematika terhadap data untuk mewujudkan peningkatan kesejahteraan masyarakat selaras dengan hadits Nabi Muhammad saw., mengenai perintah keutamaan menuntut ilmu mulai dari lahir sampai mati dan kewajiban mengamalkan ilmu yang merujuk pada hadits,

الْعِلْمُ بِلاَ عَمَلٍ كَالشَّجَرِ بِلاَ ثَمَرٍ

“Ilmu yang tidak diamalkan bagaikan pohon yang tidak berbuah.”

Dari hadits tersebut yang dapat dipetik kesimpulan bahwa apa yang dimiliki, apa yang telah diketahui, apa yang telah dipelajari, harus disampaikan agar bermanfaat bagi sesama. Bagaikan pohon yang berbuah, yang dapat memberi banyak manfaat, kebaikan, serta kebahagiaan terhadap sekitar. Makna dari kata tersebut adalah seseorang yang berilmu hendaknya mempertanggungjawabkannya pada Allah swt. dengan mengamalkan ilmunya. Sehingga ilmu yang didapatkan bermanfaat untuk diri sendiri dan bagi orang lain. Sedangkan jika mencari ilmu

namun bukan untuk diamalkan, maka ilmu yang diperoleh hanya sebagai suatu kebanggaan yang menyebabkan kesombongan.

Beberapa penelitian yang menggunakan *K-Means Clustering* antara lain penelitian oleh Tarigan dkk. (2022) yang menggunakan metode *K-Means* untuk menentukan secara akurat keluarga mana yang berhak menerima kartu Program Keluarga Harapan (PKH) yang bertujuan untuk membantu menanggulangi masalah kemiskinan dan kesejahteraan sosial agar bantuan dapat diberikan secara merata dan tepat sasaran. Penelitian Khakim (2022) membandingkan algoritma klasifikasi data kesejahteraan sosial Kabupaten Bantul untuk upaya pemerataan klasifikasi Data Terpadu Kesejahteraan Sosial (DTKS). Selain itu terdapat penelitian Rahma Meilaniwati dan Fauzan (2022) menggunakan metode *Naïve Bayes* dan *K-Neareast Neighbor* (KNN) yang mengklasifikasikan penduduk miskin penerima PKH. Penelitian Fathi dan Rosmeida (2022) menerapkan algoritma *Weighted Product Method* (WPM) pada sistem pendukung keputusan untuk mengidentifikasi calon penerima bantuan PKH.

Makalah Suarna dan Wijaya (2023) mengelompokkan data transaksi konsumen di CV. Mitra Indexindo Pratama, yang menggunakan pengujian kinerja dengan mengevaluasi *Davies Bouldin Index* (DBI) untuk efisiensi kelompok. Penelitian Wijaya (2022) menerapkan evaluasi *V-Measure* untuk melihat apakah performa kelompok yang dihasilkan oleh algoritma *Balanced Iterative Reduced and Clustering Using Hierarchies* (BIRCH) pada variabel tinggi badan dan berat badan, sudah baik atau tidak pada kelompok perolehan medali dalam olimpiade bola basket, sepak bola, voli pantai dan seluruh cabang olahraga lainnya.

Berdasarkan latar belakang, maka penelitian ini menggunakan metode *K-Means Clustering* dengan Evaluasi *V-Measure* untuk pengelompokan wilayah berdasarkan data potensi sumber kesejahteraan sosial. Penelitian ini juga menambah variabel yang lebih spesifik dari penelitian Puspaningtyas (2017) yang menggunakan metode *cluster* untuk pengelompokan kecamatan di wilayah Mojokerto berdasarkan Jenis Potensi Sumber Kesejahteraan Sosial (PSKS). Pengelompokan ini dilakukan dengan mengidentifikasi pola-pola yang muncul dari setiap indikator data pada tingkat kecamatan, yang diharapkan dapat menjadi pandangan lain untuk mengatasi permasalahan serta dapat meningkatkan kesejahteraan wilayah Mojokerto dengan memanfaatkan potensi masyarakat, yang pada masa sebelumnya menurut penelitian oleh Nella dkk., (2022) penyaluran bantuan kepada Keluarga Penerima Manfaat (KPM), hanya terpaku pada keluarga yang terdaftar dalam DTKS maupun PKH.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, maka perumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana hasil pengelompokan wilayah di Mojokerto berdasarkan data potensi sumber kesejahteraan sosial menggunakan *K-Means Clustering*?
2. Bagaimana evaluasi hasil pengelompokan wilayah di Mojokerto dengan *V-Measure*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah, maka tujuan pada penelitian ini adalah sebagai berikut:

1. Untuk mengetahui hasil pengelompokan wilayah di Mojokerto berdasarkan Data Potensi Sumber kesejahteraan Sosial dengan metode *K-Means Clustering*.
2. Untuk mengetahui evaluasi hasil pengelompokan wilayah di Mojokerto berdasarkan Data Potensi Sumber kesejahteraan Sosial dengan *V-Measure*.

1.4 Manfaat Penelitian

Berdasarkan tujuan penelitian, maka manfaat pada penelitian ini adalah sebagai berikut:

1. Bagi Peneliti

Sebagai ladang amal penerapan ilmu yang diperoleh selama masa perkuliahan, sebagai sarana belajar dalam memahami topik yang lebih mendalam untuk menambah wawasan, meningkatkan kemampuan dan kualitas sumber daya peneliti dalam merencanakan dan melaksanakan kegiatan dengan lebih efektif.

2. Bagi Pembaca

Sebagai sarana pengetahuan serta sebagai perbandingan dan evaluasi dalam menggunakan sumber acuan untuk bidang kajian yang sama dengan menggunakan variabel data yang meliputi sumber potensi kesejahteraan sosial yang berasal dari manusia/masyarakat dan alam/lingkungan baik menggunakan metode yang sama maupun yang berbeda.

3. Bagi Instansi

- a. Program Studi Matematika: Penelitian ini diharapkan dapat memberikan kontribusi positif bagi pengembangan keilmuan di lingkungan Program Studi, khususnya dalam menambah referensi penelitian yang relevan dengan bidang studi. Hasil penelitian ini dapat menjadi bahan pertimbangan, rujukan, maupun pengembangan teori dan praktik dalam mata kuliah terkait, serta mendorong mahasiswa lain untuk melakukan penelitian lanjutan yang lebih mendalam pada topik serupa.
- b. Dinas Sosial Kabupaten dan Kota Mojokerto: Penelitian ini diharapkan dapat menjadi masukan strategis bagi Dinas Sosial Kabupaten dan Kota Mojokerto dalam merumuskan, mengevaluasi, atau menyempurnakan program-program sosial yang berkaitan dengan topik penelitian, membantu dalam memahami kondisi lapangan secara lebih objektif, serta menjadi dasar dalam pengambilan kebijakan atau pengembangan layanan sosial yang lebih efektif, tepat sasaran, dan berkelanjutan.

1.5 Batasan Masalah

Berdasarkan rumusan masalah, maka batasan masalah pada penelitian ini adalah:

1. Menggunakan pengukuran jarak *Euclidean*.
2. Menggunakan metode Elbow untuk menentukan banyak kelompok yang dibagi (k kluster).
3. Pemilihan *centroid* awal dilakukan secara acak (*random*).

4. Data PSKS yang digunakan adalah data PSKS Kabupaten Mojokerto dan Kota Mojokerto tahun 2023-2024 dengan variabel yang berasal dari Sumber Daya Manusia.
5. Variabel yang digunakan mengacu pada Peraturan Menteri Sosial Republik Indonesia No.08 Tahun 2012 tentang Pedoman Pendataan dan Pengelolaan Data Pemerlu Pelayanan Kesejahteraan Sosial Potensi dan Sumber Kesejahteraan Sosial.

BAB II

KAJIAN TEORI

2.1 Teori Pendukung

2.1.1 *K-Means Clustering*

Clustering adalah suatu teknik dalam analisis statistika multivariat yang mengkaji struktur hubungan antar variabel dan objek yang bertujuan untuk mengelompokkan objek ke dalam beberapa kelompok dengan memperhatikan sifat atau karakteristik yang sama (Putri dkk., 2022). Secara umum terdapat dua metode *clustering*: *clustering* hierarki dan *clustering* non hierarki. Dengan metode *clustering* ini data akan dibagi menjadi beberapa kelompok dengan menggunakan metode *clustering* berbasis jarak dari nilai rata-rata yang dikenal dengan *K-Means*.

K-Means adalah salah satu metode *clustering* non hierarki yang membagi data menjadi beberapa kelompok sehingga data yang memiliki karakteristik yang sama dapat dikelompokkan ke dalam kelompok yang sama, dengan tujuan untuk meminimalisir variasi antar objek dan memaksimalkan variasi antar kelompok (Härdle & Simar, 2015). Persamaan karakteristik setiap data terhadap kelompok diukur dengan kedekatan nilai data terhadap nilai rata-rata pada kelompok yang disebut *centroid cluster*. Konsep dasar dari metode *K-Means* adalah menentukan nilai pusat kelompok secara iteratif, artinya menggunakan nilai awal untuk menghasilkan rangkaian solusi yang bertahap untuk suatu kelas permasalahan, menjadikannya lebih mendekati solusi optimal seiring bertambahnya banyak

iterasi. Setiap iterasi bertujuan untuk memperbaiki hasil dari iterasi sebelumnya hingga konvergensi atau batas iterasi tercapai (Rochimah dkk., 2006).

K-Means Clustering juga termasuk ke dalam *distance based clustering* yang hasilnya membagi data ke dalam k -kelompok dan memerlukan proses penentuan titik awal yang disebut k -centroid secara acak. Menurut Johnson dan Wichern (2007), tahapan pengelompokan menggunakan metode *K-Means Clustering*, sebagai berikut:

1. Proses pengelompokan diawali dengan mengidentifikasi data yang akan dikelompokkan. Berikut identifikasi umum dari data dalam penelitian:

$$x_{ij}(i = 1, 2, 3, \dots, n; j = 1, 2, 3, \dots, m) \quad (2.1)$$

dengan:

n : Banyak data yang akan dikelompokkan,

m : Banyak variabel dalam data.

2. Menentukan nilai k , sebagai nilai banyaknya kelompok yang akan dibentuk. Pada penelitian ini penentuan nilai k menggunakan *Elbow Method Criterion* yang akan dijelaskan pada subbab selanjutnya.
3. Menentukan titik pusat dari kelompok yang disebut *centroid* secara acak dari banyaknya kelompok yang telah ditentukan.
4. Kemudian dilakukan perulangan penentuan *centroid* kelompok dengan rumus berikut untuk melakukan iterasi selanjutnya sampai semua data memiliki kelompok masing-masing sesuai pada konsep *K-Means*.

$$\vec{v}_h = \frac{\sum_{i=1}^{n_h} \vec{x}_i}{n_h} \quad (2.2)$$

dengan:

$\vec{v}_h$: Vektor titik *centroid* ke- h dengan $\vec{v}_h = (v_{h1}, v_{h2}, \dots, v_{hm})^t$ atau

$$\vec{v}_h = v_{hj} \text{ dengan } j = 1, 2, 3, \dots, m.,$$

$\vec{x}_i$: Vektor data ke- i dengan $\vec{x}_i = (x_{i1}, x_{i2}, \dots, x_{im})^t$,

n_h : Banyak data dalam setiap kelompok ke- k .

5. Menghitung jarak antara setiap data dengan *centroid* kelompok menggunakan rumus berikut:

$$d_{ih} = \sqrt{\sum_{j=1}^m (x_{ij} - v_{hj})^2} \quad (2.3)$$

dengan:

d_{ih} : Jarak dari data ke- i terhadap *centroid* kelompok ke- h ,

x_{ij} : Data ke- i pada variabel ke- j ,

v_{hj} : Nilai *centroid* kelompok ke- h pada variabel ke- j ,

m : Banyak variabel.

6. Mengelompokkan data yang memiliki nilai hasil perhitungan yang minimum ke *centroid* terdekat.

Urutan dalam pengelompokan hasil perhitungan jarak pada rumus (2.3) dipilih nilai jarak yang terkecil, kemudian dilihat letak nilai tersebut berada dikolom mana, kolom tersebut yang akan menjadi nama kelompok. Berdasarkan tahapan dalam prosesnya, *K-Means Clustering* memiliki karakteristik yang sangat sensitif terhadap penentuan nilai awal secara acak. Di samping prosesnya sangat cepat, *K-Means Clustering* selalu berubah-ubah sifatnya atau sangat sulit mencapai keoptimalan.

2.1.2 Standardisasi Data

Standardisasi data dilakukan jika pada variabel dalam penelitian memiliki perbedaan ukuran satuan yang cukup besar. Perbedaan satuan ini yang akan menyebabkan perhitungan pada analisis yang akan dilakukan selanjutnya tidak valid (Rahardja dkk., 2020). Oleh karena itu, dalam penelitian ini dilakukan proses standardisasi data melalui transformasi pada data dalam bentuk *z-score*, yang menurut Lawuna dkk (2024), diperoleh dari data dikurangi *mean* dan dibagi dengan standar deviasi seperti di bawah ini:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (2.4)$$

dengan:

z_{ij} : Nilai *z-score* data ke-*i* pada variabel ke-*j*,

x_{ij} : Data ke-*i* pada variabel ke-*j*,

$\bar{x}_j$: Nilai rata-rata data pada variabel ke-*j*,

s_j : Nilai standar deviasi pada variabel ke-*j*.

2.1.3 Deteksi *Outlier*

Data yang diolah dalam proses analisis yang akan dilakukan hanya sebagian kecil dari populasi, maka agar hasil yang diperoleh dapat menggeneralisasi, data yang diolah sebaiknya bersifat representatif, yang artinya asumsi bahwa sampel yang diambil mewakili populasi. Oleh karena itu, perlu dilakukan deteksi terhadap *outlier*. Jika dalam proses analisis yang dilakukan terdeteksi *outlier*, maka langkah yang dapat dilakukan adalah dengan menghilangkan dari sampel agar mendapatkan hasil yang tidak bias. Selain itu,

juga dapat dilakukan dengan transformasi yang meliputi transformasi log, akar kuadrat (*square root*), dan invers. Transformasi ini dapat memadatkan nilai yang lebih tinggi, sehingga mengurangi efek dari nilai ekstrem. Deteksi *outlier* digunakan untuk menemukan data yang berbeda dengan data mayoritas lain. Walaupun memiliki nilai yang berbeda dan sering dinggap *noise*, data *outlier* sering kali mengandung informasi yang sangat berguna (Silvi, 2018) .

Menurut Ghozali (2011) pengecekan *outlier* data juga dapat dilakukan dengan melihat nilai standardisasi data yang telah dilakukan menggunakan nilai *z-score*. Suatu data dianggap sebagai nilai *outlier* jika nilai *z-score* kurang dari -3.00 atau lebih dari 3.00 . Dalam melakukan penghapusan variabel yang memiliki proporsi *outlier* tinggi, diperlukan metode perhitungan *Inter-Quartile Range* (IQR) untuk mendeteksi nilai *outlier* tersebut (Sihombing dkk., 2023). IQR adalah selisih antara kuartil pertama (Q_1) dan kuartil ketiga (Q_3) dalam suatu kumpulan data (Dash dkk., 2023).

2.1.4 Uji Multikolinieritas

Multikolinieritas adalah kondisi yang terjadi karena adanya kemungkinan hubungan atau korelasi antar variabel (Zaki dkk., 2022). Dalam melakukan pengelompokan sebaiknya tidak terdapat korelasi antar variabel, apabila terdapat korelasi maka dianjurkan untuk menghilangkan beberapa variabel yang memiliki nilai korelasi yang cukup besar (Santoso, 2015). Nilai yang dapat digunakan untuk mengetahui adanya multikolinieritas dengan memperhatikan faktor inflasi ragam atau *Variance Inflation Factor* yang dapat disebut VIF. VIF digunakan sebagai kriteria mendeteksi multikolinieritas yang melibatkan lebih dari dua

variabel bebas dengan menyimpulkan hasilnya (Zaki dkk., 2022). Kriteria nilai VIF menurut Ghazali (2011) menunjukkan bahwa jika nilai VIF dari data lebih besar dari 10 maka data tersebut terindikasi adanya multikolinieritas, sehingga perlu dilakukan tindakan perbaikan seperti dengan transformasi logaritma natural, metode *Principal Component Analysis* (PCA), dan metode-metode lainnya. Menurut Ghazali (2011) uji hipotesis untuk uji multikolinieritas data dapat dilakukan sebagai berikut:

1. Hipotesis:

H_0 : Tidak terdapat multikolinieritas

H_1 : Terdapat multikolinieritas

2. Statistik uji VIF:

$$VIF_j = \frac{1}{1 - R_{x_j(x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_m)}^2} \quad (2.5)$$

$$R^2 = \frac{SSR}{SST} \quad (2.6)$$

dengan:

VIF_j : *Variance Inflation Factor* untuk variabel x_j ,

$R_{x_j(x_1, x_2, x_{j-1}, x_{j+1}, \dots, x_m)}^2$: Koefisien determinasi regresi dari variabel $(x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_m)$ terhadap x_j .

SSR = Jumlah kuadrat regresi

SSE = Jumlah kuadrat total

3. Daerah Kritis

Nilai $VIF \leq 10$ maka terima H_0

Nilai $VIF > 10$ maka tolak H_0

4. Keputusan:

Jika nilai $VIF \leq 10$ maka terima H_0 yang menunjukkan tidak terjadinya multikolinieritas. Jika nilai $VIF > 10$ maka tolak H_0 yang menunjukkan terjadinya multikolinieritas.

2.1.5 *Elbow Method Criterion*

Elbow Method Criterion atau yang dapat disebut dengan metode *Elbow* adalah salah satu metode yang dapat digunakan dalam menentukan banyaknya kelompok terbaik yang dapat dibentuk. Penentuan tersebut dapat dilakukan dengan melihat persentase nilai pada grafik setiap kelompok yang membentuk sudut siku-siku pada titik tertentu. Metode *Elbow* ini bertujuan untuk memilih nilai k yang kecil dan memiliki nilai *withinss* yang rendah (Maori & Evanita, 2023). Nilai k yang terdapat pada kombinasi siku dengan *K-Means* adalah grafik hubungan kelompok dengan penurunan *error*.

Banyak kelompok sebanyak k yang dihasilkan dari pengujian dengan *K-Means* dievaluasi dengan SSE. SSE atau *Sum of Square Error* merupakan rumus yang digunakan untuk mengukur perbedaan antara data yang telah dilakukan sebelumnya (Solikhun dkk., 2022). Penggunaan nilai SSE dalam metode *elbow* memberikan ukuran kuantitatif tentang seberapa baik data terkelompok dalam suatu kelompok (Dewi dkk., 2019). Semakin kecil nilai SSE yang diperoleh maka semakin baik data terkelompok. Adapun persamaan yang digunakan dalam metode *Elbow* menurut Maori dan Evanita (2023) yaitu nilai total WSS (*Within Cluster Sum of Squares*) atau biasa disebut dengan SSE (*Sum Square Error*) dengan rumus:

$$SSE = \sum_{h=1}^k \sum_{i=1}^{c_h} (\vec{x}_i - \vec{v}_h)^2 \quad (2.7)$$

dengan:

c_h : Banyak data yang masuk dalam kelompok ke- h ,

Metode *Elbow* dalam menentukan nilai k pada konsep *K-Means* dilakukan

dengan langkah berikut:

1. Menentukan awal nilai k ,
2. Menentukan hasil SSE dari setiap nilai k dengan rumus (2.7),
3. Analisis hasil SSE dari nilai k yang mengalami penurunan secara drastis.
4. Cari dan tetapkan nilai k yang berbentuk siku atau yang memiliki sudut paling kecil.

Perhitungan metode *Elbow* dengan SSE ini lebih sering digunakan dalam proses pengelompokan seperti halnya dalam konsep *K-Means Clustering* karena berhubungan dengan optimalisasi kelompok. Jika dibandingkan dengan metode penentuan banyak kelompok lainnya, metode *Elbow* lebih berfokus pada pengoptimalkan banyak kelompok berdasarkan penurunan nilai SSE. Sedangkan metode lainnya seperti *Silhouette Coefficient* lebih berfokus pada mengevaluasi kualitas banyak kelompok yang terbentuk (Ekasetya & Jananto, 2020).

2.1.6 V-Measure

Pada subbab sebelumnya telah dijelaskan mengenai penentuan nilai k yang optimal dan proses pengelompokan data penelitian. Selanjutnya pada subbab ini akan dijelaskan mengenai validasi dan evaluasi hasil pengelompokan tersebut, agar hasil yang diperoleh memiliki kriteria yang sangat kurang akurat

hingga sangat akurat dengan menggunakan metode *V-Measure*. *V-Measure* adalah suatu metode perhitungan evaluasi kelompok eksternal yang berbasis entropi, menggambarkan ketidakaturan dalam suatu sistem, yang dikembangkan oleh Andrew Rosenberg dan Julia Hirschberg dari Columbia University, New York. Perhitungan ini dirancang untuk mengatasi masalah yang memengaruhi perhitungan pada pengelompokan sebelumnya, seperti adanya ketergantungan pada suatu kumpulan data, dimana suatu kelompok hanya sebagian data yang dapat dievaluasi (Rosenberg & Hirschberg, 2007). Pada *V-Measure* evaluasi dilakukan secara akurat dengan penggabungan dua aspek yang penting dalam suatu *clustering*, yaitu aspek homogenitas dan kelengkapan. Sehingga secara eksplisit, *V-Measure* mengukur seberapa akurat aspek homogenitas dan kelengkapan dari hasil pengelompokan dapat terpenuhi.

V-Measure dihitung sebagai rata-rata harmonik dari nilai homogenitas dan kelengkapan, seperti pada hal presisi dan perolehan yang biasanya digabungkan dalam perhitungan *F-measure* (Rosenberg & Hirschberg, 2007). *V-Measure* bertujuan untuk mengevaluasi kualitas kelompok yang dihasilkan pada proses sebelumnya karena bersifat ambigu, dimana satu data dapat tergabung dalam beberapa kelompok sekaligus. Aspek homogenitas dalam *V-Measure* mengukur seberapa homogen anggota kelompok dalam kelompoknya. Sedangkan aspek kelengkapan mengukur seberapa lengkap anggota kelompok dalam kelompoknya tidak mengalami perubahan dalam proses pengelompokan. Dengan kedua aspek ini, *V-Measure* dapat memberikan gambaran yang komprehensif tentang kualitas kelompok yang dihasilkan (Rosenberg & Hirschberg, 2007).

Kedua aspek tersebut memberikan nilai yang digunakan untuk memperoleh nilai *V-Measure*. Kriteria nilai untuk kedua aspek tersebut disajikan dalam Tabel 2.1.

Tabel 2.1 Kriteria Nilai Homogenitas dan Kelengkapan

Nilai	Kategori	Keterangan
$0 < H \leq 0.2$ dan $0 < K \leq 0.2$	Sangat Buruk	Kelompok yang terbentuk cenderung heterogen, dengan anggota yang berasal dari kelompok yang berbeda (Rosenberg & Hirschberg, 2007).
$0.2 < H \leq 0.4$ dan $0.2 < K \leq 0.4$	Buruk	Kelompok yang terbentuk cenderung mencerminkan suatu karakteristik tertentu, namun masih terdapat pencampuran anggota dari berbagai kelompok dalam satu kelompok (Rosenberg & Hirschberg, 2007).
$0.4 < H \leq 0.6$ dan $0.4 < K \leq 0.6$	Cukup Baik	Kelompok yang terbentuk memiliki pembagian kelompok yang cenderung lebih homogen, meskipun masih terdapat pencampuran antar kelompok (Rosenberg & Hirschberg, 2007).
$0.6 < H \leq 0.8$ dan $0.6 < K \leq 0.8$	Baik	Kelompok yang terbentuk cukup homogen dan representatif; sebagian besar anggota dalam satu kelompok berasal dari kelompok yang sama, dengan pencampuran yang masih minimal (Rosenberg & Hirschberg, 2007).
$0.8 < H < 1$ dan $0.8 < H < 1$	Sangat Baik	Pengelompokan yang terbentuk hampir sepenuhnya mencerminkan struktur kelompok sebenarnya; hanya sedikit pencampuran, dan cluster yang terbentuk bersifat sangat konsisten (Rosenberg & Hirschberg, 2007).
$H = 1$	Sempurna	Hasil pengelompokan yang terbentuk terdiri dari anggota kelompok yang sama, tidak terdapat pencampuran anggota dari kelompok yang berbeda (Rosenberg & Hirschberg, 2007).

Sumber: (Rosenberg & Hirschberg, 2007)

Kedua aspek dalam *V-Measure* diasumsikan sebanyak N titik data yang terdiri dari dua pengelompokan utama, yaitu banyak data dalam satu kelompok, $C = \{C_c | c = 1, \dots, l\}$ dan banyak data yang terbentuk dari hasil pengelompokan acuan, $Y = \{Y_h | h = 1, \dots, k\}$. Secara keseluruhan menurut Rosenberg dan Hirschberg (2007) perhitungan *V-Measure* dapat dilakukan dengan langkah-langkah berikut:

1. Menentukan distribusi data dari setiap kelompok.
2. Menghitung nilai entropi bersama kelompok C

$$H(C) = - \sum_{c=1}^{|C|} \frac{\sum_{h=1}^{|Y|} a_{ch}}{n} \log_2 \frac{\sum_{h=1}^{|Y|} a_{ch}}{n} \quad (2.8)$$

dengan:

$H(C)$: Nilai entropi bersama kelompok C .

3. Menghitung nilai entropi bersyarat dari C diberikan hasil pengelompokan acuan Y

$$H(C|Y) = - \sum_{h=1}^{|Y|} \sum_{c=1}^{|C|} \frac{a_{ch}}{N} \log_2 \frac{a_{ch}}{\sum_{c=1}^{|C|} a_{ch}} \quad (2.9)$$

dengan:

$H(C|Y)$: Nilai entropi bersyarat dari C diberikan hasil pengelompokan acuan Y ,

a_{ch} : Data ke- c dalam kelompok h .

4. Menghitung nilai Homogenitas

$$H = 1 - \frac{H(C|Y)}{H(C)} \quad (2.10)$$

dengan:

H : Nilai homogenitas.

5. Menentukan distribusi data dari setiap hasil pengelompokan acuan.
6. Menghitung nilai entropi bersama hasil pengelompokan acuan Y

$$H(Y) = - \sum_{h=1}^{|Y|} \frac{\sum_{c=1}^{|C|} a_{ch}}{n} \log_2 \frac{\sum_{c=1}^{|C|} a_{ch}}{n} \quad (2.11)$$

dengan:

$H(Y)$: Nilai entropi bersama dari hasil pengelompokan acuan Y

7. Menghitung nilai entropi bersyarat dari hasil pengelompokan acuan Y diberikan C

$$H(Y|C) = - \sum_{c=1}^{|C|} \sum_{h=1}^{|Y|} \frac{a_{ch}}{N} \log_2 \frac{a_{ch}}{\sum_{h=1}^{|Y|} a_{ch}} \quad (2.12)$$

dengan:

$H(Y|C)$: Nilai entropi bersyarat hasil pengelompokan acuan Y diberikan hasil pengelompokan C .

8. Menghitung nilai kelengkapan

$$K = 1 - \frac{H(Y|C)}{H(Y)} \quad (2.13)$$

dengan:

K : Nilai kelengkapan

Berdasarkan persamaan pada (2.20) dan (2.23) diperoleh rumus *V-Measure* sebagai berikut:

$$v = \frac{2(H.K)}{H+K} \quad (2.14)$$

dengan:

v : Nilai *V-Measure*.

Pengukuran *V-Measure* juga menggunakan *log* dalam perhitungannya yang berfungsi untuk menghindari bias atau kecenderungan yang mungkin terjadi ketika membandingkan banyak elemen yang berbeda dalam kelompok dan kategori. Penggunaan *log* membantu menyeimbangkan kontribusi dari kelompok yang lebih besar dan lebih kecil, sehingga hasil yang diperoleh lebih adil, representatif, dan analisis yang dilakukan lebih stabil (Boutaba dkk., 2018).

Dengan menggunakan rumus pada (2.14), *V-Measure* memberikan nilai yang menggambarkan seberapa baik kelompok-kelompok tersebut

mempertahankan strukturnya, dapat diartikan bahwa perhitungan tersebut memberikan informasi tentang sejauh mana suatu kelompok berhasil mengelompokkan data yang seharusnya berada dalam kelompok yang sama ke dalam kelompok yang sesuai. Kriteria dalam evaluasi *V-Measure* bernilai 0 jika nilai homogenitas dan kelengkapan menghasilkan nilai 0, dan akan memberikan nilai antara 0 dan 1. Nilai sebesar 1 menunjukkan bahwa hasil pengelompokan tersebut sempurna dalam hal homogenitas dan kelengkapan (Rosenberg & Hirschberg, 2007).

Tabel 2.2 Kriteria Nilai *V-Measure*

Nilai	Kriteria	Keterangan
$0 < v \leq 0.5$	Tidak Akurat	Hasil pengelompokan terdapat campuran anggota yang memiliki kelas yang berbeda menjadi satu kelompok dan anggota dari kelas yang sama tidak dalam satu kelompok (Lancia dkk., 2019).
$0.5 < v \leq 0.9$	Akurat	Hasil pengelompokan terdapat beberapa anggota dalam kelompok yang mayoritas berada dalam kelas yang sama dan dalam <i>cluster</i> yang sama sehingga dalam satu kelompok cukup homogen dan lengkap, meskipun masih terdapat anggota dari kelas yang berbeda (Lancia dkk., 2019).
$v = 1$	Sempurna	Hasil pengelompokan menunjukkan kesamaan yang moderat, di mana semua anggota dalam kelompok berasal dari kelas yang sama dan semua anggota dari kelas tertentu terkelompok bersama (Lancia dkk., 2019).

Sumber: (Lancia dkk., 2019)

Menurut Carlo Lancia dkk. (2019), secara rinci rentang nilai antara 0 dan 1 yang diberikan untuk kriteria dan interpretasi nilai *V-Measure* disajikan dalam Tabel 2.2. Metode matematika *V-Measure* ini tidak bergantung pada banyak data, banyak kelompok, dan ukuran data sehingga memberikan cara yang sistematis dan terukur untuk mengevaluasi kualitas kelompok-kelompok dalam analisis data dengan mengukur homogenitas dan kelengkapan kelompok (Ardali dkk., 2024).

V-Measure memberikan gambaran tentang sejauh mana suatu kelompok mencerminkan distribusi sampel data yang sebenarnya dari data dan seberapa konsisten serta lengkap suatu kelompok dalam tugas pengelompokan data.

2.1.7 Data Potensi Sumber Kesejahteraan Sosial

Dalam Peraturan Menteri Sosial Republik Indonesia No.5 Tahun 2019 tentang Pengelolaan Data Terpadu Kesejahteraan Sosial Potensi dan Sumber Kesejahteraan Sosial, mendefinisikan bahwa potensi sumber kesejahteraan sosial adalah segala sesuatu yang berharga bersifat manusiawi, sosial, dan alam, baik perseorangan, keluarga, kelompok, dan/atau masyarakat yang dapat berperan serta untuk menjaga, menciptakan, mendukung, dan memperkuat usaha penyelenggaraan kesejahteraan sosial. Berdasarkan Peraturan Menteri Sosial Republik Indonesia No.08 Tahun 2012 tentang Pedoman Pendataan dan Pengelolaan Data Pemerlu Pelayanan Kesejahteraan Sosial Potensi dan Sumber Kesejahteraan Sosial, meliputi:

1. Pekerja Sosial Profesional (PSP) adalah individu yang bekerja di organisasi pemerintah atau swasta yang memiliki keterampilan, minat, dan karier pekerjaan sosialnya yang dapat diperoleh melalui pelatihan.
2. Pekerja Sosial Masyarakat (PSM) adalah anggota komunitas yang secara individu sukarela mengabdikan diri di bidang kesejahteraan sosial dengan kesadaran diri akan kepedulian dan karakteristik sosial lainnya.
3. Taruna Siaga Bencana (TAGANA) adalah relawan masyarakat yang berminat dan aktif dalam penanggulangan bencana.

4. Lembaga Kesejahteraan Sosial (LKS) adalah organisasi kemasyarakatan atau perkumpulan sosial yang menyelenggarakan kegiatan kesejahteraan sosial yang didirikan oleh masyarakat, baik berbadan hukum maupun tidak.
5. Karang Taruna merupakan organisasi sosial di pedesaan yang berfungsi sebagai wadah dan wahana pengembangan seluruh anggota masyarakat yang bergerak di bidang kesejahteraan sosial.
6. Lembaga Konsultasi Kesejahteraan Keluarga (LK3) merupakan lembaga atau organisasi yang memberikan konsultasi, jasa konsultasi, penyediaan atau penyebarluasan informasi dan pemberdayaan keluarga secara lebih profesional dan mendalam.
7. Keluarga Pioner adalah keluarga yang memiliki kemampuan dalam menyelesaikan masalahnya dengan efisien dan dapat dijadikan teladan bagi keluarga lain.
8. Tenaga Kesejahteraan Sosial Kecamatan (TKSK) pejabat inti yang bertanggung jawab menyelenggarakan kegiatan jaminan sosial di tingkat kabupaten/kota.
9. Wahana Kesejahteraan Sosial Keluarga Berbasis Masyarakat (WKSBM) merupakan wadah kerjasama antar lembaga pelayanan sosial termasuk kelompok usaha dan jaringan pendukung di masyarakat.
10. Wanita Pemimpin Kesejahteraan Sosial adalah perempuan yang menginspirasi dan mendorong pelaksanaan program perlindungan sosial di lingkungannya.

11. Penyuluh Sosial merupakan Pegawai Negeri Sipil (PNS) maupun tokoh masyarakat yang mempunyai tanggung jawab, wewenang dan kewajiban memberikan nasihat di bidang kesejahteraan sosial.
12. Dunia Usaha adalah organisasi yang bergerak di bidang usaha, industri atau produk dan jasa yang berkoordinasi dengan badan usaha milik negara, badan usaha milik negara, pengusaha dan jaringannya untuk ikut serta dalam organisasi perlindungan sosial sebagai bentuk tanggung jawab sosial.

2.2 Kajian Integrasi Topik dengan al-Quran/Hadits

Sebagaimana hadits yang telah tercantum pada latar belakang mengenai keutamaan menuntut ilmu dan perintah untuk mengamalkan ilmu yang berkenaan dengan tujuan penelitian mengenai penerapan metode *K-Means Clustering* dengan evaluasi menggunakan metode *V-Measure* dalam data potensi sumber kesejahteraan sosial untuk meningkatkan kesejahteraan masyarakat dan pengentasan kemiskinan, selaras dengan QS. al-Baqarah (2) ayat 177:

لَيْسَ الْبِرَّ أَنْ تُوَلُّوا وُجُوهَكُمْ قَبْلَ الْمَشْرِقِ وَالْمَغْرِبِ وَلَكِنَّ الْبِرَّ مَنْ آمَنَ بِاللَّهِ وَالْيَوْمِ الْآخِرِ
وَالْمَلَائِكَةِ وَالْكِتَابِ وَالنَّبِيِّينَ وَآتَى الْمَالَ عَلَى حُبِّهِ ذَوِي الْقُرْبَىٰ وَالْيَتَامَىٰ وَالْمَسْكِينِ وَابْنَ السَّبِيلِ
وَفِي الرِّقَابِ وَأَقَامَ الصَّلَاةَ وَآتَى الزَّكَاةَ وَالْمُوفُونَ بِعَهْدِهِمْ إِذَا عَاهَدُوا وَالصَّابِرِينَ فِي الْبَأْسَاءِ وَالضَّرَّاءِ وَحِينَ
الْبَأْسِ فَمَا أُولَٰئِكَ إِلَّا الَّذِينَ صَدَقُوا وَمِنْكُمْ هُمُ الْمُتَّقُونَ

Artinya: “Kebajikan itu bukanlah menghadapkan wajahmu ke arah timur dan barat, melainkan kebajikan itu ialah (kebajikan) orang yang beriman kepada Allah, hari Akhir, malaikat-malaikat, kitab suci, dan nabi-nabi; memberikan harta yang dicintainya kepada kerabat, anak yatim, orang miskin, musafir, peminta-minta, dan (memerdekakan) hamba sahaya; melaksanakan salat; menunaikan zakat; menepati janji apabila berjanji; sabar dalam kemelaratan, penderitaan, dan pada masa peperangan. Mereka itulah orang-orang yang benar dan mereka itulah orang-orang yang bertakwa.”(QS. al-Baqarah/2:177)

QS. al-Baqarah (2) ayat 177 menjelaskan bahwa kebajikan yang sesungguhnya adalah beriman kepada Allah dengan tidak menyekutukan-Nya; beriman kepada hari akhir yaitu hari pembalasan segala amal perbuatan selama di dunia, sehingga mendorong manusia untuk selalu berbuat baik; beriman kepada malaikat-malaikat yang taat menjalankan perintah Allah dan tidak pernah berbuat maksiat sehingga mendorong manusia untuk meneladani ketaatannya; beriman kepada kitab-kitab yang diturunkan oleh Allah kepada para rasul; beriman kepada nabi-nabi yang selalu menyampaikan kebenaran meskipun banyak yang memusuhinya (Purnamasari, 2023).

Kebajikan orang yang memberikan harta yang dicintainya kepada kerabat yang kurang mampu, anak yatim, karena mereka sudah kehilangan orang tua, sehingga setiap orang beriman patut memberikan kebaikan kepada mereka, orang-orang miskin yang hidupnya serba kekuarangan dalam memenuhi kebutuhan sehari-hari, orang-orang yang dalam perjalanan atau musafir yang kehabisan bekal perjalanan, peminta-minta untuk meringankan penderitaan dan kekurangannya, dan untuk memerdekakan hamba sahaya yang timbul akibat praktik perbudakan (Purnamasari, 2023).

Selain itu, kebajikan orang yang melaksanakan salat dengan khusyuk dan memenuhi syarat dan rukunnya, menunaikan zakat sesuai ketentuan dan tidak menunda-nunda pelaksanaannya, orang-orang yang menepati janji, orang yang sabar saat mengalami musibah, kesengsaraan, kepedihan dan berbagai macam kekurangan. Orang yang mempunyai sifat-sifat ini, mereka itulah orang-orang yang benar keimanannya, dan mereka itulah orang-orang yang bertakwa kepada Allah (Purnamasari, 2023).

Perintah Allah untuk mensejahterakan umat manusia juga ditegaskan oleh Nabi Muhammad saw yang terkandung dalam kitab *Arba'in Nawawi* karangan Imam an-Nawawi:

عَنْ أَبِي هُرَيْرَةَ رَضِيَ اللَّهُ عَنْهُ، عَنِ النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ: «مَنْ نَفَسَ عَنْ مُؤْمِنٍ كُرْبَةً مِنْ كُرْبِ الدُّنْيَا، نَفَسَ اللَّهُ عَنْهُ كُرْبَةً مِنْ كُرْبِ يَوْمِ الْقِيَامَةِ. وَمَنْ يَسَّرَ عَلَى مُعْسِرٍ، يَسَّرَ اللَّهُ عَلَيْهِ فِي الدُّنْيَا وَالْآخِرَةِ. وَمَنْ سَتَرَ مُسْلِمًا سَتَرَهُ اللَّهُ فِي الدُّنْيَا وَالْآخِرَةِ. وَاللَّهُ فِي عَوْنِ الْعَبْدِ مَا كَانَ الْعَبْدُ فِي عَوْنِ أَخِيهِ. وَمَنْ سَلَكَ طَرِيقًا يَلْتَمِسُ فِيهِ عِلْمًا سَهَّلَ اللَّهُ لَهُ بِهِ طَرِيقًا إِلَى الْجَنَّةِ. وَمَا اجْتَمَعَ قَوْمٌ فِي بَيْتٍ مِنْ بُيُوتِ اللَّهِ يَتْلُونَ كِتَابَ اللَّهِ وَيَتَذَكَّرُونَ بَيْنَهُمْ إِلَّا نَزَلَتْ عَلَيْهِمُ السَّكِينَةُ، وَعَشِيَتْهُمْ الرَّحْمَةُ، وَخَفَّتْهُمُ الْمَلَائِكَةُ، وَذَكَرَهُمُ اللَّهُ فِيمَنْ عِنْدَهُ، وَمَنْ بَطَأَ بِهِ عَمَلُهُ لَمْ يُسْرِعْ بِهِ نَسَبُهُ» رَوَاهُ مُسْلِمٌ بِهَذَا اللَّفْظِ.

Artinya: “*Dari Abu Hurairah ra, Nabi Muhammad saw. bersabda, “Barangsiapa yang menghilangkan kesusahan dari kesusahan-kesusahan dunia orang mukmin, maka Allah akan menghilangkan kesusahan dari kesusahan-kesusahan hari kiamat. Barangsiapa yang memberi kemudahan orang yang kesulitan (utang), maka Allah akan memberi kemudahan baginya di dunia dan akhirat. Siapa saja yang menolong saudaranya, maka Allah akan menolongnya sebagaimana ia menolong saudaranya. Barangsiapa yang menempuh perjalanan dalam rangka menuntut ilmu, maka Allah akan mudahkan baginya jalan menuju surga. Tidaklah berkumpul sekelompok orang di salah satu rumah Allah (masjid) untuk membaca Kitabullah dan saling mempelajarinya di antara mereka, melainkan akan turun kepada mereka ketenangan, rahmat meliputinya, para malaikat mengelilinginya, dan Allah menyanjung namanya kepada Malaikat yang ada di sisi-Nya. Barangsiapa yang lambat amalnya, maka tidak akan bisa dikejar oleh nasabnya (garis keturunannya yang mulia).” (HR. Muslim) (Yahya bin Syaraf al-Nawawi, 1984).*

Penelitian ini berlandaskan prinsip kebajikan yang diuraikan dalam QS. al-Baqarah (2) ayat 177 dan hadits Nabi Muhammad SAW yang menekankan pentingnya menuntut ilmu, melaksanakan kebaikan, serta memberi bantuan kepada mereka yang membutuhkan. Dengan menerapkan metode *K-Means Clustering* dan evaluasi menggunakan *V-Measure* pada data potensi sumber kesejahteraan sosial, penelitian ini bertujuan untuk meningkatkan kesejahteraan masyarakat dan mengentaskan kemiskinan, selaras dengan ajaran Islam tentang kewajiban menolong sesama dan meningkatkan kesejahteraan umat manusia.

2.3 Kajian Topik dengan Teori Pendukung

Penelitian ini disusun berdasarkan beberapa teori pendukung terkait metode *K-Means Clustering*. Topik yang dibahas adalah penerapan *K-Means Clustering* pada indikator di dalam data PSKS yang terdiri dari 12 indikator utama di wilayah Mojokerto. Terdapat beberapa penelitian yang menggunakan metode *K-Means Clustering*, namun masih belum ditemukan evaluasi hasil dari *K-Means Clustering* dengan metode *V-Measure*. Diantaranya penelitian Suhartini dan Yuliani (2021). Penelitian ini menyajikan hasil penggunaan algoritma *K-Means* untuk mengelompokkan masyarakat miskin di Dusun Bagik Endep dan menentukan kelompok masyarakat miskin sesuai kelompok yang diharapkan. Selain itu terdapat beberapa penelitian sebelumnya yang menggunakan beberapa metode lain dalam data DTKS dan PKH. Seperti pada penelitian Rahma Meilaniwati dan Fauzan (2022) menggunakan menggunakan metode *Naïve Bayes* dan *K-Neareast Neighbor* (KNN) yang mengklasifikasikan penduduk miskin penerima PKH.

Penggunaan metode *K-Means Clustering* dilakukan dengan menghitung jarak setiap data terhadap *centroid* menggunakan metode *Euclidean Distance*. Misalkan suatu titik data memiliki jarak terdekat ke *centroid* pertama, maka titik data tersebut dimasukkan kedalam kelompok pertama. Setelah semua data dikelompokkan berdasarkan *centroid* terdekat, akan diperoleh hasil pengelompokan yang optimal dengan varians dalam setiap kelompok diminimalkan yang kemudian dapat diinterpretasikan sebagai hasil penelitian. Kemudian, dilakukan evaluasi dengan menggunakan metode *V-Measure* yang memberikan gambaran tentang sejauh mana kelompok mencerminkan distribusi kelas yang sebenarnya dalam data dan seberapa konsisten serta lengkap kelompok dalam tugas pengelompokan data.

BAB III

METODE PENELITIAN

3.1 Jenis Penelitian

Jenis penelitian yang digunakan adalah penelitian deskriptif kuantitatif yang bersifat analitik yang berfokus pada analisis data dan interpretasi dari hasil yang telah diperoleh. Metode deskriptif digunakan untuk menjelaskan data yang akan dianalisis berdasarkan keadaan yang sebenarnya. Selanjutnya proses pengolahan data secara kuantitatif yang diimplementasikan sesuai dengan kebutuhan dalam metode *K-Means Clustering* dan evaluasi *V-Measure*.

3.2 Data dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh melalui *website* resmi Badan Pusat Statistik (BPS) Kabupaten Mojokerto di dalam publikasi Kecamatan dalam Angka 2024 dan Dinas Sosial Kota Mojokerto. Data dalam penelitian ini dilampirkan pada Lampiran 1.

3.3 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini merupakan 12 indikator dalam data potensi sumber kesejahteraan sosial Kabupaten Mojokerto dan Kota Mojokerto tahun 2023-2024. Berdasarkan Peraturan Menteri Sosial Republik Indonesia No.08 Tahun 2012 tentang Pedoman Pendataan dan Pengelolaan Data Pemerlu Pelayanan Kesejahteraan Sosial Potensi dan Sumber Kesejahteraan Sosial yang disajikan pada Tabel 3.1.

Tabel 3.1 Indikator Data Potensi Sumber Kesejahteraan Sosial

No.	Variabel	Keterangan	Satuan	Skala
1.	x_1	Pekerja Sosial Profesional (PSP)	Orang	Rasio
2.	x_2	Pekerja Sosial Masyarakat (PSM)	Orang	Rasio
3.	x_3	Taruna Siaga Bencana (TAGANA)	Orang	Rasio
4.	x_4	Lembaga Kesejahteraan Sosial (LKS)	Organisasi	Rasio
5.	x_5	Karang Taruna	Organisasi	Rasio
6.	x_6	Lembaga Konsultasi Kesejahteraan Keluarga (LK3)	Organisasi	Rasio
7.	x_7	Keluarga Pioner	Unit	Rasio
8.	x_8	Wahana Kesejahteraan Sosial Keluarga Berbasis Masyarakat (WKSBM)	Organisasi	Rasio
9.	x_9	Wanita Pemimpin Kesejahteraan Sosial	Orang	Rasio
10.	x_{10}	Penyuluh Sosial	Orang	Rasio
11.	x_{11}	Tenaga Kesejahteraan Sosial Kecamatan (TKSK)	Orang	Rasio
12.	x_{12}	Dunia Usaha	Organisasi	Rasio

3.4 Teknik Analisis Data

Untuk mengimplementasikan metode *K-Means Clustering* yang dilanjutkan proses evaluasi dengan *V-Measure* berdasarkan data potensi sumber kesejahteraan sosial dalam penelitian ini, dapat dilakukan dengan langkah-langkah sebagai berikut:

1. Melakukan analisis statistika deskriptif pada data berdasarkan indikator dalam data potensi sumber kesejahteraan sosial Kabupaten dan Kota Mojokerto, yang digunakan sebagai variabel, seperti yang telah disajikan pada Tabel 3.1, yang terdiri dari nilai rata-rata, nilai standar deviasi, nilai minimum dan nilai maksimum, nilai kuartil satu dan nilai kuartil tiga.
2. Melakukan tahapan standardisasi data menggunakan metode transformasi dengan *z-score* pada rumus (2.4) dalam data potensi sumber kesejahteraan sosial Kabupaten dan Kota Mojokerto.
3. Melakukan deteksi *outlier* dengan data yang telah distandardisasi dengan nilai *z-score*, yang digunakan dalam pengelompokan wilayah di Kabupaten dan

Kota Mojokerto berdasarkan data potensi sumber kesejahteraan sosial. Data dengan nilai *z-score* dianggap sebagai nilai *outlier* jika nilai *z-score* kurang dari -3.00 atau lebih dari 3.00 . Jika dalam data penelitian terdeteksi *outlier*, maka variabel yang memiliki nilai *outlier* yang berbeda secara signifikan akan dilakukan penanganan, seperti transformasi data hingga penghapusan data *outlier*.

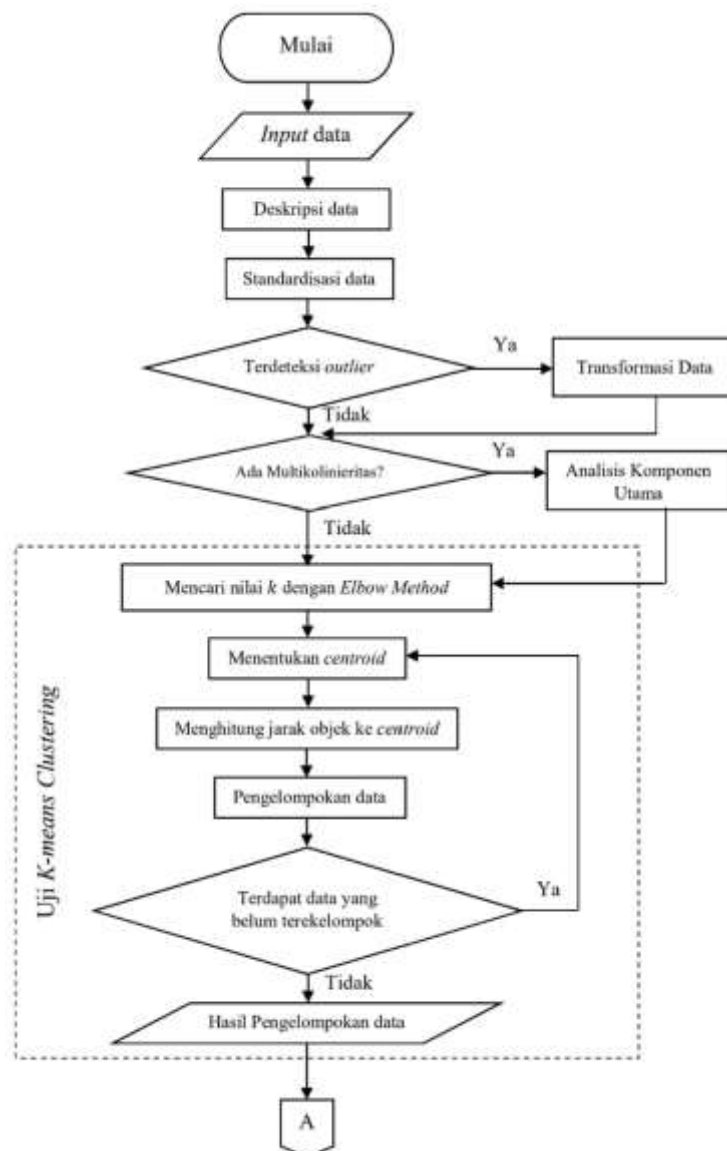
4. Melakukan uji asumsi multikolinieritas menggunakan nilai VIF dengan persamaan pada (2.5) berdasarkan data potensi sumber kesejahteraan sosial. Jika dalam data penelitian memiliki korelasi yang sangat besar antar variabel data maka dilakukan pengolahan data menggunakan metode *Principal Component Analysis* (PCA).
5. Melakukan analisis *K-Means Clustering* untuk pengelompokan wilayah di Kabupaten dan Kota Mojokerto berdasarkan data potensi sumber kesejahteraan sosial. Berikut tahapan yang dilakukan dalam analisis *K-Means Clustering*:
 - a) Menentukan banyaknya kelompok terbaik yang dapat dibentuk menggunakan metode *Elbow* dengan persamaan (2.7),
 - b) Menentukan titik pusat dari kelompok yang disebut *centroid* secara acak,
 - c) Menghitung jarak antara setiap data dengan *centroid* kelompok menggunakan rumus (2.3),
 - d) Mengelompokkan data yang memiliki nilai hasil perhitungan yang minimum ke *centroid* terdekat,

- e) Melakukan perulangan penentuan *centroid* kelompok dengan rumus (2.2) untuk melakukan iterasi selanjutnya sampai semua data memiliki kelompok masing-masing,
 - f) Interpretasi hasil pengelompokan dengan metode *K-Means Clustering*.
6. Melakukan evaluasi *V-Measure* terhadap hasil pengelompokan wilayah di Kabupaten dan Kota Mojokerto berdasarkan data potensi sumber kesejahteraan sosial. Berikut tahapan yang dilakukan dalam evaluasi hasil pengelompokan:
- a) Menentukan distribusi data dalam setiap kelompok,
 - b) Menghitung nilai entropi bersama dari kelompok C dengan rumus (2.8),
 - c) Mendefinisikan data kelompok C dalam hasil pengelompokan acuan Y ,
 - d) Menghitung nilai entropi bersyarat dari kelompok C dengan hasil pengelompokan acuan Y dengan rumus (2.9),
 - e) Menghitung nilai homogenitas menggunakan rumus (2.10),
 - f) Menentukan distribusi data dalam setiap hasil pengelompokan acuan Y ,
 - g) Menghitung nilai entropi bersama hasil pengelompokan acuan Y menggunakan rumus (2.11),
 - h) Mendefinisikan data hasil pengelompokan acuan Y dalam kelompok C
 - i) Menghitung nilai entropi bersyarat dari hasil pengelompokan acuan Y dengan kelompok C menggunakan rumus (2.12),
 - j) Menghitung nilai kelengkapan menggunakan rumus pada (2.13),
 - k) Menghitung nilai *V-Measure* menggunakan rumus pada (2.14),
 - l) Mengevaluasi hasil perhitungan *V-Measure* sesuai dengan kriteria *V-Measure* pada Tabel 2.2.

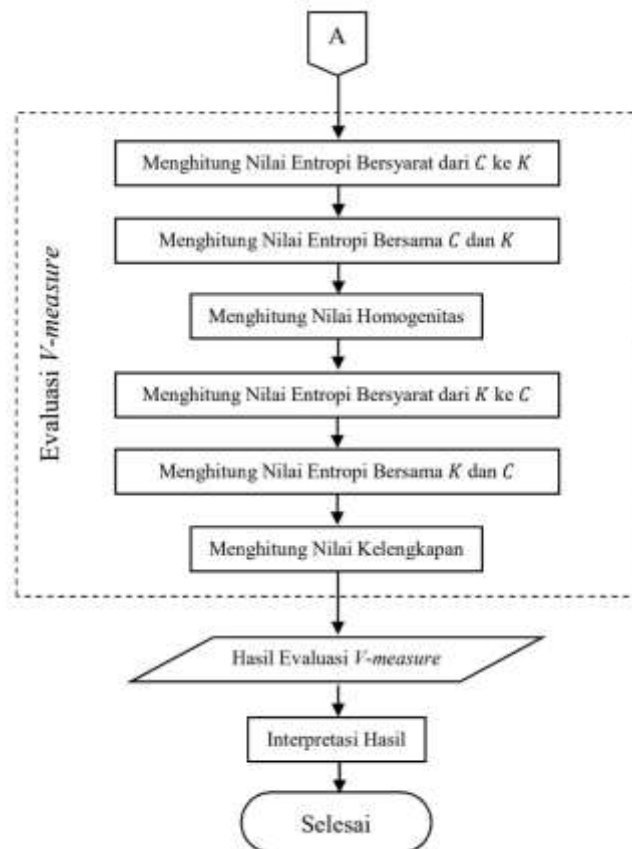
7. Interpretasi hasil evaluasi *V-Measure* terhadap hasil pengelompokan dengan metode *K-Means Clustering*.

3.5 Diagram Alir Analisis Data

Diagram alir berikut menggambarkan langkah-langkah sistematis dalam analisis data implementasi metode *K-Means Clustering* kemudian evaluasi hasil pengelompokan dengan menggunakan metode *V-Measure* yang diawali dari proses *input* data hingga evaluasi hasil pengelompokan.



Gambar 3.1 Diagram Alir Analisis Data



Lanjutan Gambar 3.1 Diagram Alir Analisis Data

BAB IV

HASIL DAN PEMBAHASAN

4.1 Analisis Statistika Deskriptif

Analisis deskriptif digunakan untuk memberikan informasi mengenai karakteristik variabel data yang digunakan secara umum agar lebih mudah dipahami. Data pada penelitian ini disajikan dalam Lampiran 1 yang terdiri dari 21 data PSKS Per Kecamatan dalam wilayah Kabupaten dan Kota Mojokerto dengan 12 indikator yang menjadi variabel dalam penelitian ini, sebagaimana telah disajikan dalam Tabel 3.1. Analisis deskriptif yang disajikan pada Tabel 4.1 terdiri dari rata-rata (*mean*), standar deviasi, nilai minimum, dan nilai maksimum.

Tabel 4.1 Hasil Statistika Deskriptif Variabel Penelitian

Variabel	Rata-rata	Std. Deviasi	Nilai Minimum	Nilai Maksimum
Pekerja Sosial Profesional (x_1)	2.333	4.305	0	18
Pekerja Sosial Masyarakat (x_2)	17.285	18.420	0	68
Taruna Siaga Bencana (x_3)	2.523	2.960	0	9
Lembaga Kesejahteraan Sosial (x_4)	127.952	111.003	0	361
Karang Taruna (x_5)	15.380	4.165	7	20
Lembaga Konsultasi Kesejahteraan Keluarga (x_6)	0.238	0.538	0	2
Keluarga Pioner (x_7)	5.380	8.529	0	28
Wahana Kesejahteraan Sosial Berbasis Masyarakat (x_8)	114.714	128.244	0	563

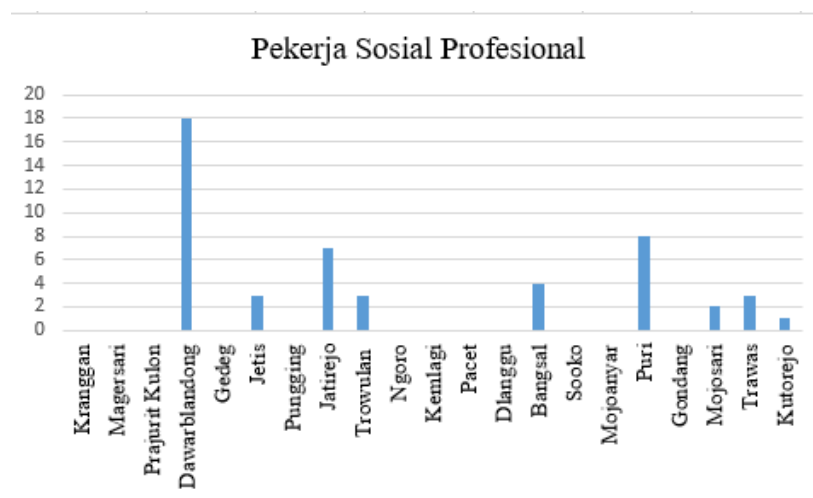
Lanjutan Tabel 4.1 Hasil Statistika Deskriptif Variabel Penelitian

Variabel	Rata-rata	Std. Deviasi	Nilai Minimum	Nilai Maksimum
Wanita Pemimpin Kesejahteraan Sosial (x_9)	29.904	23.058	0	88
Penyuluh Sosial (x_{10})	1.809	2.040	0	6
Tenaga Kesejahteraan Sosial Kecamatan (x_{11})	0.714	0.462	0	1
Dunia Usaha (x_{12})	5.238	7.582	0	25

Berdasarkan Tabel 4.1 hasil analisis statistik mengenai indikator data potensi sumber kesejahteraan sosial diketahui bahwa nilai minimum pada seluruh variabel yaitu variabel x_1 sampai x_{12} sebesar nol, kecuali pada variabel x_5 yang memiliki nilai minimum sebesar tujuh yang dimiliki oleh Kecamatan Kranggan, Magersari, dan Prajurit Kulon. Nilai minimum sebesar nol menunjukkan bahwa pada setiap kecamatan tidak semua memiliki indikator potensi dan sumber kesejahteraan sosial secara lengkap. Nilai minimum sebesar tujuh pada variabel karang taruna (x_5) dipengaruhi oleh banyak desa dalam satu kecamatan, sehingga menunjukkan bahwa tiga kecamatan tersebut memiliki banyak desa yang sama dalam setiap kecamatan.

Pada kolom nilai rata-rata, variabel yang memiliki nilai rata-rata paling rendah sebesar 0.238 adalah variabel lembaga konsultasi kesejahteraan keluarga (x_6) yang mencerminkan bahwa dalam setiap kecamatan, menyediakan lembaga terkait sangat terbatas. Lembaga tersebut hanya tersedia di Kecamatan Jatirejo, Trowulan, dan Kemlagi sebanyak satu lembaga, dan dua lembaga di Kecamatan Sooko, serta tidak ada lembaga sama sekali yang berada di kecamatan lainnya. Kecamatan yang memiliki lembaga konsultasi kesejahteraan keluarga dipengaruhi

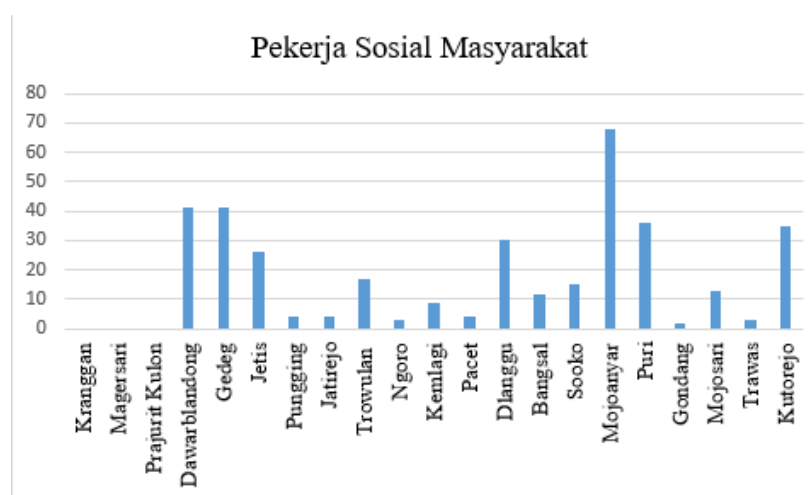
oleh banyaknya masyarakat yang mengunjungi wilayah tersebut dikarenakan adanya sektor wisata, seperti adanya wisata sejarah Kerajaan Majapahit yang terletak di Kecamatan Trowulan. Hal ini berbanding terbalik dengan nilai rata-rata sebesar 127.952 dari variabel lembaga kesejahteraan sosial (x_4), memiliki jangkauan layanan yang lebih luas dan mendalam, mencakup berbagai aspek kehidupan masyarakat, seperti kesehatan, pendidikan, dan ekonomi. Sehingga, lembaga ini otomatis akan lebih banyak tersedia dibandingkan lembaga konsultasi kesejahteraan keluarga (x_6) yang lebih mencakup untuk meningkatkan kesejahteraan dalam lingkup keluarga. Selain itu, deskripsi data dalam bentuk grafik juga dilakukan untuk mempermudah dalam melakukan analisis pada penelitian ini sebanyak 21 data PSKS Per Kecamatan dalam wilayah Kabupaten dan Kota Mojokerto tahun 2023-2024.



Gambar 4.1 Diagram Batang Variabel Pekerja Sosial Profesional (x_1)

Pekerja sosial profesional sebagai indikator pertama dalam mengukur data potensi sumber kesejahteraan sosial memiliki nilai rata-rata sebesar 2.333 dan nilai standar deviasinya sebesar 4.305, sebagaimana didasarkan pada Tabel 4.1. Pekerja sosial profesional di Kabupaten dan Kota Mojokerto divisualisasikan dalam

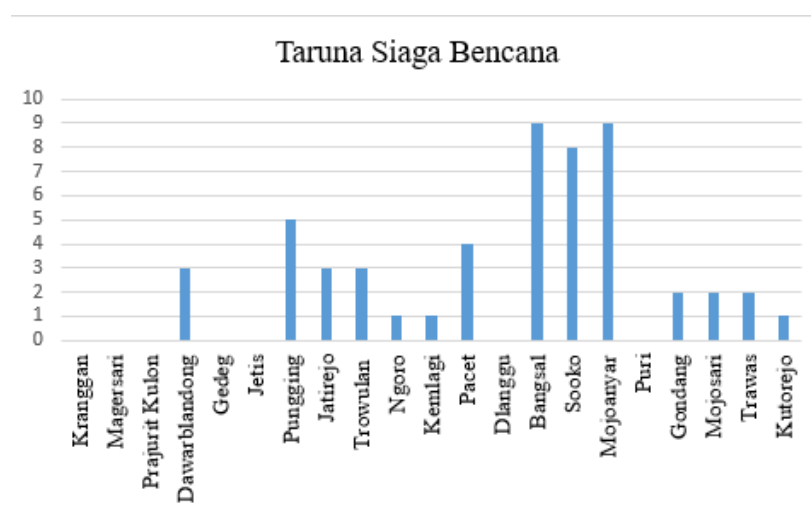
Gambar 4.1. Pada gambar tersebut banyak pekerja sosial profesional terbanyak terletak di Kecamatan Dawarblandong sebesar 18 orang. Hal ini disebabkan karena sebagian wilayah di Kecamatan Dawarblandong digunakan sebagai hutan produksi. Selain itu, wilayah ini berbatasan langsung dengan dua Kabupaten yang lebih besar, sehingga mengalami pembangunan fisik yang cukup pesat dan menjadikan wilayah tersebut mendapat dukungan dan perhatian dari pemerintah, seperti dengan meningkatkan banyak tenaga pekerja sosial profesional. Terdapat 11 kecamatan lainnya yang tidak tersedianya pekerja sosial profesional karena tingkat permasalahan sosial yang rendah sehingga pekerja sosial profesional yang tersedia langsung bersal dari tingkat kabupaten/kota dengan mencakup beberapa wilayah sekaligus, seperti dalam Kecamatan Kranggan, Magersari, dan Prajurit Kulon, pekerja sosial profesional yang tersedia berasal dari pihak Dinas Sosial Kota Mojokerto.



Gambar 4.2 Diagram Batang Variabel Pekerja Sosial Masyarakat (x_2)

Pekerja sosial masyarakat sebagai indikator kedua dalam mengukur data potensi sumber kesejahteraan sosial yang berperan dalam mengidentifikasi dan memanfaatkan sumber daya yang ada di masyarakat untuk mendorong peningkatan

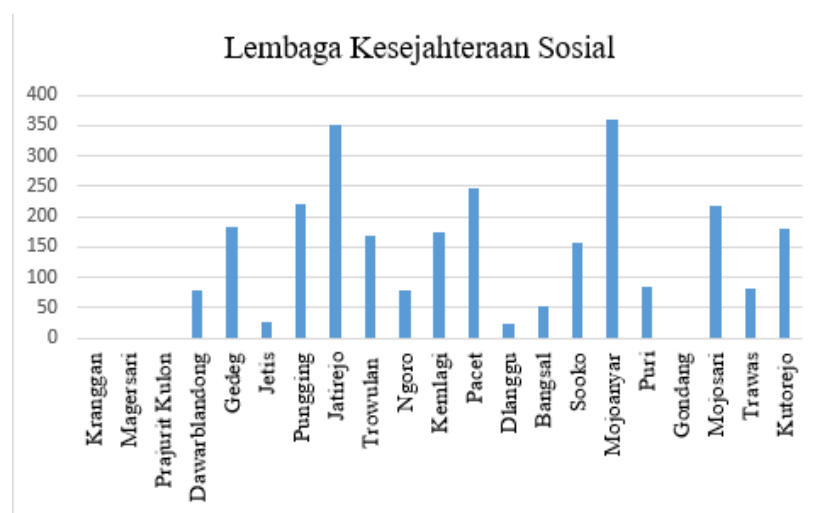
kesejahteraan sosial tingkat desa dan kelurahan. Berdasarkan Tabel 4.2, pekerja sosial masyarakat memiliki nilai rata-rata sebesar 17.285 dengan nilai standar deviasinya sebesar 18.420. Pekerja sosial masyarakat terbanyak sebesar 68 orang terletak di Kecamatan Mojoanyar sebagaimana yang telah disajikan dalam Gambar 4.2, karena wilayah ini memiliki potensi sosial yang kuat. Faktor pendukungnya karena tingginya aksesibilitas seperti adanya kawasan jalan arteri dan berdekatan dengan pusat pemerintah daerah yang memudahkan mobilisasi dan koordinasi antar lembaga dalam dukungan dan upaya pengembangan kesejahteraan sosial. Sedangkan dalam Kecamatan Kranggan, Magersari, dan Prajurit Kulon, pekerja sosial masyarakat diutus langsung dari pemerintah kota bukan atas inisiatif masyarakat yang sukarelawan.



Gambar 4.3 Diagram Batang Variabel Taruna Siaga Bencana (x_3)

Taruna siaga bencana merupakan indikator yang ketiga dalam mengukur data potensi sumber kesejahteraan sosial yang diketahui berdasarkan Tabel 4.1 memiliki nilai rata-rata sebesar 2.523 dan nilai standar deviasinya sebesar 2.960. Taruna siaga bencana berperan dalam mendorong pemanfaatan potensi masyarakat dalam penanggulangan bencana khususnya pada tingkat desa dan kelurahan yang

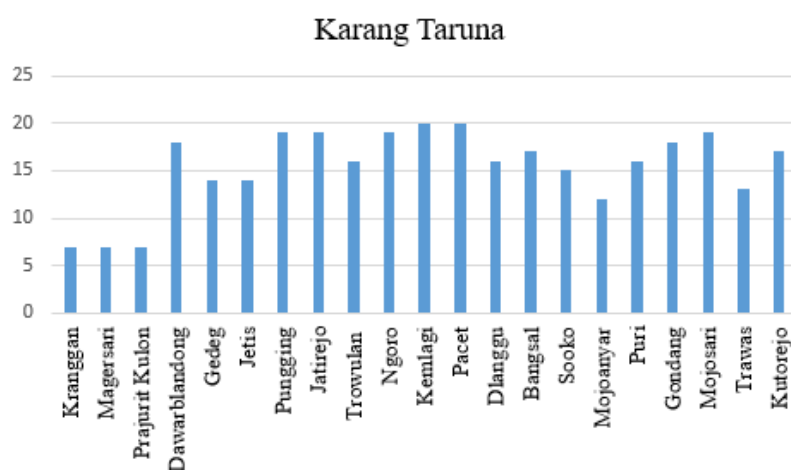
mencakup kegiatan sebelum bencana terjadi, saat keadaan darurat berlangsung, serta dalam tahap pemulihan setelah bencana, termasuk menangani dampak sosial yang ditimbulkan. Taruna siaga bencana dengan jumlah terbanyak pada Kecamatan Bangsal dan Mojoanyar sebanyak sembilan orang. Hal ini dikarenakan banyaknya riwayat bencana alam yang terjadi pada wilayah tersebut seperti pada bencana alam kekeringan pada tahun 2018, Kecamatan Bangsal menjadi wilayah yang mengalami kekeringan paling tinggi di Kabupaten Mojokerto.



Gambar 4.4 Diagram Batang Variabel Lembaga Kesejahteraan Sosial (x_4)

Lembaga kesejahteraan sosial, indikator keempat dalam mengukur data potensi sumber kesejahteraan sosial diketahui berdasarkan Tabel 4.1 memiliki nilai rata-rata sebesar 127.952 dan nilai standar deviasinya sebesar 111.003. Lembaga kesejahteraan sosial berperan dalam memenuhi kebutuhan dasar masyarakat, seperti pangan, sandang, dan papan. Lembaga kesejahteraan sosial meliputi panti asuhan, lembaga kesejahteraan sosial anak (LKSA), organisasi sosial kemasyarakatan, dan lain-lain. Kecamatan Mojoanyar memiliki banyak lembaga kesejahteraan sosial terbanyak dalam wilayah Kabupaten dan Kota Mojokerto

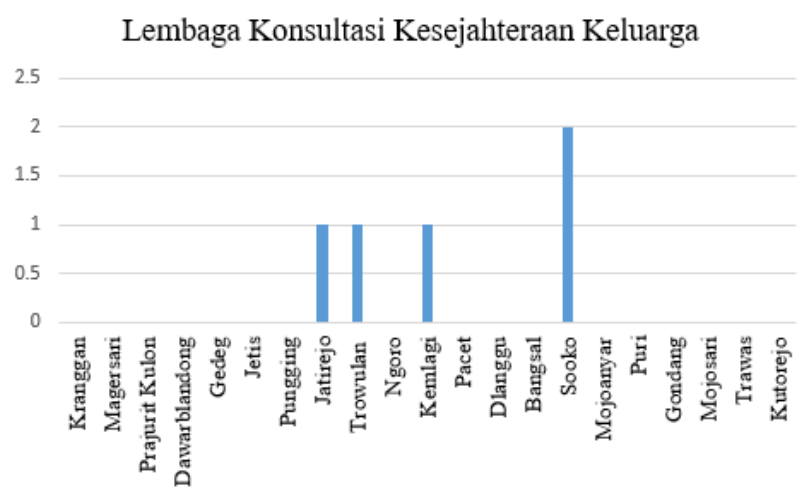
sebanyak 361 unit yang berupa lembaga, yayasan, organisasi, dan bentuk lainnya. Hal ini dipengaruhi oleh letak geografis Mojoanyar yang berada di dekat pusat pemerintahan daerah, peralihan antara wilayah urban dan rural, yang memberikan kemudahan dalam akses koordinasi dan pelayanan sosial. Selain itu, adanya partisipasi aktif dari masyarakat dalam bidang sosial dengan adanya peran tokoh masyarakat, tokoh agama, serta relawan dalam mendirikan dan mengelola berbagai lembaga sosial yang di dukung dan difasilitasi oleh pemerintah daerah.



Gambar 4.5 Diagram Batang Variabel Karang Taruna (x_5)

Karang taruna sebagai indikator kelima dalam mengukur data potensi sumber kesejahteraan sosial diketahui berdasarkan Tabel 4.1 memiliki nilai rata-rata sebesar 15.380 dan nilai nilai standar deviasi sebesar 4.165. Banyak karang taruna yang tersedia dalam setiap kecamatan dipengaruhi oleh banyak desa dalam suatu kecamatan. Gambar 4.5 karang taruna terbanyak dimiliki oleh kecamatan Kemlagi dan Pacet sebanyak 20 kelompok atau organisasi. Hal ini disebabkan karena kedua wilayah tersebut memiliki karakteristik yang didominasi oleh lingkungan pedesaan yang menyebabkan tingkat partisipasi pemuda aktif dalam pencegahan dan

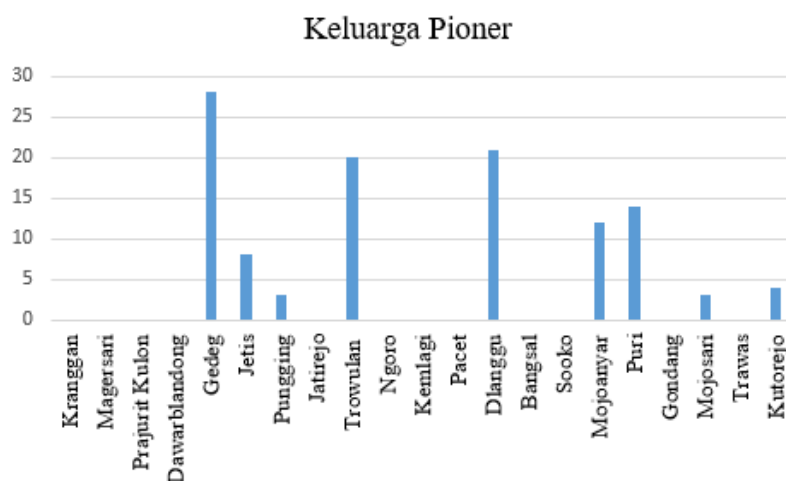
penanggulangan masalah sosial sangat tinggi. Keberadaan karang taruna yang aktif ini mencerminkan kesadaran kolektif pemuda dalam mengambil peran strategis untuk mendukung pembangunan sosial di tingkat desa dan kelurahan. Selain itu, kerja sama antara masyarakat dan pemerintah daerah yang saling berkontribusi dalam mensejahterakan masyarakat kecamatan tersebut. Selama tahun 2023 terjadi penambahan jumlah Rukun Warga (RW) dan Rukun Tetangga (RT).



Gambar 4.6 Diagram Batang Variabel Lembaga Konsultasi Kesejahteraan Keluarga (x_6)

Lembaga konsultasi kesejahteraan keluarga sebagai indikator keenam dalam mengukur data potensi sumber kesejahteraan sosial diketahui berdasarkan Tabel 4.1 memiliki nilai rata-rata sebesar 0.238 dan nilai standar deviasi sebesar 0.538. Lembaga konsultasi kesejahteraan keluarga menangani masalah psikososial keluarga untuk mewujudkan ketahanan keluarga yang mencakup konseling advokasi, dan sebagainya. Lembaga konsultasi kesejahteraan keluarga tersedia hanya pada daerah yang memiliki tingkat perceraian atau tingkat permasalahan dalam keluarga yang cukup tinggi. Kecamatan Sooko memiliki lembaga konsultasi kesejahteraan keluarga sebanyak dua lembaga, dan kecamatan Jatirejo, Trowulan,

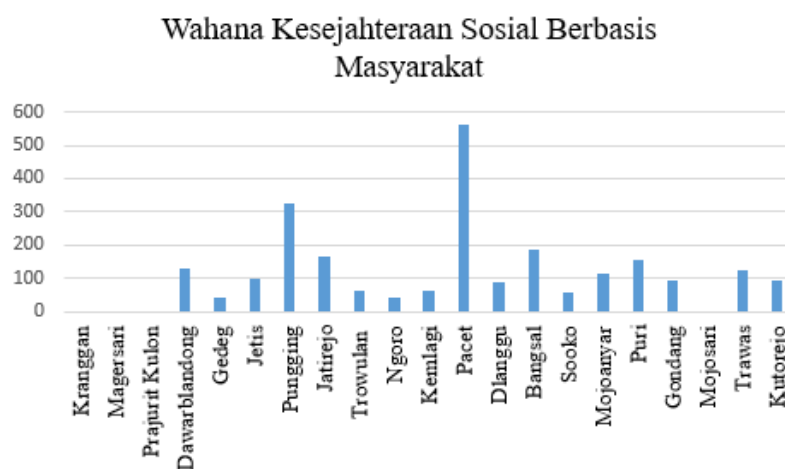
dan Kemlagi sebanyak satu lembaga. Hal ini dikarenakan wilayah yang menyediakan lembaga konsultasi kesejahteraan keluarga memiliki kebutuhan yang lebih besar terhadap layanan konseling dan pendampingan keluarga, seperti tingginya nilai rata-rata banyaknya anggota dalam keluarga. Kecamatan Sooko menjadi kecamatan terpadat pada tahun 2023 dengan kepadatan penduduk mencapai 3.872,8 jiwa per km². Hal ini karena Kecamatan Sooko banyak dibangun perumahan dan daerah instansi-instansi pemerintah Kabupaten Mojokerto.



Gambar 4.7 Diagram Batang Variabel Keluarga Pioner (x_7)

Keluarga pioner sebagai indikator ketujuh dalam mengukur data potensi sumber kesejahteraan sosial diketahui berdasarkan Tabel 4.1 memiliki nilai rata-rata sebesar 5.380 dan nilai standar deviasi sebesar 8.529. Gambar 4.7 menunjukkan bahwa keluarga pioneer terbanyak sebesar 28 keluarga atau unit berasal dari Kecamatan Gedeg. Wilayah ini memiliki 28 keluarga yang mampu mengatasi masalah dalam lingkup keluarga dengan cara-cara efektif dan dapat dijadikan panutan bagi keluarga lainnya, dibuktikan dengan banyaknya pemerlu

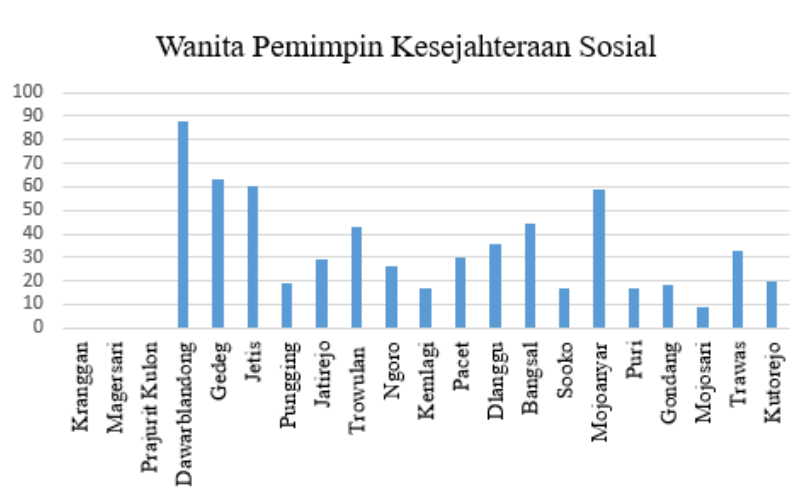
pelayanan kesejahteraan sosial di dalam Kecamatan Gedeg yang sangat sedikit dibandingkan dengan kecamatan yang lainnya. Sedangkan pada beberapa kecamatan lain, yang tidak tersedia keluarga pioner memiliki kualitas ketahanan keluarga dalam mengelola permasalahan sosial dalam keluarga lebih tinggi. Namun tetap diperlukan adanya intervensi lebih lanjut dari pemerintah daerah untuk membina dan membentuk keluarga pioner di wilayah-wilayah tersebut.



Gambar 4.8 Diagram Batang Variabel Wahana Kesejahteraan Sosial Berbasis Masyarakat (x_8)

Wahana kesejahteraan sosial berbasis masyarakat sebagai indikator kedelapan dalam mengukur data potensi sumber kesejahteraan sosial diketahui berdasarkan Tabel 4.1 memiliki nilai rata-rata sebesar 114.714 dan nilai standar deviasi sebesar 128.244. Kecamatan Pacet berdasarkan pada Gambar 4.8 memiliki nilai wahana kesejahteraan sosial berbasis masyarakat yang sangat tinggi mencapai 563 unit, yang terdiri dari Kecamatan Pacet berdasarkan pada Gambar 4.8 memiliki nilai wahana kesejahteraan sosial berbasis masyarakat yang sangat tinggi mencapai 563 unit, yang terdiri dari berbagai jenis lembaga dan kelompok sosial yang aktif mendukung kesejahteraan masyarakat. Wahana ini meliputi organisasi

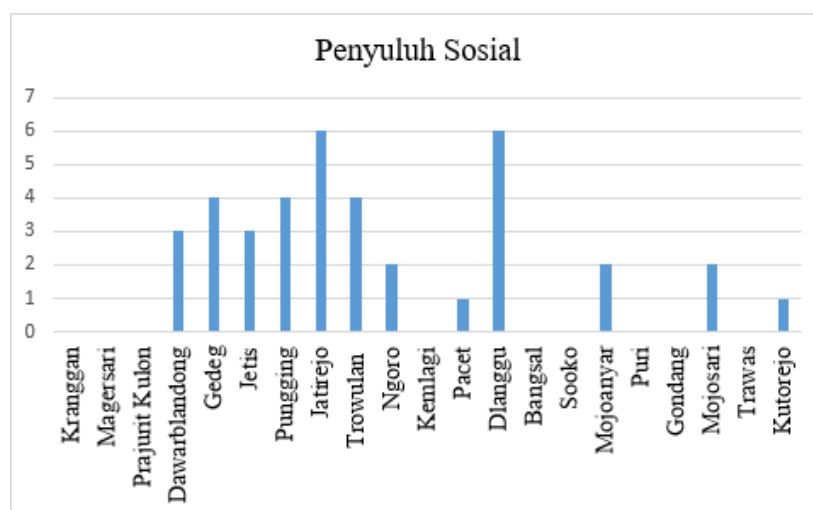
kemasyarakatan seperti posyandu, kelompok PKK, kelompok usaha bersama (KUB), serta fasilitas pelayanan sosial lainnya yang tersebar di seluruh desa dalam Kecamatan Pacet. Keberadaan berbagai wahana ini menunjukkan tingginya partisipasi dan solidaritas sosial masyarakat Pacet dalam meningkatkan kualitas hidup dan menangani berbagai masalah sosial secara mandiri. Hal ini juga mencerminkan efektivitas pembinaan dan pendampingan oleh pemerintah daerah serta organisasi sosial yang berperan dalam mengoptimalkan potensi masyarakat setempat sebagai pilar utama kesejahteraan sosial berbasis komunitas.



Gambar 4.9 Diagram Batang Variabel Wanita Pemimpin Kesejahteraan Sosial (x_9)

Wanita pemimpin kesejahteraan sosial sebagai indikator kesembilan dalam mengukur data potensi sumber kesejahteraan sosial diketahui berdasarkan Tabel 4.1 memiliki nilai rata-rata sebesar 29.904 dan nilai standar deviasi sebesar 23.058. Kecamatan Dawarblandong memiliki banyak wanita pemimpin kesejahteraan sosial terbanyak sebesar 88 orang. Hal ini disebabkan oleh kemampuan daerah tersebut dalam menghasilkan Pendapatan Asli Daerah (PAD) yang cukup besar, yang sebagian besar berasal dari banyaknya usaha perdagangan

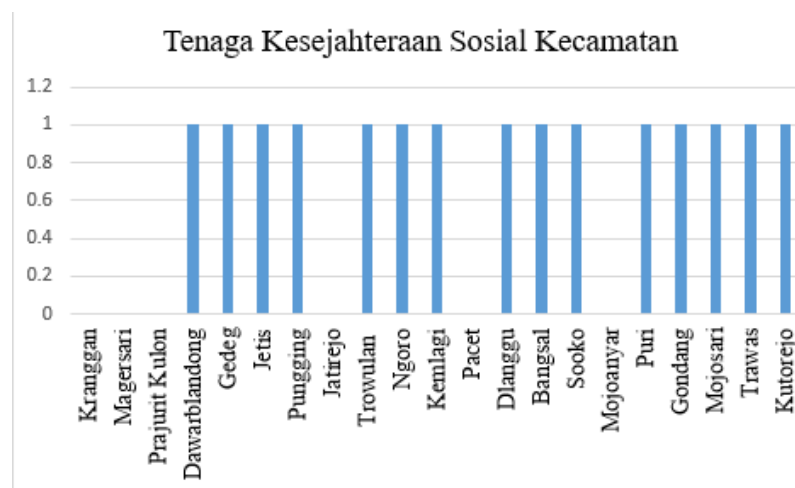
yang dijalankan oleh masyarakat, khususnya para wanita. Keterkaitan erat antara peran wanita sebagai pemimpin kesejahteraan sosial dengan aktivitas ekonomi masyarakat menjadikan mereka sebagai pengelola utama yang berfokus pada pemenuhan kebutuhan para pekerja dan anggota komunitas. Keberadaan wanita pemimpin kesejahteraan sosial yang cukup banyak ini turut berkontribusi positif dalam menjaga kondisi sosial di wilayah tersebut. Peran wanita yang aktif dalam mengorganisasi dan mengatasi berbagai permasalahan kesejahteraan, sehingga menciptakan lingkungan sosial yang lebih stabil dan harmonis.



Gambar 4.10 Diagram Batang Variabel Penyuluh Sosial (x_{10})

Penyuluh sosial sebagai indikator kesepuluh dalam mengukur data potensi sumber kesejahteraan sosial diketahui berdasarkan Tabel 4.1 memiliki nilai rata-rata sebesar 1.809 dengan nilai standar deviasinya sebesar 2.040. Kecamatan Jatirejo dan Dlanggu memiliki individu sebagai penyuluh sosial paling banyak yaitu enam orang. Hal ini menunjukkan bahwa penyuluh sosial yang relatif lebih banyak ini menunjukkan peran penting mereka dalam memberikan edukasi dan bimbingan kepada masyarakat, terutama dalam hal adaptasi terhadap perubahan yang terjadi di berbagai aspek kehidupan sosial dan ekonomi. Penyuluh sosial berfungsi sebagai

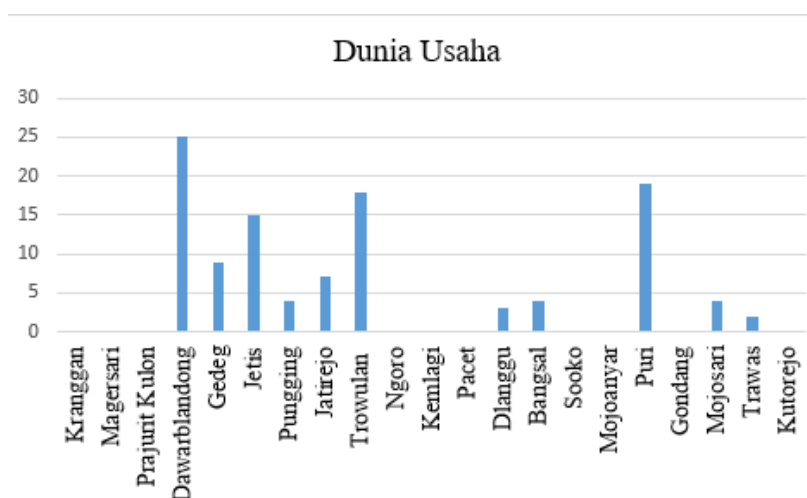
penghubung antara pemerintah dan masyarakat dalam menyampaikan informasi, serta membantu masyarakat mengatasi berbagai permasalahan sosial dengan pendekatan yang tepat. Sedangkan kecamatan-kecamatan lain yang tidak tersedia penyuluh sosial secara langsung kemungkinan besar menerima layanan penyuluhan sosial dari pemerintah di tingkat kabupaten atau kota. Hal ini mengindikasikan adanya distribusi sumber daya manusia penyuluh sosial yang bersifat terpusat dengan mekanisme penyuluhan dari tingkat yang lebih tinggi dan dapat menjangkau masyarakat secara menyeluruh.



Gambar 4.11 Diagram Batang Variabel Tenaga Kesejahteraan Sosial Berbasis Masyarakat (x_{11})

Tenaga kesejahteraan sosial kecamatan sebagai indikator kesebelas dalam mengukur data potensi sumber kesejahteraan sosial diketahui berdasarkan Tabel 4.1 memiliki nilai rata-rata sebesar 0.714 dengan nilai standar deviasinya sebesar 0.462. Kecamatan Dawarblandong, Gedeg, Jetis, Pungging, Trowulan, Ngoro, Kemlagi, Dlanggu, Bangsal, Sooko, Puri, Gondang, Mojosari, Trawas, dan Kutorejo masing-masing memiliki tenaga kesejahteraan sosial sebanyak satu orang. Hal ini menunjukkan bahwa setiap kecamatan memiliki seorang koordinator yang

dapat membantu dalam proses penyelenggaraan layanan sosial, bantuan kesejahteraan, dan sebagainya. Pada Kecamatan Kranggan, Magersari, dan Prajurit Kulon, tenaga kesejahteraan sosial kecamatan yang tersedia sudah dilakukan secara terpusat di tingkat pemerintah kota. Hal ini menunjukkan adanya variasi dalam sistem pengelolaan tenaga kesejahteraan sosial yang disesuaikan dengan kebutuhan dan kebijakan daerah masing-masing.



Gambar 4.12 Diagram Batang Variabel Dunia Usaha (x_{12})

Dunia usaha sebagai indikator kedua belas dalam mengukur data potensi sumber kesejahteraan sosial diketahui berdasarkan Tabel 4.1 memiliki nilai rata-rata sebesar 5.238 dengan nilai standar deviasinya sebesar 7.582. Kecamatan Dawarblandong berdasarkan Gambar 4.12 memiliki nilai dunia usaha paling tinggi yaitu sebesar 25 unit usaha yang berbentuk organisasi, lembaga, maupun kelompok. Keberadaan dunia usaha yang cukup besar ini mencerminkan dinamika ekonomi yang aktif di Kecamatan Dawarblandong, yang turut berkontribusi pada peningkatan kesejahteraan sosial masyarakat setempat. Berbagai unit usaha tersebut tidak hanya berperan sebagai sumber pendapatan, tetapi juga sebagai wadah

pemberdayaan masyarakat melalui penciptaan lapangan kerja dan pengembangan kapasitas usaha lokal.

4.2 Standardisasi Data

Penelitian ini menggunakan data dengan skala yang sama, tetapi dengan satuan yang berbeda, mengakibatkan standardisasi data perlu dilakukan untuk mempersempit distribusi data agar tidak mempengaruhi proses pengelompokan. Pada penelitian ini, semua data diubah menjadi *z-score* untuk menyamakan satuan data sehingga memiliki nilai mean nol dan standar deviasi nol. Nilai *z-score* diperoleh dengan mengurangi nilai data dengan nilai rata-rata dan kemudian dibagi dengan nilai standar deviasi. Hasil perhitungan rata-rata dan standar deviasi dari setiap variabel data telah dilakukan, selengkapnya telah disajikan pada tabel 4.1. Berikut merupakan contoh perhitungan manual untuk memperoleh nilai *z-score* menggunakan rumus (2.4).

Nilai *z-score* data ke-1 variabel 1 (z_{11}) dan data ke-1 variabel 2 (z_{12})

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

$$z_{11} = \frac{0 - 2.333}{4.305} = -0.542$$

$$z_{12} = \frac{0 - 17.285}{18.420} = -0.938$$

Hasil standardisasi data yang telah dilakukan dengan menggunakan nilai *z-score*, selengkapnya dapat disajikan pada Tabel 4.2 dan Lampiran 2.

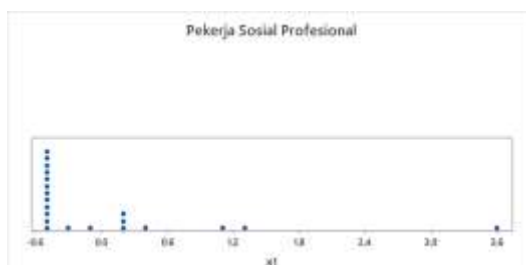
Tabel 4.2 Tabel Data Standardisasi

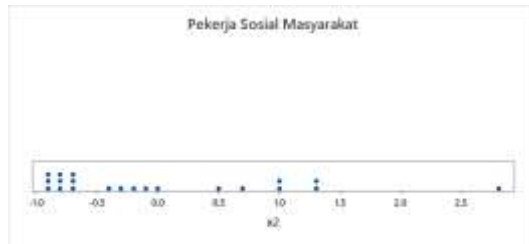
No	Kecamatan	x_1	x_2	...	x_{11}	x_{12}
1	Kranggan	-0.542	-0.938	...	-1.543	-0.691
2	Magersari	-0.542	-0.938	...	-1.543	-0.691
3	Prajurit Kulon	-0.542	-0.938	...	-1.543	-0.691
4	Dawarblandong	3.639	1.287	...	0.617	2.606

Lanjutan Tabel 4.2 Tabel Data Standardisasi

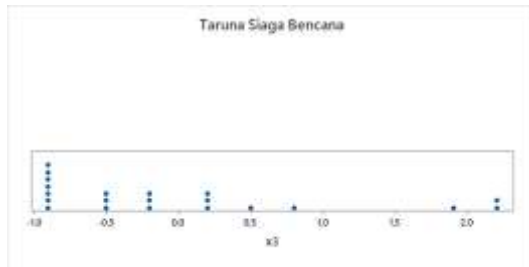
No	Kecamatan	x_1	x_2	...	x_{11}	x_{12}
5	Gedeg	-0.542	1.287	...	0.617	0.496
6	Jetis	0.155	0.473	...	0.617	1.287
7	Pungging	-0.542	-0.721	...	0.617	-0.163
8	Jatirejo	1.084	-0.721	...	-1.543	0.232
9	Trowulan	0.155	-0.016	...	0.617	1.683
10	Ngoro	-0.542	-0.776	...	0.617	-0.691
11	Kemlagi	-0.542	-0.450	...	0.617	-0.691
12	Pacet	-0.542	-0.721	...	-1.543	-0.691
13	Dlanggu	-0.542	0.690	...	0.617	-0.295
14	Bangsals	0.387	-0.287	...	0.617	-0.163
15	Soko	-0.542	-0.124	...	0.617	-0.691
16	Mojoanyar	-0.542	2.753	...	-1.543	-0.691
17	Puri	1.316	1.016	...	0.617	1.815
18	Gondang	-0.542	-0.830	...	0.617	-0.691
19	Mojosari	-0.077	-0.233	...	0.617	-0.163
20	Trawas	0.155	-0.776	...	0.617	-0.427
21	Kutorejo	-0.310	0.962	...	0.617	-0.691

Berdasarkan Tabel 4.2 yang merupakan data penelitian yang telah dilakukan standardisasi dengan *z-score*, selanjutnya dilakukan visualisasi data untuk setiap variabel guna mendeteksi adanya nilai *outlier*. Visualisasi ini disajikan pada Gambar 4.13 hingga Gambar 4.24. Setiap gambar merepresentasikan distribusi masing-masing variabel dalam bentuk grafik yang dapat memudahkan dalam mengidentifikasi nilai-nilai *outlier* dalam data penelitian.

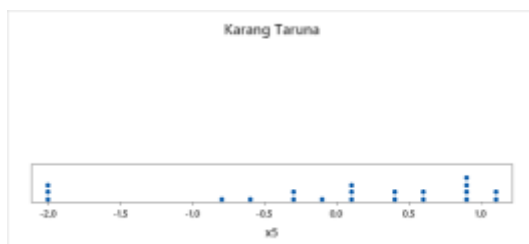
Gambar 4.13 Dotplot Variabel x_1 Setelah Standardisasi



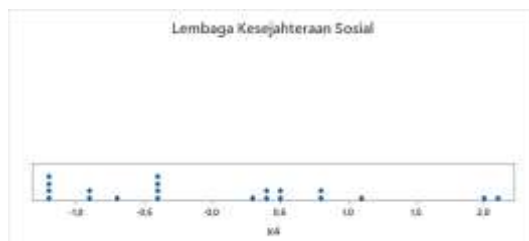
Gambar 4.14 Dotplot Variabel x_2 Setelah Standardisasi



Gambar 4.15 Dotplot Variabel x_3 Setelah Standardisasi



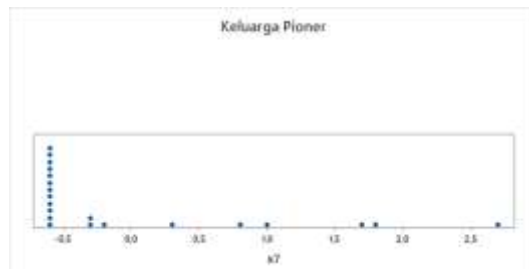
Gambar 4.16 Dotplot Variabel x_4 Setelah Standardisasi



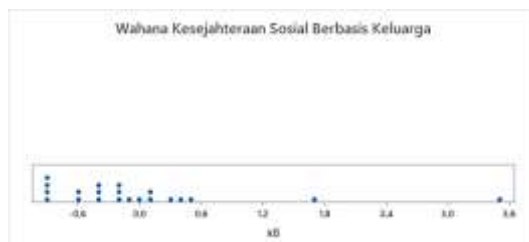
Gambar 4.17 Dotplot Variabel x_5 Setelah Standardisasi



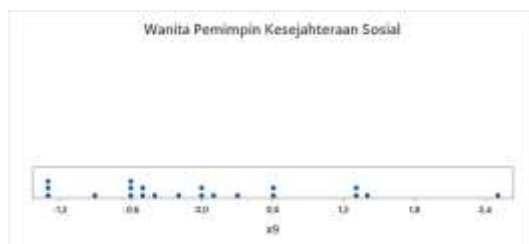
Gambar 4.18 Dotplot Variabel x_6 Setelah Standardisasi



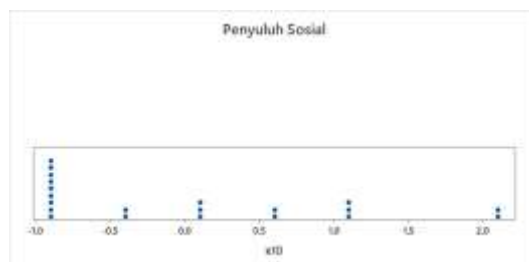
Gambar 4.19 Dotplot Variabel x_7 Setelah Standardisasi



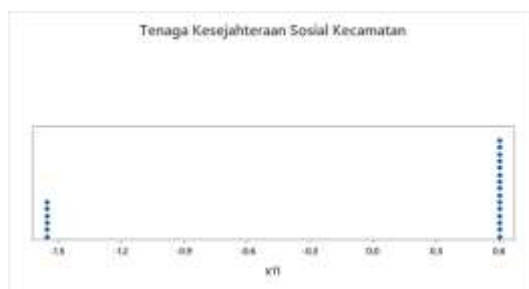
Gambar 4.20 Dotplot Variabel x_8 Setelah Standardisasi



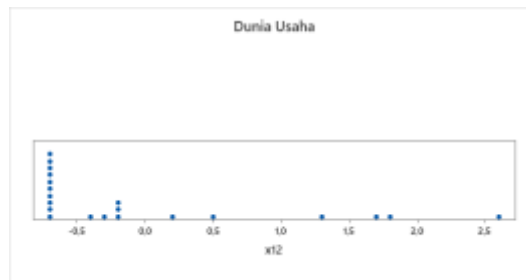
Gambar 4.21 Dotplot Variabel x_9 Setelah Standardisasi



Gambar 4.22 Dotplot Variabel x_{10} Setelah Standardisasi



Gambar 4.23 Dotplot Variabel x_{11} Setelah Standardisasi



Gambar 4.24 Dotplot Variabel x_{12} Setelah Standardisasi

Hasil visualisasi pada Gambar 4.13 hingga Gambar 4.24 menunjukkan bahwa terdapat beberapa variabel yang memiliki nilai *outlier*. Keberadaan *outlier* ini teridentifikasi melalui penyimpangan nilai yang signifikan dari pola distribusi data yang umum. Oleh karena itu, pada tahap selanjutnya dilakukan proses identifikasi nilai *outlier* karena nilai-nilai ekstrem tersebut berpotensi memengaruhi hasil analisis, khususnya dalam proses pengelompokan.

4.3 Deteksi Outlier

Sebelum melakukan proses selanjutnya, data yang telah distandardisasi perlu dilakukan deteksi data terhadap *outlier* untuk mengidentifikasi adanya nilai pencilan dalam distribusi data. Jika terdeteksi terdapat *outlier*, maka terdapat data yang menyimpang secara signifikan dari pola umum dalam kumpulan data. Pengecekan *outlier* data dilakukan dengan melihat *z-score*, dengan suatu data dikategorikan sebagai nilai *outlier* jika nilai *z-score* kurang dari -3.00 atau lebih dari 3.00 .

Berdasarkan nilai *z-score* yang disajikan pada Tabel 4.2 yang kemudian divisualisasikan dalam Gambar 4.13 sampai Gambar 4.24, dari seluruh data yang digunakan dalam penelitian, terdapat tiga data yang terdeteksi *outlier*. Nilai *outlier* yang pertama terletak pada variabel x_1 yakni variabel pekerja sosial profesional,

dari Kecamatan Dawarblandong sebesar 3.639 yang lebih dari 3.00 setelah dilakukan proses standardisasi dengan *z-score*. Nilai ini menunjukkan adanya kebutuhan pekerja sosial profesional untuk Kecamatan Dawarblandong lebih tinggi dibandingkan dengan kecamatan lainnya. Hal ini disebabkan oleh karakteristik wilayah Kecamatan Dawarblandong yang sebagian besar menjadi kawasan hutan industri yang dikelola oleh pemerintah, dengan akses pengawasan dan mobilisasi yang relatif jauh dari pusat pemerintah daerah. Nilai *outlier* yang kedua terletak pada variabel x_8 yakni variabel wahana kesejahteraan sosial berbasis masyarakat dari Kecamatan Pacet, yang memiliki nilai sebesar 3.496 yang lebih dari 3.00 setelah dilakukan proses standardisasi menggunakan nilai *z-score*. Nilai ini menunjukkan adanya wahana kesejahteraan sosial yang berbasis masyarakat untuk Kecamatan Pacet lebih dominan dibandingkan dengan kecamatan lainnya. Hal ini disebabkan karena adanya solidaritas dan partisipasi masyarakat yang tinggi dalam meningkatkan kualitas daerah yang didukung dan ditunjang oleh pemerintah daerah. Sehingga, dengan mempertimbangkan kondisi dan karakteristik khusus dari wilayah tersebut, data *outlier* ini tetap dipertahankan karena mencerminkan realitas kebutuhan lapangan yang valid dan signifikan. Penghapusan data ini justru berisiko menghilangkan informasi penting dan dapat memengaruhi hasil analisis secara keseluruhan.

Selain itu, variabel lembaga konsultasi kesejahteraan keluarga (x_6) tidak digunakan dalam analisis lebih lanjut karena fungsi dan peran dari lembaga konsultasi ini secara substantif telah tercermin dalam dua variabel lain, yaitu variabel lembaga kesejahteraan sosial (x_3) dan variabel keluarga pioneer (x_7). Penghapusan variabel lembaga konsultasi kesejahteraan keluarga (x_6) dilakukan

untuk menghindari redundansi data antar variabel yang memiliki keterkaitan konsep, serta untuk menjaga efisiensi dan kestabilan model analisis karena kontribusi informatif variabel lembaga konsultasi kesejahteraan keluarga (x_6) yang relatif rendah.

4.4 Uji Multikolinieritas

Uji multikolinieritas dilakukan untuk menguji adanya kesamaan yang dimiliki oleh setiap variabel dalam data penelitian. Nilai yang dapat digunakan untuk mengetahui adanya multikolinieritas adalah dengan menentukan nilai *Variance Inflation Factor* (VIF) yang melibatkan lebih dari dua variabel bebas dengan menyimpulkan hasilnya. Berikut adalah contoh perhitungan manual untuk memperoleh nilai VIF menggunakan persamaan (2.5).

1. Hipotesis masalah

H_0 : Tidak terdapat multikolinieritas

H_1 : Terdapat multikolinieritas

2. Statistik Uji VIF

Variabel pekerja sosial profesional (x_1)

$$VIF_1 = \frac{1}{1 - \frac{17.699}{20}} = \frac{1}{1 - 0.885} = 8.695$$

Nilai VIF yang diperoleh dari perhitungan menggunakan persamaan (2.5), untuk variabel lainya disajikan dalam Tabel 4.3.

3. Daerah Kritis

Nilai $VIF \leq 10$ maka terima H_0

Nilai $VIF > 10$ maka tolak H_0

4. Keputusan:

Berdasarkan Tabel 4.3 nilai VIF dari seluruh variabel data penelitian kurang dari 10, maka terima H_0 yang menunjukkan tidak terjadi multikolinieritas dari seluruh variabel dan semua variabel dapat digunakan dalam proses analisis selanjutnya.

Tabel 4.3 Hasil Nilai *Variance Inflation Factor* (VIF)

Variabel	SSR	SST	R^2	VIF	Keputusan
x_1	17.699	20	0.885	8.695	Tidak Terjadi Multikolinieritas
x_2	15.489	20	0.774	4.434	Tidak Terjadi Multikolinieritas
x_3	10.315	20	0.516	2.065	Tidak Terjadi Multikolinieritas
x_4	14.775	20	0.739	3.827	Tidak Terjadi Multikolinieritas
x_5	16.715	20	0.722	0.836	Tidak Terjadi Multikolinieritas
x_7	16.948	20	0.532	0.848	Tidak Terjadi Multikolinieritas
x_8	11.530	20	0.577	2.361	Tidak Terjadi Multikolinieritas
x_9	14.611	20	0.731	3.711	Tidak Terjadi Multikolinieritas
x_{10}	13.022	20	0.651	2.866	Tidak Terjadi Multikolinieritas
x_{11}	15.769	20	0.788	4.728	Tidak Terjadi Multikolinieritas
x_{12}	17.653	20	0.883	8.520	Tidak Terjadi Multikolinieritas

4.5 Penerapan *Elbow Method Criterion*

Menentukan banyak kelompok optimal yang akan digunakan dalam *K-Means Clustering* menggunakan metode *Elbow*. Metode ini menentukan banyaknya kelompok terbaik yang dapat dibentuk dalam grafik dari nilai *Sum of Square Error*. Titik optimal k ditentukan dari grafik SSE yang membentuk siku atau yang memiliki sudut paling kecil. Titik optimal ini menunjukkan banyak kelompok yang memberikan keseimbangan terbaik antara kompleksitas model dengan tingkat kesalahan.

Berikut perhitungan nilai SSE menggunakan rumus (2.7) terhadap data dalam Tabel 4.2 untuk nilai k pertama yang diawali dengan menghitung $\vec{v}_h$ sebagai

centroid yang diperoleh dari nilai rata-rata setiap data standardisasi dengan rumus pada persamaan (2.2).

$$\vec{v}_h = \frac{\sum_{i=1}^{n_h} \vec{x}_i}{n_h}.$$

Nilai rata-rata *z-score* variabel pekerja sosial profesional ($\vec{v}_1$)

$$\vec{v}_1 = \frac{(-0.542)+(-0.542)+(-0.542)+\dots+(0.155)+(-0.310)}{21} = 0$$

Menghitung nilai *Sum of Square Error* (SSE) untuk $k = 1$ menggunakan persamaan (2.7).

$$\begin{aligned} SSE &= \sum_{h=1}^k \sum_{i=1}^{c_h} (\vec{x}_i - \vec{v}_h)^2 \\ &= (-0.542 - 0)^2 + (-0.938 - 0)^2 + \dots + (0.617 - 0)^2 + (-0.691 - 0)^2 = 220.000 \end{aligned}$$

Hasil perhitungan manual SSE secara lebih rinci disajikan pada Tabel 4.4.

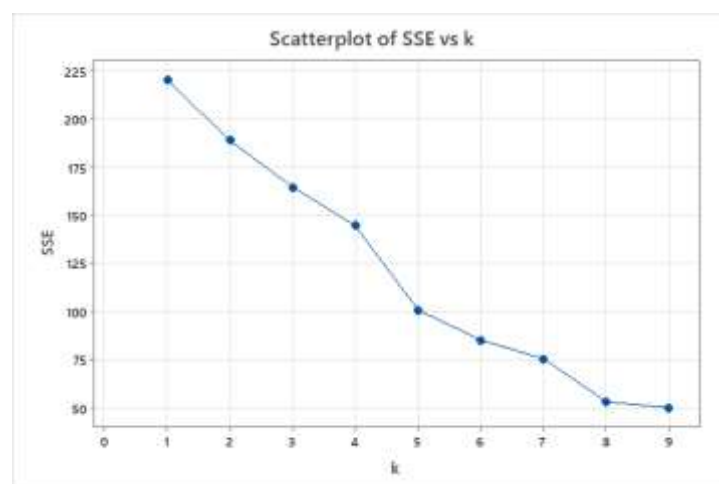
Tabel 4.4 Nilai *Sum of Square Error* (SSE)

<i>k</i>	SSE	Selisih
1	220.000	0
2	188.731	31.269
3	164.613	24.118
4	144.517	20.096
5	100.872	43.645
6	85.376	15.4965
7	75.651	9.7246
8	53.414	22.237

Nilai SSE pada Tabel 4.4 yang merupakan jumlah jarak kuadrat antara titik dalam kelompok dan *centroid* direpresentasikan ke dalam grafik visualisasi metode *Elbow* pada Gambar 4.25. Grafik tersebut pada sumbu *y* menunjukkan nilai WCSS/nilai SSE dan pada sumbu *x* menunjukkan nilai *k*. Pada Gambar 4.25, dapat dilihat dan dipilih bahwa nilai *k* yang dapat ditentukan terletak pada titik siku yang terbentuk, dimana laju penurunan nilai SSE berubah secara drastis, jika terjadi penambahan nilai *k* (diluar titik siku) akan menyebabkan *overfitting* dan jika

mengurangi nilai k tidak akan efektif dalam mencakup variabilitas data. Gambar 4.13 yang menunjukkan adanya penurunan yang paling signifikan terletak pada $k = 3$ dengan nilai selisih SSE sebesar 24.118, namun pada penurunan tersebut kurang optimal dikarenakan grafik yang tidak membentuk pola *elbow*, menyiku atau tidak berbentuk siku.

Setelah nilai $k = 3$, yaitu pada nilai $k = 4$ masih terdapat adanya penurunan grafik yang signifikan. Namun, penurunan tersebut mulai melandai dan membentuk sudut menyerupai siku dengan besar sudut yang paling kecil pada $k = 5$. Sehingga dapat diperoleh nilai k yang optimal yang dapat ditentukan dalam penelitian ini adalah $k = 5$ dengan nilai SSE sebesar 43.645. Setelah nilai $k = 5$, perubahan nilai SSE yang mengalami penurunan menjadi lebih kecil menunjukkan bahwa menambahkan kelompok lebih banyak tidak memberikan peningkatan signifikan dalam data.



Gambar 4.25 Grafik Metode *Elbow* pada Nilai SSE

4.6 Penerapan Metode *K-Means Clustering*

Data yang digunakan dalam proses pengelompokan dengan metode *K-Means Clustering* tersaji dalam Tabel 4.2. Proses *K-Means Clustering* diawali dengan

menentukan banyaknya jumlah kelompok (k) terbaik yang dapat dibentuk menggunakan metode *Elbow* berdasarkan persamaan (2.7). Hasil dari proses analisis tersebut, memperoleh nilai yang telah tersaji dalam Tabel 4.4 dan menunjukkan bahwa banyak kelompok (k) terbaik yang dapat terbentuk adalah lima kelompok ($k = 5$).

Kemudian dilakukan penentuan titik pusat awal dari kelompok yang disebut *centroid* secara acak. *Centroid* yang dapat dipilih sebanyak k optimal yang telah ditentukan, sebanyak lima titik pada setiap data. Berikut nilai *centroid* awal yang diambil dari data PSKS di wilayah Kecamatan Kranggan, Kecamatan Gedeg, Kecamatan Dlanggu, Kecamatan Gondang, dan Kecamatan Kutorejo yang disajikan dalam Tabel 4.5.

Tabel 4.5 Nilai *Centroid* Awal untuk Iterasi ke-1

Titik Pusat	x_1	x_2	x_3	x_4	x_5	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
$C1$	-0.542	-0.938	-0.853	-1.153	-2.012	-0.631	-0.894	-1.297	-0.887	-1.543	-0.691
$C2$	-0.542	1.287	-0.853	0.496	-0.332	2.652	-0.567	1.435	1.074	0.617	0.496
$C3$	-0.542	0.690	-0.853	-0.936	0.149	1.831	-0.185	0.264	2.054	0.617	-0.295
$C4$	-0.542	-0.830	-0.177	-1.153	0.629	-0.631	-0.169	-0.516	-0.887	0.617	-0.691
$C5$	-0.310	0.962	-0.515	0.478	0.389	-0.162	-0.154	-0.430	-0.397	0.617	-0.691

Tahap berikutnya yaitu menghitung jarak antara setiap data dengan *centroid* kelompok menggunakan metode *Euclidean distance* pada rumus (2.3). Berikut adalah contoh perhitungan jarak *Euclidean* untuk data ke-1 (Kecamatan Kranggan) terhadap *centroid* di setiap kelompok.

$$d_{1h} = \sqrt{\sum_{h=1}^m (x_{1j} - v_{hj})^2}$$

Jarak data ke-1 terhadap *centroid* kelompok ke-1 ($C1$)

$$d_{11} = \sqrt{(-0.542 - (-0.542))^2 + (-0.542 - (-0.938))^2 + \dots + (-0.691 - (-0.691))^2} = 0$$

Jarak data ke-1 terhadap *centroid* kelompok ke-2 (C2)

$$d_{12} = \sqrt{(-0.542 - (-0.542))^2 + (-0.542 - 1.287)^2 + \dots + (0.775 - (-0.691))^2} = 7.223$$

Jarak data ke-1 terhadap *centroid* kelompok ke-3 (C3)

$$d_{13} = \sqrt{(-0.542 - (-0.542))^2 + (-0.542 - 0.690)^2 + \dots + (0.775 - 0.496)^2} = 5.638$$

Jarak data ke-1 terhadap *centroid* kelompok ke-4 (C4)

$$d_{14} = \sqrt{(-0.542 - (-0.542))^2 + (-0.542 - (-0.830))^2 + \dots + (0.775 + 0.737)^2} = 3.454$$

Jarak data ke-1 terhadap *centroid* kelompok ke-5 (C5)

$$d_{15} = \sqrt{(-0.542 - (-0.310))^2 + (-0.542 - 0.962)^2 + \dots + (0.775 + 0.453)^2} = 3.906$$

Setelah menghitung jarak *Euclidean*, diperoleh nilai jarak setiap data terhadap *centroid* yang disajikan dalam Tabel 4.6.

Tabel 4.6 Hasil Perhitungan Jarak *Euclidean* Iterasi ke-1

No	C1	C2	C3	C4	C5	Jarak Minimum	Kelompok
1	0.000	6.226	5.463	3.639	4.316	0.000	1
2	0.000	6.226	5.463	3.639	4.316	0.000	1
3	0.000	6.226	5.463	3.639	4.316	0.000	1
4	8.014	6.115	6.416	6.702	6.153	6.115	2
5	6.226	0.000	2.533	5.346	3.991	0.000	2
6	4.932	3.113	2.973	3.728	3.334	2.973	3
7	5.588	5.104	4.224	3.550	3.286	3.286	5
8	5.896	5.568	5.128	5.248	4.472	4.472	5
9	5.719	2.556	2.952	4.471	3.750	2.556	2
10	3.994	4.734	3.664	1.386	2.201	1.386	4
11	4.241	5.025	4.446	1.735	1.765	1.735	4
12	6.146	6.770	6.268	4.925	4.814	4.814	5
13	5.463	2.533	0.000	4.304	3.594	0.000	3
14	5.214	5.610	5.155	3.015	3.607	3.015	4
15	4.366	5.426	5.005	2.677	2.767	2.677	4
16	6.720	5.009	5.869	6.471	4.828	4.828	5
17	5.312	4.199	4.310	4.166	3.426	4.166	4
18	3.639	5.346	4.304	0.000	2.562	0.000	4
19	4.421	4.484	3.914	2.562	1.803	1.803	5
20	3.361	4.866	4.323	1.736	2.446	1.736	4
21	4.316	3.991	3.594	2.562	0.000	0.000	5

Dalam Tabel 4.6 kemudian dipilih nilai yang terkecil atau nilai jarak minimum dari setiap kolom $C1, C2, C3, C4$, dan $C5$ untuk setiap data, kemudian ditentukan penamaan kelompok atau golongan kelompok berdasarkan kolom dari nilai minimum yang diperoleh. Diketahui dari Tabel 4.6, pada data ke-1 Kecamatan Kranggan nilai jarak *Euclidean* data terhadap *centroid* ke-1 ($C1$), ke-2 ($C2$), ke-3 ($C3$), ke-4 ($C4$), dan ke-5 ($C5$) secara berurutan sebesar 0.000, 6.226, 5.463, 3.639, dan 4.316. Berdasarkan nilai-nilai jarak tersebut, diperoleh nilai jarak minimum untuk data ke-1 Kecamatan Kranggan terhadap *centroid* yaitu sebesar 0.000, jarak data ke-1 terhadap *centroid* 1 (d_{11}). Sehingga dapat diambil kesimpulan, bahwa data ke-1, Kecamatan Kranggan, masuk dalam kelompok ke-1. Tahapan pengelompokan tersebut dilakukan hingga data ke-21, Kecamatan Kutorejo, memiliki kelompok masing-masing.

Kemudian dilakukan perulangan penentuan *centroid* kelompok dengan rumus (2.2) untuk melakukan iterasi selanjutnya hingga semua data memiliki kelompok masing-masing dan tidak mengalami perubahan anggota kelompok dalam setiap kelompok. Pada setiap iterasi, penentuan *centroid* dilakukan dengan mencari nilai rata-rata, yang diperoleh dengan cara menjumlahkan semua nilai pada tiap kelompok, lalu dibagi dengan total data pada kelompok tersebut. Setelah memperoleh nilai rata-rata tersebut, kemudian menghitung nilai *centroid* baru dengan menggunakan data pada Tabel 4.2. Berikut adalah contoh perhitungan nilai *centroid* baru untuk iterasi ke-2 pada setiap kelompok menggunakan rumus (2.2).

1. *Centroid* ke-2 pada kelompok ke-1:

$$\hat{v}_1 = \frac{(-0.542) + (-0.542) + (-0.542)}{3} = -0.542$$

2. *Centroid* ke-2 pada kelompok ke-2:

$$\vec{v}_2 = \frac{3.639 + (-0.542) + 0.155}{3} = 1.084$$

3. *Centroid* ke-2 pada kelompok ke-3:

$$\vec{v}_3 = \frac{0.155 - 0.542}{2} = -0.194$$

4. *Centroid* ke-2 pada kelompok ke-4:

$$\vec{v}_4 = \frac{(-0.542) + (-0.542) + 0.387 + (-0.542) + 1.316 + (-0.542) + 0.155}{7} = -0.044$$

5. *Centroid* ke-2 pada kelompok ke-5:

$$\vec{v}_{51} = \frac{(-0.542) + 1.084 + (-0.542) + (-0.542) + (-0.077) + (-0.310)}{6} = -0.155$$

Berdasarkan perhitungan pada rumus (2.2) untuk menentukan *centroid* baru, diperoleh hasil pada Tabel 4.7, yang merupakan nilai *centroid* baru untuk iterasi ke-2 pada setiap kelompok.

Tabel 4.7 Nilai *Centroid* Baru untuk Iterasi ke-2

Titik Pusat	x_1	x_2	x_3	x_4	x_5	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
$C1$	-0.542	-0.938	-0.853	-1.153	-2.012	-0.631	-0.894	-1.297	-0.887	-1.543	-0.691
$C2$	1.084	0.853	-0.177	0.142	0.149	1.245	-0.284	1.508	0.910	0.617	1.595
$C3$	-0.194	0.582	-0.853	-0.923	-0.091	1.069	-0.162	0.785	1.319	0.617	0.496
$C4$	-0.044	-0.318	0.257	-0.347	0.354	-0.396	-0.089	-0.231	-0.747	0.617	-0.220
$C5$	-0.155	0.220	0.499	1.218	0.549	-0.201	0.747	-0.097	0.420	-0.463	-0.361

Pada Tabel 4.7 dapat diketahui bahwa terdapat perubahan *centroid* baru pada $C1, C2, C3, C4$, dan $C5$, sehingga diperlukan langkah perulangan perhitungan jarak data terhadap *centroid* dengan cara yang sama seperti sebelumnya yakni perhitungan jarak *Euclidean* dengan menggunakan nilai *centroid* baru pada Tabel 4.13. Berikut adalah contoh perhitungan jarak *Euclidean* dengan rumus (2.3) pada data ke-4, Kecamatan Dawarblandong, terhadap setiap *centroid* baru untuk setiap kelompok sebagai iterasi ke-2.

Jarak data ke-4 terhadap *centroid* kelompok ke-1 (C1)

$$\begin{aligned} d_{41} &= \sqrt{(3.639 - (-0.542))^2 + (1.287 - (-0.938))^2 + \dots + (2.606 - (-0.691))^2} \\ &= 8.014 \end{aligned}$$

Jarak data ke-4 terhadap *centroid* kelompok ke-2 (C2)

$$\begin{aligned} d_{42} &= \sqrt{(3.639 - 1.084)^2 + (1.287 - 0.853)^2 + \dots + (2.606 - 1.595)^2} \\ &= 3.635 \end{aligned}$$

Jarak data ke-4 terhadap *centroid* kelompok ke-3 (C3)

$$\begin{aligned} d_{43} &= \sqrt{(3.639 - (-0.194))^2 + (1.287 - 0.582)^2 + \dots + (2.606 - 0.496)^2} \\ &= 5.826 \end{aligned}$$

Jarak data ke-4 terhadap *centroid* kelompok ke-4 (C4)

$$\begin{aligned} d_{44} &= \sqrt{(3.639 - (-0.044))^2 + (1.287 - (-0.318))^2 + \dots + (2.606 - (-0.220))^2} \\ &= 5.600 \end{aligned}$$

Jarak data ke-4 terhadap *centroid* kelompok ke-5 (C5)

$$\begin{aligned} d_{25} &= \sqrt{(3.639 - (-0.155))^2 + (1.287 - 0.220)^2 + \dots + (2.606 - (-0.361))^2} \\ &= 5.985 \end{aligned}$$

Perhitungan jarak pada iterasi ke-2 bertujuan untuk mengevaluasi kedekatan setiap data terhadap *centroid* yang telah diperbarui dari hasil iterasi pertama. Hasil perhitungan *centroid* baru dari setiap iterasi dilampirkan pada Lampiran 3. Langkah iterasi penting dilakukan untuk mengidentifikasi apakah terdapat perubahan keanggotaan data dalam masing-masing klaster. Proses ini dilakukan dengan membandingkan jarak setiap data ke seluruh *centroid*, dan data akan diklasifikasikan ke dalam klaster dengan jarak terpendek. Melalui perbandingan ini, dapat diketahui apakah struktur klaster mengalami perubahan atau telah mencapai kestabilan. Hasil perhitungan jarak dalam iterasi ke-2 disajikan dalam Tabel 4.8,

yang digunakan sebagai dasar dalam pembentukan komposisi anggota kelompok terbaru pada proses *K-Means Clustering*.

Tabel 4.8 Hasil Perhitungan Jarak *Euclidean* Iterasi ke-2

No	C1	C2	C3	C4	C5	Jarak Minimum	Kelompok
1	0.000	6.127	4.988	3.933	4.775	0.000	1
2	0.000	6.127	4.988	3.933	4.775	0.000	1
3	0.000	6.127	4.988	3.933	4.775	0.000	1
4	8.014	3.635	5.286	5.600	5.985	3.635	2
5	6.226	2.635	2.417	4.574	4.354	2.417	3
6	4.932	1.986	1.487	2.907	3.769	1.487	3
7	5.588	4.502	3.983	2.995	2.002	2.002	5
8	5.896	4.561	4.799	4.483	2.787	2.787	5
9	5.719	1.711	2.279	3.364	3.511	1.711	2
10	3.994	4.310	3.179	1.775	2.885	1.775	4
11	4.241	4.660	3.946	1.770	2.707	1.770	4
12	6.146	6.247	5.913	4.651	3.326	3.326	5
13	5.463	3.327	1.486	4.026	3.996	1.486	3
14	5.214	4.411	4.399	2.317	3.269	2.317	4
15	4.366	4.904	4.511	2.123	2.863	2.123	4
16	6.720	5.245	5.626	5.528	4.100	4.100	5
17	5.312	2.970	3.387	3.121	4.105	2.970	2
18	3.639	4.844	3.742	1.554	3.394	1.554	4
19	4.421	3.974	3.500	1.948	2.366	1.948	4
20	3.361	4.148	3.463	1.397	3.011	1.397	4
21	4.316	3.863	3.132	2.009	2.245	2.009	4

Pada hasil iterasi ke-2 ini, terdapat kelompok yang mengalami perubahan anggota dari pengelompokan pada iterasi sebelumnya. Wilayah yang mengalami perubahan kelompok yaitu Kecamatan Gedeg pada iterasi ke-1 berada dalam kelompok ke-2, pada iterasi ke-2 berada dalam kelompok ke-3. Perubahan tersebut juga terjadi pada Kecamatan Puri, pada iterasi ke-1 berada dalam kelompok ke-4 dan pada iterasi ke-2 berada dalam kelompok ke-2. Oleh sebab itu, proses iterasi ini akan dilanjutkan ke iterasi-3.

Pada iterasi ke-3 masih terjadi perubahan anggota kelompok dalam setiap kelompok yaitu pada anggota kelompok ke-5, Kecamatan Mojoanyar mengalami perubahan kelompok, sehingga berada dalam kelompok ke-3. Karena masih terdapat perubahan, maka perlu dilakukan proses iterasi ke-4 untuk melihat adanya perubahan yang terjadi pada anggota dalam setiap kelompok yang terbentuk meskipun dengan nilai *centroid* yang berbeda. Hasil perubahan kelompok untuk iterasi ke-3 disajikan dalam Tabel 4.9. Sehingga proses iterasi perlu dilakukan lagi pada iterasi ke-4.

Tabel 4.9 Hasil Perhitungan Jarak *Euclidean* Iterasi ke-3

No	C1	C2	C3	C4	C5	Jarak Minimum	Kelompok
1	0.000	5.946	5.311	3.882	5.333	0.000	1
2	0.000	5.946	5.311	3.882	5.333	0.000	1
3	0.000	5.946	5.311	3.882	5.333	0.000	1
4	8.014	3.055	5.458	6.092	6.197	3.055	2
5	6.226	3.762	1.611	4.670	4.817	1.611	3
6	4.932	2.230	1.845	3.346	4.350	1.845	3
7	5.588	4.519	4.238	2.867	2.195	2.195	5
8	5.896	4.640	4.939	4.385	2.618	2.618	5
9	5.719	2.339	2.084	3.738	3.979	2.084	3
10	3.994	4.302	3.593	1.335	3.773	1.335	4
11	4.241	4.500	4.183	1.265	3.663	1.265	4
12	6.146	6.077	6.107	4.580	2.824	2.824	5
13	5.463	4.057	1.521	3.998	4.548	1.521	3
14	5.214	4.117	4.701	2.498	3.623	2.498	4
15	4.366	4.789	4.699	1.827	3.513	1.827	4
16	6.720	5.766	5.307	5.385	3.936	5.307	3
17	5.312	2.068	3.497	3.738	4.753	2.068	2
18	3.639	4.589	4.191	1.368	4.217	1.368	4
19	4.421	3.890	3.684	1.629	3.371	1.629	4
20	3.361	3.905	3.819	1.454	3.755	1.454	4
21	4.316	3.855	3.248	1.661	3.246	1.661	4

Pada iterasi ke-4 ini, tidak ada perubahan kelompok dan anggota dari setiap kelompok, artinya setiap kelompok pada iterasi ke-4 memiliki anggota yang sama

dari iterasi ke-3. Pada kelompok ke-1 memiliki tiga anggota, kelompok ke-2 memiliki dua anggota, kelompok ke-3 memiliki lima anggota, kelompok ke-4 memiliki delapan anggota, dan kelompok ke-5 memiliki tiga anggota. Hasil pengelompokan dari iterasi ke-4 disajikan dalam Tabel 4.10.

Tabel 4.10 Hasil Perhitungan Jarak *Euclidean* Iterasi ke-4

No	C1	C2	C3	C4	C5	Jarak Minimum	Kelompok
1	0.000	6.387	5.238	3.882	5.455	0.000	1
2	0.000	6.387	5.238	3.882	5.455	0.000	1
3	0.000	6.387	5.238	3.882	5.455	0.000	1
4	8.014	2.332	5.310	6.092	6.394	2.332	2
5	6.226	4.698	1.680	4.670	5.429	1.680	3
6	4.932	2.919	2.204	3.346	4.585	2.204	3
7	5.588	5.228	3.861	2.867	1.711	1.711	5
8	5.896	5.196	4.402	4.385	2.442	2.442	5
9	5.719	3.510	1.828	3.738	4.283	1.828	3
10	3.994	4.919	3.656	1.335	3.633	1.335	4
11	4.241	5.033	4.017	1.265	3.613	1.265	4
12	6.146	6.447	5.626	4.580	2.365	2.365	5
13	5.463	4.944	2.227	3.998	4.794	2.227	3
14	5.214	4.475	4.194	2.498	3.883	2.498	4
15	4.366	5.355	4.100	1.827	3.942	1.827	4
16	6.720	6.284	4.199	5.385	5.248	4.199	3
17	5.312	2.332	3.459	3.738	4.925	2.332	2
18	3.639	5.070	4.225	1.368	4.079	1.368	4
19	4.421	4.577	3.402	1.629	3.436	1.629	4
20	3.361	4.356	3.698	1.454	3.793	1.454	4
21	4.316	4.400	2.957	1.661	3.629	1.661	4

Pada setiap iterasi, data ke 1, 2, 3, dan 4 yaitu Kecamatan Kranggan, Magersari, dan Prajurit Kulon tidak mengalami perubahan kelompok sama sekali pada setiap iterasi yang dilakukan. Dalam iterasi ke-2 hingga iterasi ke-4 kelompok ke-4 juga tidak mengalami perubahan yang beranggotakan Kecamatan Ngoro, Kemlagi, Bangsal, Sooko, Gondang, Mojosari, Trawas, dan Kutorejo. Sedangkan pada kelompok ke-2, Kelompok ke-3, dan kelompok ke-5

mengalami perubahan anggota kelompok sebanyak satu anggota pada setiap iterasi, hingga mencapai iterasi ke-4, yang menunjukkan bahwa pada iterasi tersebut adalah pengelompokan yang baik, stabil, dan optimal. Berdasarkan tahapan dalam prosesnya pengelompokan dengan metode *K-Means Clustering* yang telah dilakukan, diperoleh hasil akhir pengelompokan wilayah di Kabupaten dan Kota Mojokerto berdasarkan data potensi sumber kesejahteraan sosial yang telah dilakukan standardisasi yang disajikan dalam Tabel 4.2.

Tabel 4.11 Hasil Metode *K-Means Clustering*

No	Kecamatan	x_1	x_2	...	x_{12}	Kelompok
1	Kranggan	-0.542	-0.938	...	-0.691	1
2	Magersari	-0.542	-0.938	...	-0.691	1
3	Prajurit Kulon	-0.542	-0.938	...	-0.691	1
4	Dawarblandong	3.639	1.287	...	2.606	2
5	Gedeg	-0.542	1.287	...	0.496	3
6	Jetis	0.155	0.473	...	1.287	3
7	Pungging	-0.542	-0.721	...	-0.163	5
8	Jatirejo	1.084	-0.721	...	0.232	5
9	Trowulan	0.155	-0.016	...	1.683	3
10	Ngoro	-0.542	-0.776	...	-0.691	4
11	Kemlagi	-0.542	-0.450	...	-0.691	4
12	Pacet	-0.542	-0.721	...	-0.691	5
13	Dlanggu	-0.542	0.690	...	-0.295	3
14	Bangsals	0.387	-0.287	...	-0.163	4
15	Sooko	-0.542	-0.124	...	-0.691	4
16	Mojoanyar	-0.542	2.753	...	-0.691	3
17	Puri	1.316	1.016	...	1.815	2
18	Gondang	-0.542	-0.830	...	-0.691	4
19	Mojosari	-0.077	-0.233	...	-0.163	4
20	Trawas	0.155	-0.776	...	-0.427	4
21	Kutorejo	-0.310	0.962	...	-0.691	4

Pada Tabel 4.11 menunjukkan hasil pengelompokan dengan lima kelompok dari data yang terdiri dari empat komponen utama sebagai pengganti dari variabel-variabel data yang merupakan indikator dari data potensi sumber kesejahteraan sosial per Kecamatan di Kabupaten dan Kota Mojokerto. Berdasarkan Tabel 4.11

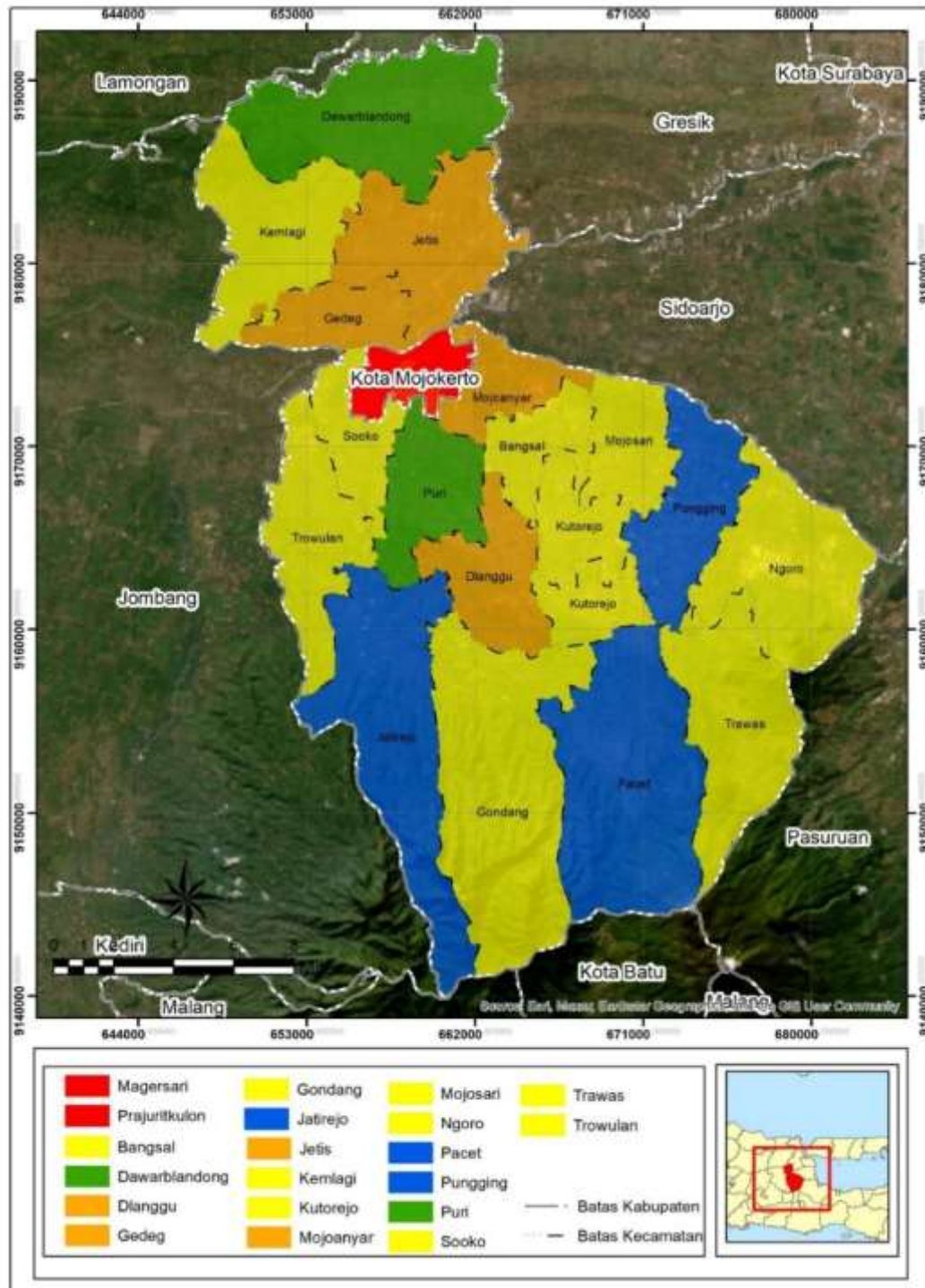
hasil pengelompokan terdapat pada kolom “Kelompok” yang terdiri dari lima kelompok sesuai dengan angka yang tertera dalam kolom tersebut, yang didasarkan pada kedekatan data dengan *centroid*. Untuk memudahkan identifikasi dan analisis lebih lanjut, masing-masing kelompok yang terbentuk kemudian diberikan penamaan khusus. Penamaan ini bertujuan untuk membedakan antar kelompok dan menggambarkan ciri umum dari masing-masing kelompok. Penamaan dan deskripsi masing-masing kelompok hasil pengelompokan tersebut disajikan secara rinci dalam Tabel 4.12.

Tabel 4.12 Hasil Pengelompokan Wilayah dengan *K-Means Clustering*

Keterangan	Penamaan Khusus	Kecamatan di Kabupaten-Kota Mojokerto	Banyak Anggota
Kelompok 1	Kawasan Perkotaan	Kecamatan Kranggan, Kecamatan Magersari, Kecamatan Prajurit Kulon.	3
Kelompok 2	Kawasan Perhutanan	Kecamatan Dawarblandong dan Kecamatan Puri	2
Kelompok 3	Kawasan Bersejarah	Kecamatan Jetis, Kecamatan Gedeg, Kecamatan Trowulan, Kecamatan Dlanggu, dan Kecamatan Mojoanyar.	5
Kelompok 4	Kawasan Industri	Kecamatan Ngoro, Kecamatan Kemlagi, Kecamatan Bangsal, Kecamatan Sooko, , Kecamatan Gondang, Kecamatan Mojosari, Kecamatan Trawas, Kecamatan Kutorejo.	8
Kelompok 5	Kawasan Pegunungan	Kecamatan Pungging, Kecamatan Jatirejo, dan Kecamatan Pacet.	3

Setelah proses pengelompokan dilakukan menggunakan metode *K-Means Clustering* dan hasil pengelompokan yang diperoleh dievaluasi menggunakan metode *V-Measure*, tahap selanjutnya adalah menyajikan hasil pengelompokan tersebut dalam bentuk visualisasi wilayah. Visualisasi ini bertujuan untuk memberikan gambaran spasial mengenai distribusi setiap kelompok yang terbentuk berdasarkan karakteristik data. Dengan demikian, pola persebaran antar kelompok

dapat dianalisis secara geografis dan lebih mudah diinterpretasikan. Hasil visualisasi pengelompokan wilayah tersebut ditampilkan pada Gambar 4.26.



Gambar 4.26 Visualisasi Hasil Pengelompokan *K-Means Clustering*

Berdasarkan Tabel 4.12 dan Gambar 4.26, dapat diketahui bahwa hasil pengelompokan wilayah berdasarkan data potensi sumber kesejahteraan sosial di Kabupaten-Kota Mojokerto terbagi menjadi lima kelompok. Pengelompokan ini diperoleh dari proses *K-Means Clustering* sehingga setiap kelompok merepresentasikan karakteristik data yang serupa berdasarkan indikator-indikatornya. Dari keseluruhan 21 Kecamatan, pada Kelompok 1 terdiri dari tiga Kecamatan dari Kota Mojokerto yang direpresentasikan dengan titik berwarna merah. Ketiga kecamatan ini memiliki kesamaan nilai jarak *Euclidean* data terhadap *centroid* dari setiap iterasi dengan memiliki nilai data yang mayoritas sama dari setiap variabel dari data potensi sumber kesejahteraan sosial. Selain itu, ketiga kecamatan ini diberi nama kawasan perkotaan sebagai label atau penamaan khusus, karena terletak di wilayah Kota Mojokerto dengan karakteristik perkotaan dan pusat administratif pemerintahan. Karakteristik khas perkotaan, seperti kepadatan penduduk yang tinggi, dominasi sektor perdagangan dan jasa, serta keberadaan fasilitas publik yang lengkap seperti pusat pemerintahan, pendidikan, dan kesehatan.

Wilayah berwarna hijau yang disajikan dalam Gambar 4.26 merepresentasikan Kelompok 2 terdiri dari dua kecamatan dari Kabupaten Mojokerto yakni Kecamatan Dawablandong dan Kecamatan Puri. Kecamatan ini memiliki nilai unggul secara signifikan pada variabel pekerja sosial profesional, variabel taruna siaga bencana, dan variabel wanita pemimpin kesejahteraan sosial. Keunggulan pada indikator-indikator ini menjadikan kedua kecamatan ini berbeda secara signifikan dibandingkan dengan kecamatan lain, sehingga membentuk kelompok tersendiri dalam hasil pengelompokan. Selain itu, kedua kecamatan ini

diberi nama kawasan perhutanan sebagai label atau penamaan khusus, karena sebagian besar wilayahnya merupakan kawasan hutan industri yang terletak di bagian barat wilayah kabupaten, berbatasan langsung dengan wilayah Kabupaten Lamongan dan Gresik.

Kelompok 3 yang direpresentasikan dengan wilayah yang berwarna coklat yang terdiri dari lima anggota dari wilayah Kabupaten Mojokerto yang meliputi Kecamatan Jetis, Kecamatan Gedeg, Kecamatan Trowulan, Kecamatan Dlanggu, dan Kecamatan Mojoanyar. Kelima kecamatan tersebut memiliki kesamaan nilai yang relatif tinggi dan hampir seimbang pada variabel lembaga kesejahteraan sosial dan variabel karang taruna. Kesamaan ini menunjukkan bahwa peran serta keberadaan lembaga-lembaga sosial serta organisasi kepemudaan seperti karang taruna cukup signifikan dan merata di wilayah ini. Kesamaan nilai tersebut menyebabkan jarak antar kecamatan dalam ruang data menjadi lebih dekat, sehingga membentuk satu kelompok tersendiri dalam hasil pengelompokan. Selain itu, kelima wilayah ini diberi penamaan khusus pada nama kelompoknya, yaitu kawasan bersejarah. Hal ini dikarenakan wilayah-wilayah tersebut memiliki nilai historis yang tinggi, terutama sebagai pusat peradaban Kerajaan Majapahit. Salah satu contohnya adalah Kecamatan Trowulan yang dikenal sebagai situs arkeologi peninggalan Majapahit, serta kecamatan lainnya yang berada di sekitarnya dan turut mendukung pelestarian cagar budaya, museum, dan potensi wisata sejarah yang ada di Mojokerto.

Kelompok 4 yang direpresentasikan dengan wilayah yang berwarna kuning terdiri dari delapan Kecamatan yaitu Kecamatan Ngoro, Kecamatan Kemlagi, Kecamatan Bangsal, Kecamatan Sooko, Kecamatan Gondang, Kecamatan

Mojosari, Kecamatan Trawas, Kecamatan Kutorejo. Sembilan kecamatan tersebut dapat menjadi satu kelompok karena memiliki persamaan karakteristik yang menonjol pada nilai variabel pekerja sosial masyarakat yang nilainya lebih rendah dibandingkan dengan anggota kelompok lainnya. Nilai yang relatif kecil pada variabel ini menunjukkan bahwa aktivitas atau keberadaan pekerja sosial masyarakat di kecamatan-kecamatan ini tidak begitu dominan, sehingga menjadikannya berbeda dari kelompok lain dan secara alami mengelompok bersama dalam analisis klaster. Selain itu kesembilan wilayah ini diberi label khusus sebagai nama kelompok yaitu, kawasan industri. Hal ini dikarenakan beberapa kecamatan di dalamnya, seperti Kecamatan Ngoro dan Mojosari, merupakan pusat aktivitas industri dan perdagangan yang berkembang pesat. Keberadaan kawasan industri besar, sentra UMKM, dan dukungan infrastruktur penunjang ekonomi menjadikan wilayah ini memiliki karakter ekonomi yang lebih kuat dibandingkan kelompok lainnya. Label ini juga mencerminkan orientasi pembangunan wilayah yang lebih berfokus pada sektor industri dan pertumbuhan ekonomi.

Wilayah yang berwarna biru merepresentasikan Kelompok 5 yang terdiri dari tiga Kecamatan dari Kabupaten Mojokerto, yaitu Kecamatan Pungging, Kecamatan Jatirejo, dan Kecamatan Pacet. Dalam visualisasi pada Gambar 4.14, ketiga kecamatan ini secara geografis terletak saling berdekatan. Berdasarkan karakteristik data, ketiganya memiliki kesamaan nilai yang tinggi pada variabel wahana kesejahteraan sosial berbasis masyarakat, yang menjadi indikator utama dalam pembentukan kelompok ini. Kesamaan geografis dan indikator dominan yang serupa mengindikasikan adanya pola keterkaitan antara kondisi wilayah dengan

tingkat ketersediaan dan keberfungsian wahana kesejahteraan sosial di masyarakat. Hal ini menunjukkan bahwa wilayah-wilayah dalam Kelompok 5 memiliki potensi yang kuat dalam pengembangan partisipasi sosial berbasis komunitas. Selain itu, ketiga kecamatan ini diberi nama kawasan pegunungan sebagai label atau penamaan khusus, karena penamaan khusus karena secara geografis terletak di wilayah dataran tinggi atau pegunungan. Mayoritas wilayahnya dimanfaatkan sebagai kawasan wisata alam, seperti pemandian air panas, air terjun, serta area perkemahan dan agrowisata, yang menjadi daya tarik utama bagi wisatawan lokal maupun luar daerah.

4.7 Evaluasi Hasil Pengelompokan dengan Metode *V-Measure*

Hasil pengelompokan wilayah Kabupaten dan Kota Mojokerto yang telah diperoleh pada subbab 4.6 dengan metode *K-Means Clustering* disajikan dalam Tabel 4.12, kemudian dilakukan evaluasi menggunakan metode *V-Measure*. *V-Measure* mengombinasikan dua aspek penting, yaitu homogenitas dan kelengkapan. Homogenitas mengukur sejauh mana data dalam satu kelompok berasal dari kelompok yang sama, sementara kelengkapan mengukur sejauh mana data dari kelompok yang sama berada dalam kelompok yang sama. Evaluasi ini dilakukan untuk menilai ketepatan pembagian kelompok yang dihasilkan dengan memberikan nilai kuantitatif dan objektif dalam menilai kualitas hasil pengelompokan.

Penerapan evaluasi *V-Measure* dilakukan dengan cara membandingkan dua hasil pengelompokan dalam mengukur kualitas kelompok yang terbentuk. Dalam penelitian ini, metode evaluasi *V-Measure* digunakan untuk membandingkan hasil

pengelompokan data potensi sumber kesejahteraan sosial menggunakan metode *K-Means Clustering* dan *Agglomerative Hierarchical Clustering* dengan metode pendekatan *Centroid Linkage*. Pemilihan metode *Centroid Linkage* sebagai pembanding didasarkan pada kemiripan konsep dengan *K-Means Clustering*, yang keduanya menggunakan prinsip titik pusat atau *centroid* sebagai dasar pengelompokan, serta menghasilkan pola pengelompokan yang paling serupa baik dari segi distribusi maupun struktur kelompok yang terbentuk. Oleh karena itu, metode ini dianggap paling sesuai untuk dievaluasi lebih lanjut melalui pendekatan *V-Measure*. Hasil pengelompokan wilayah berdasarkan data potensi sumber kesejahteraan dengan metode *Centroid Linkage* diberi penamaan “*Cluster*” untuk setiap kelompok, yang disajikan dalam Tabel 4.13.

Tabel 4.13 Tabel Hasil Pengelompokan Data Metode *Centroid Linkage*

Keterangan	Penamaan Khusus	Kecamatan di Kabupaten-Kota Mojokerto	Banyak Anggota
<i>Cluster 1</i>	Kawasan Perkotaan	Kecamatan Kraggan, Kecamatan Prajurit Kulon dan Kecamatan Magersari.	3
<i>Cluster 2</i>	Kawasan Bersejarah	Kecamatan Dawarblandong, Kecamatan Jetis, Kecamatan Gedeg, Kecamatan Trowulan, Kecamatan Dlanggu, dan Kecamatan Puri	6
<i>Cluster 3</i>	Kawasan Industri	Kecamatan Bangsal, Kecamatan Sooko, , Kecamatan Ngoro, Kecamatan Gondang, Kecamatan Mojosari, Kecamatan Pacet, Kecamatan Trawas, Kecamatan Kutorejo, Kecamatan Pungging, dan Kecamatan Kemlagi.	10
<i>Cluster 4</i>	Kawasan Agraris	Kecamatan Jatirejo	1
<i>Cluster 5</i>	Kawasan Pengembangan	Kecamatan Mojoanyar	1

Berdasarkan Tabel 4.12 dan Tabel 4.13 disusun Tabel 4.14 sebagai hasil integrasi dari kedua tabel tersebut.

Tabel 4.14 Tabel Hasil Pengelompokan Data

No	Kecamatan	Kelompok (<i>C</i>)	Cluster (<i>Y</i>)
1	Kranggan	1	1
2	Magersari	1	1
3	Prajurit Kulon	1	1
4	Dawarblandong	2	2
5	Gedeg	3	2
6	Jetis	3	2
7	Pungging	5	3
8	Jatirejo	5	4
9	Trowulan	3	2
10	Ngoro	4	3
11	Kemlagi	4	3
12	Pacet	5	3
13	Dlanggu	3	2
14	Bangsals	4	3
15	Sooko	4	3
16	Mojoanyar	3	5
17	Puri	2	2
18	Gondang	4	3
19	Mojosari	4	3
20	Trawas	4	3
21	Kutorejo	4	3

Berdasarkan Tabel 4.14, kolom Kelompok merupakan hasil pengelompokan wilayah berdasarkan data PSKS di Kabupaten dan Kota Mojokerto dengan metode *K-Means Clustering*. Setiap kelompok memiliki anggota sebagai berikut: sebanyak tiga kecamatan dalam kelompok ke-1 ($C_1 = \{1, 2, 3\}$), dua kecamatan dalam kelompok ke-2 ($C_2 = \{4, 5\}$), dalam kelompok ke-3 terdapat lima kecamatan ($C_3 = \{6, 7, 8, 9, 10\}$), dalam kelompok ke-4 terdapat delapan kecamatan ($C_4 = \{11, 12, 13, 14, 15, 16, 17, 18\}$), dan dalam kelompok ke-5 terdapat tiga kecamatan ($C_5 = \{19, 20, 21\}$). Pada kolom *Cluster* merupakan hasil pengelompokan wilayah di Kabupaten dan Kota Mojokerto dengan metode *Centroid Linkage*. Setiap *cluster* memiliki anggota sebagai berikut: sebanyak tiga kecamatan dalam *cluster* ke-1

($Y_1 = \{1, 2, 3\}$), enam kecamatan dalam *cluster* ke-2 ($Y_1 = \{1, 2, \dots, 6\}$), dalam *cluster* ke-3 terdapat sepuluh kecamatan ($Y_3 = \{1, 2, 3, \dots, 10\}$), dalam *cluster* ke-4 terdapat satu kecamatan ($Y_4 = \{1\}$), dan dalam *cluster* ke-5 terdapat satu kecamatan ($Y_5 = \{1\}$). Berikut merupakan perhitungan dengan menggunakan metode *V-Measure* dari hasil pengelompokan data potensi sumber kesejahteraan sosial dengan metode *K-Means Clustering* dan metode *Centroid Linkage*.

- 1) Menentukan distribusi data setiap kelompok dari Tabel 4.14 dalam kolom Kelompok dengan $n = 21$.

Tabel 4.15 Distribusi Data Setiap Kelompok

Kelompok	Banyak Anggota	Probabilitas $P(C)$
1	3	$\frac{3}{21}$
2	2	$\frac{2}{21}$
3	5	$\frac{5}{21}$
4	8	$\frac{8}{21}$
5	3	$\frac{3}{21}$

- 2) Menghitung nilai entropi bersama untuk setiap kelompok ($H(C)$)

Berdasarkan Tabel 4.15, diperoleh probabilitas setiap kelompok yang akan digunakan untuk menghitung nilai entropi bersama. Seperti pada kelompok ke-1, terdapat tiga anggota, sehingga probabilitas yang diperoleh yaitu $\frac{3}{21}$, yang artinya dalam kelompok ke-1 terdiri dari tiga kecamatan dari 21 kecamatan lain yang berada di dalam wilayah Kabupaten dan Kota Mojokerto. Berikut perhitungan manual yang dilakukan menggunakan rumus (2.8).

Kelompok ke-1(C_1) memiliki nilai probabilitas $P(1) = \frac{3}{21}$

$$\begin{aligned}
 H(C_1) &= \frac{3}{21} \log_2 \frac{3}{21} = \frac{1}{7} \times \left(\frac{\log_2 \frac{1}{7}}{\log 2} \right) = \frac{1}{7} \times \left(\frac{-0.844}{0.301} \right) \\
 &= 0.143 \times (-2.807) = -0.401
 \end{aligned}$$

Kelompok ke-2 (C_2) memiliki nilai probabilitas $P(2) = \frac{2}{21}$

$$\begin{aligned}
 H(C_2) &= \frac{2}{21} \log_2 \frac{2}{21} = \frac{2}{21} \times \left(\frac{\log_2 \frac{2}{21}}{\log 2} \right) = \frac{2}{21} \times \left(\frac{-1.021}{0.301} \right) \\
 &= 0.095 \times (-3.392) = -0.323
 \end{aligned}$$

Kelompok ke-3 (C_3) memiliki nilai probabilitas $P(3) = \frac{5}{21}$

$$\begin{aligned}
 H(C_3) &= \frac{5}{21} \log_2 \frac{5}{21} = \frac{5}{21} \times \left(\frac{\log_2 \frac{5}{21}}{\log 2} \right) = \frac{5}{21} \times \left(\frac{-0.623}{0.301} \right) \\
 &= 0.238 \times (-2.070) = -0.493
 \end{aligned}$$

Kelompok ke-4 (C_4) memiliki nilai probabilitas $P(4) = \frac{8}{21}$

$$\begin{aligned}
 H(C_4) &= \frac{8}{21} \log_2 \frac{8}{21} = \frac{8}{21} \times \left(\frac{\log_2 \frac{8}{21}}{\log 2} \right) = \frac{8}{21} \times \left(\frac{-0.419}{0.301} \right) \\
 &= 0.381 \times (-1.392) = -0.531
 \end{aligned}$$

Kelompok ke-5 (C_5) memiliki nilai probabilitas $P(5) = \frac{3}{21}$

$$\begin{aligned}
 H(C_5) &= \frac{3}{21} \log_2 \frac{3}{21} = \frac{1}{7} \times \left(\frac{\log_2 \frac{1}{7}}{\log 2} \right) = \frac{1}{7} \times \left(\frac{-0.844}{0.301} \right) \\
 &= 0.143 \times (-2.807) = -0.401
 \end{aligned}$$

Total nilai entropi bersama

$$\begin{aligned}
 H(C) &= - \sum_{c=1}^{|C|} \frac{\sum_{h=1}^{|Y|} a_{ch}}{n} \log_2 \frac{\sum_{h=1}^{|Y|} a_{ch}}{n} \\
 &= -((-0.401) + (-0.323) + (-0.493) + (-0.531) + (-0.401)) \\
 &= 2.149
 \end{aligned}$$

- 3) Mendefinisikan data untuk menghitung nilai entropi bersyarat dengan menggunakan hasil pengelompokan dari Tabel 4.14 pada kolom Kelompok yang diberikan hasil pengelompokan pada kolom *Cluster* untuk mengevaluasi

sejauh mana hasil pengelompokan mendekati distribusi kelas yang sebenarnya.

Tabel 4.16 Tabel Hubungan Antara Anggota Kelompok dan *Cluster*

Kelompok (C)	Cluster (Y)					Total	Probabilitas ($P(C Y)$)
	1	2	3	4	5		
1	3	0	0	0	0	3	$\frac{3}{3}$
2	0	2	0	0	0	2	$\frac{2}{2}$
3	0	4	0	0	1	5	$\frac{4}{5}, \frac{1}{5}$
4	0	0	8	0	0	8	$\frac{8}{8}$
5	0	0	2	1	0	3	$\frac{2}{3}, \frac{1}{3}$

Dalam menghitung nilai entropi bersyarat disajikan Tabel 4.16 sebagai tabel yang menunjukkan hubungan antara Kelompok dengan *Cluster*. Dalam setiap baris, menunjukkan banyak anggota suatu kelompok yang berasal dari *cluster* tertentu. Berikut interpretasi dari Tabel 4.16 secara lebih rinci.

Kelompok 1 (C_1) : Tiga anggota Kelompok dari *Cluster* 1.

Kelompok 2 (C_2) : Dua anggota Kelompok dari *Cluster* 2.

Kelompok 3 (C_3) : Empat anggota Kelompok dari *Cluster* 2 dan satu anggota Kelompok dari *Cluster* 5.

Kelompok 4 (C_4) : Delapan anggota Kelompok dari *Cluster* 3.

Kelompok 5 (C_5) : Dua anggota Kelompok dari *Cluster* 3 dan satu anggota Kelompok dari *Cluster* 4.

4) Menghitung nilai entropi bersyarat $H(C|Y)$ menggunakan rumus (2.9).

Kelompok 1 (C_1) memiliki nilai probabilitas $P(C_1|Y_1) = \frac{3}{3}$

$$H(C|Y_1) = \frac{3}{3} \log_2 \frac{3}{3} = 0$$

Kelompok 2 (C_2) memiliki nilai probabilitas $P(C_2|Y_2) = \frac{2}{2}$

$$H(C|Y_2) = \frac{2}{2} \log_2 \frac{2}{2} = 0$$

Kelompok 3 (C_3) memiliki nilai probabilitas $P(C_3|Y_2) = \frac{4}{5}$ dan nilai

probabilitas pada $P(C_3|Y_5) = \frac{1}{5}$

$$\begin{aligned} H(C|Y_3) &= \frac{4}{5} \log_2 \frac{4}{5} + \frac{1}{5} \log_2 \frac{1}{5} = \frac{4}{5} \left(\frac{\log_2 \frac{4}{5}}{\log 2} \right) + \frac{1}{5} \left(\frac{\log_2 \frac{1}{5}}{\log 2} \right) \\ &= \left(\frac{4}{5} \times \frac{-0.097}{0.301} \right) + \left(\frac{1}{5} \times \frac{-0.699}{0.301} \right) \\ &= (0.8 \times (-0.322)) + (0.2 \times (-2.322)) \\ &= -0.258 - 0.465 = -0.723 \end{aligned}$$

Kelompok 4 (C_4) memiliki nilai probabilitas $P(C_4|Y_3) = \frac{8}{8}$

$$H(C|Y_4) = \frac{8}{8} \log_2 \frac{8}{8} = 0$$

Kelompok 5 (C_5) memiliki nilai probabilitas $P(C_5|Y_3) = \frac{2}{3}$ dan nilai

probabilitas $P(C_5|Y_4) = \frac{1}{3}$

$$\begin{aligned} H(C|Y_5) &= \frac{2}{3} \log_2 \frac{2}{3} + \frac{1}{3} \log_2 \frac{1}{3} = \frac{2}{3} \left(\frac{\log_2 \frac{2}{3}}{\log 2} \right) + \frac{1}{3} \left(\frac{\log_2 \frac{1}{3}}{\log 2} \right) \\ &= \left(\frac{2}{3} \times \frac{-0.176}{0.301} \right) + \left(\frac{1}{3} \times \frac{-0.477}{0.301} \right) \\ &= (0.666 \times (-0.585)) + (0.333 \times (-1.585)) \\ &= -0.528 - 0.390 = -0.919 \end{aligned}$$

Total nilai entropi bersyarat

$$\begin{aligned} H(C|Y) &= - \left(\frac{(3 \times 0) + (2 \times 0) + (5 \times (-0.723)) + (8 \times 0) + (3 \times (-0.919))}{21} \right) \\ &= - \left(\frac{-6.372}{21} \right) = 0.303 \end{aligned}$$

5) Menghitung nilai homogenitas menggunakan rumus pada (2.10).

$$H = 1 - \frac{H(C|Y)}{H(C)} = 1 - \frac{0.303}{2.149} = 1 - 0.141 = 0.859$$

- 6) Menentukan distribusi data setiap kelompok dari Tabel 4.14 dalam kolom *Cluster* dengan $n = 21$.

Tabel 4.17 Distribusi Data dari Setiap *Cluster*

<i>Cluster</i>	Banyak Anggota	Probabilitas $P(Y)$
1	3	$\frac{3}{21}$
2	6	$\frac{6}{21}$
3	10	$\frac{10}{21}$
4	1	$\frac{1}{21}$
5	1	$\frac{1}{21}$

- 7) Menghitung nilai entropi bersama untuk setiap *cluster* ($H(Y)$)

Berdasarkan Tabel 4.17, diperoleh probabilitas yang akan digunakan untuk menghitung nilai entropi bersama untuk setiap *cluster*. Seperti pada *Cluster* 1, terdapat tiga anggota, sehingga probabilitas yang diperoleh yaitu $\frac{3}{21}$, yang artinya dalam label 1 terdiri dari tiga kecamatan dari 21 kecamatan lain yang berada di dalam wilayah Kabupaten dan Kota Mojokerto. Berikut perhitungan manual yang dilakukan menggunakan rumus (2.11).

Cluster 1 (Y_1) memiliki nilai probabilitas $P(1) = \frac{3}{21}$

$$\begin{aligned}
 H(Y_1) &= \frac{3}{21} \log_2 \frac{3}{21} = \frac{3}{21} \times \left(\frac{\log_{21} 3}{\log 2} \right) = \frac{3}{21} \times \left(\frac{-0.844}{0.301} \right) \\
 &= 0.143 \times (-2.807) = -0.401
 \end{aligned}$$

Cluster 2 (Y_2) memiliki nilai probabilitas $P(2) = \frac{6}{21}$

$$\begin{aligned}
 H(Y_2) &= \frac{6}{21} \log_2 \frac{6}{21} = \frac{6}{21} \times \left(\frac{\log_{21} 6}{\log 2} \right) = \frac{6}{21} \times \left(\frac{-0.544}{0.301} \right) \\
 &= 0.286 \times (-1.807) = -0.517
 \end{aligned}$$

Cluster 3 (Y_3) memiliki nilai probabilitas $P(3) = \frac{10}{21}$

$$\begin{aligned}
 H(Y_3) &= \frac{10}{21} \log_2 \frac{10}{21} = \frac{10}{21} \times \left(\frac{\log \frac{10}{21}}{\log 2} \right) = \frac{10}{21} \times \left(\frac{-0.322}{0.301} \right) \\
 &= 0.476 \times (-1.070) = -0.509
 \end{aligned}$$

Cluster 4 (Y_4) memiliki nilai probabilitas $P(4) = \frac{1}{21}$

$$\begin{aligned}
 H(Y_5) &= \frac{1}{21} \log_2 \frac{1}{21} = \frac{1}{21} \times \left(\frac{\log \frac{1}{21}}{\log 2} \right) = \frac{1}{21} \times \left(\frac{-1.322}{0.301} \right) \\
 &= 0.048 \times (-4.393) = -0.210
 \end{aligned}$$

Cluster 5 (Y_5) memiliki nilai probabilitas $P(5) = \frac{1}{21}$

$$\begin{aligned}
 H(Y_5) &= \frac{1}{21} \log_2 \frac{1}{21} = \frac{1}{21} \times \left(\frac{\log \frac{1}{21}}{\log 2} \right) = \frac{1}{21} \times \left(\frac{-1.322}{0.301} \right) \\
 &= 0.048 \times (-4.393) = -0.210
 \end{aligned}$$

Total nilai entropi bersama untuk setiap *Cluster*

$$\begin{aligned}
 H(Y) &= -((-0.401) + (-0.517) + (-0.509) + (-0.210) + (-0.210)) \\
 &= 1.847
 \end{aligned}$$

- 8) Mendefinisikan data untuk menghitung nilai entropi bersyarat dengan menggunakan hasil pengelompokan dari Tabel 4.14 pada kolom *Cluster* dengan diberikan hasil pengelompokan pada kolom Kelompok.

Tabel 4.18 Tabel Hubungan Anggota *Cluster* dengan Kelompok

<i>Cluster</i> (Y)	Kelompok (C)					Total	Probabilitas ($P(Y C)$)
	1	2	3	4	5		
1	3	0	0	0	0	3	$\frac{3}{3}$
2	0	2	4	0	0	6	$\frac{2}{6}, \frac{4}{6}$
3	0	0	0	8	2	10	$\frac{8}{10}, \frac{2}{10}$
4	0	0	0	0	1	1	$\frac{1}{1}$
5	0	0	1	0	0	1	$\frac{1}{1}$

Dalam menghitung nilai entropi bersyarat disajikan Tabel 4.18 sebagai tabel yang menunjukkan hubungan antara *Cluster* dengan Kelompok. Berdasarkan

Tabel 4.18, setiap baris menunjukkan banyak anggota suatu kelompok yang berasal dari label tertentu. Berikut interpretasi dari Tabel 4.18 secara lebih rinci.

Cluster 1 (Y_1) : Tiga anggota *Cluster* dari Kelompok 1.

Cluster 2 (Y_2) : Dua anggota *Cluster* dari Kelompok 2 dan empat anggota *Cluster* dari Kelompok 3.

Cluster 3 (Y_3) : Delapan anggota *Cluster* dari Kelompok 4 dan dua anggota *Cluster* dari Kelompok 5.

Cluster 4 (Y_4) : Satu anggota *Cluster* dari Kelompok 5.

Cluster 5 (Y_5) : Satu anggota *Cluster* dari Kelompok 3.

- 9) Menghitung nilai entropi bersyarat $H(Y|C)$ dengan menggunakan rumus pada (2.12) berdasarkan hasil nilai probabilitas pada Tabel 4.18.

Cluster 1 (Y_1) memiliki nilai probabilitas $P(Y_1|C_1) = \frac{3}{3}$

$$H(Y|C_1) = \frac{3}{3} \log_2 \frac{3}{3} = 0$$

Cluster 2 (Y_2) memiliki nilai probabilitas $P(Y_2|C_2) = \frac{2}{6}$ dan nilai probabilitas

$$P(Y_2|C_3) = \frac{4}{6}$$

$$\begin{aligned} H(Y|C_2) &= \frac{2}{6} \log_2 \frac{2}{6} + \frac{4}{6} \log_2 \frac{4}{6} = \left(\frac{2}{6} \times \left(\frac{\log_2 \frac{2}{6}}{\log_2 2} \right) \right) + \left(\frac{4}{6} \times \left(\frac{\log_2 \frac{4}{6}}{\log_2 2} \right) \right) \\ &= \left(0.333 \times \left(\frac{-0.477}{0.301} \right) \right) + \left(0.666 \times \left(\frac{-0.176}{0.301} \right) \right) \\ &= (0.333 \times (-1.585)) + (0.666 \times (-0.585)) \\ &= -0.918 \end{aligned}$$

Cluster 3 (Y_3) memiliki nilai probabilitas $P(Y_3|C_4) = \frac{8}{10}$, dan nilai

$$\text{probabilitas } P(Y_3|C_5) = \frac{2}{10}$$

$$\begin{aligned}
H(Y|C_3) &= \frac{8}{10} \log_2 \frac{8}{10} + \frac{2}{10} \log_2 \frac{2}{10} = \left(\frac{8}{10} \times \left(\frac{\log_2 \frac{8}{10}}{\log_2 2} \right) \right) + \left(\frac{2}{10} \times \left(\frac{\log_2 \frac{2}{10}}{\log_2 2} \right) \right) \\
&= \left(0.8 \times \left(\frac{-0.097}{0.301} \right) \right) + \left(0.2 \times \left(\frac{-0.699}{0.301} \right) \right) \\
&= (0.8 \times (-0.322)) + (0.2 \times (-2.322)) \\
&= -0.721
\end{aligned}$$

Cluster 4 (Y_4) memiliki nilai probabilitas $P(Y_4|C_5) = \frac{1}{1}$

$$H(Y|C_4) = \frac{1}{1} \log_2 \frac{1}{1} = 0$$

Cluster 5 (Y_5) memiliki nilai probabilitas $P(Y_5|C_3) = \frac{1}{1}$

$$H(Y|C_5) = \frac{1}{1} \log_2 \frac{1}{1} = 0$$

Total nilai entropi bersyarat untuk setiap label yang diperoleh adalah

$$\begin{aligned}
H(Y|C) &= - \left(\frac{(3 \times 0) + (6 \times (-0.918)) + (10 \times (-0.721)) + (1 \times 0) + (1 \times 0)}{21} \right) \\
&= - \left(\frac{-12.718}{21} \right) = 0.606
\end{aligned}$$

10) Menghitung nilai kelengkapan K menggunakan rumus pada (2.13).

$$\begin{aligned}
K &= 1 - \frac{H(Y|C)}{H(Y)} \\
&= 1 - \frac{0.606}{1.847} = 1 - 0.328 = 0.672
\end{aligned}$$

11) Menghitung nilai V -Measure dengan menggunakan rumus (2.14).

$$v = 2 \times \frac{(H.K)}{H+K} = 2 \times \frac{(0.859 \times 0.672)}{0.859 + 0.672} = 2 \times \frac{0.577}{1.531} = 2 \times 0.378 = 0.756$$

Berdasarkan hasil perhitungan homogenitas dan kelengkapan secara berurutan menggunakan rumus pada persamaan (2.10) dan (2.13), untuk metode evaluasi V -Measure, diperoleh nilai V -Measure sebesar 0.756. Nilai kuantitatif yang mendekati nilai 1, menunjukkan bahwa kualitas pengelompokan yang dihasilkan oleh Metode K -Means Clustering berdasarkan data potensi sumber

kesejahteraan sosial bersifat akurat sebagaimana ditunjukkan dalam Tabel 2.2. Hal ini mengindikasikan bahwa hasil pengelompokan yang diperoleh memiliki tingkat homogenitas dan kelengkapan yang cukup baik dan mencerminkan keseimbangan kedua konsep tersebut secara signifikan.

Nilai homogenitas yang diperoleh sebesar 0.859 berdasarkan Tabel 2.1 menunjukkan bahwa hasil pengelompokan memiliki kualitas yang sangat baik, artinya setiap anggota dalam kelompok cenderung memiliki kesamaan dalam hal label yang diberikan. Sedangkan, nilai kelengkapan yang diperoleh sebesar 0.672 berdasarkan Tabel 2.1 menunjukkan bahwa hasil pengelompokan dapat dianggap sudah baik dalam mencakup seluruh label yang ada. Dengan demikian, metode *K-Means Clustering* dan metode *Agglomerative Hierarchical Clustering, Centroid Linkage*, dapat dikatakan akurat dalam menjaga integritas hasil pengelompokan data potensi sumber kesejahteraan sosial di wilayah Kabupaten dan Kota Mojokerto, karena mampu membentuk kelompok yang sesuai dengan karakteristik data secara optimal.

4.8 Kajian Penelitian dalam Perspektif Islam

Kesejahteraan dalam Islam dimaknai dengan istilah *falah* yaitu kesejahteraan yang bersifat holistik dan seimbang antara aspek material dan spritual, individual-sosial dan kesejahteraan di kehidupan duniawi dan akhirat. Kesejahteraan di dunia dapat diartikan sebagai segala sesuatu yang dapat membuat kenikmatan hidup indrawi, baik jasmani, intelektual, biologis maupun material. Adapun kesejahteraan akhirat merupakan kenikmatan yang akan diperoleh setelah manusia meninggal dunia (Ilyas, 2016). Kesejahteraan masyarakat dimaknai oleh Imam al-Ghazali

dengan tercapainya kemaslahatan yang dapat dirasakan setelah terwujudnya kesejahteraan yang sesungguhnya dari seluruh manusia yaitu tercukupinya kebutuhan materi maupun rohani (Almahmudi, 2019). Imam al-Ghazali menjelaskan tentang sumber-sumber kesejahteraan, yaitu: terpeliharanya agama, jiwa, akal, keturunan dan harta (Rohman, 2010).

Oleh karena itu Allah swt., memerintahkan untuk melakukan kebajikan dan mensejahterakan umat manusia yang tersurat dalam QS. al Baqarah (2) ayat 177 yang menjelaskan bahwa kebajikan yang sesungguhnya adalah beriman kepada Allah dengan tidak menyekutukan-Nya; beriman kepada hari akhir yaitu hari pembalasan segala amal perbuatan selama di dunia, sehingga mendorong manusia untuk selalu berbuat baik; beriman kepada malaikat-malaikat yang taat menjalankan perintah Allah dan tidak pernah berbuat maksiat sehingga mendorong manusia untuk meneladani ketaatannya; beriman kepada kitab-kitab yang diturunkan oleh Allah kepada para rasul; beriman kepada nabi-nabi yang selalu menyampaikan kebenaran meskipun banyak yang memusuhinya (Purnamasari, 2023).

Kebajikan orang yang memberikan harta yang dicintainya kepada kerabat yang kurang mampu, anak yatim, karena mereka sudah kehilangan orang tua, sehingga setiap orang beriman patut memberikan kebaikan kepada mereka, orang-orang miskin yang hidupnya serba kekuarangan dalam memenuhi kebutuhan sehari-hari, orang-orang yang dalam perjalanan atau musafir yang kehabisan bekal perjalanan, peminta-minta untuk meringankan penderitaan dan kekurangannya, dan untuk memerdekakan hamba sahaya yang timbul akibat praktik perbudakan (Purnamasari, 2023).

Selain itu, kebajikan orang yang melaksanakan salat dengan khusyuk dan memenuhi syarat dan rukunnya, menunaikan zakat sesuai ketentuan dan tidak menunda-nunda pelaksanaannya, orang-orang yang menepati janji, orang yang sabar saat mengalami musibah, kesengsaraan, kepedihan dan berbagai macam kekurangan. Orang yang mempunyai sifat-sifat ini, mereka itulah orang-orang yang benar keimanannya, dan mereka itulah orang-orang yang bertakwa kepada Allah (Purnamasari, 2023).

BAB V

PENUTUP

5.1 Kesimpulan

Sesuai dengan penelitian yang telah dilakukan oleh peneliti pada data potensi sumber kesejahteraan sosial Kabupaten dan Kota Mojokerto dengan menggunakan metode *K-Means Clustering* dan dilanjutkan dengan evaluasi hasil pengelompokan dengan metode *V-Measure*, diperoleh kesimpulan sebagai berikut:

1. Hasil pengelompokan wilayah di Mojokerto menggunakan metode *K-Means Clustering* didasarkan pada banyak kelompok optimal yang ditentukan melalui metode *Elbow*, yakni pada nilai $k = 5$. Dengan demikian, wilayah Mojokerto terbagi menjadi lima kelompok sebagai berikut:
 - a. Kelompok 1: Kecamatan Kranggan, Kecamatan Magersari, dan Kecamatan Prajuritkulon.
 - b. Kelompok 2: Kecamatan Dawarblandong dan Kecamatan Puri dari Kabupaten Mojokerto.
 - c. Kelompok 3: Kecamatan Jetis, Kecamatan Gedeg, Kecamatan Trowulan, Kecamatan Dlanggu, dan Kecamatan Mojoanyar.
 - d. Kelompok 4: Kecamatan Ngoro, Kecamatan Kemlagi, Kecamatan Bangsal, Kecamatan Sooko, Kecamatan Gondang, Kecamatan Mojosari, Kecamatan Trawas, Kecamatan Kutorejo.
 - e. Kelompok 5: Kecamatan Pungging, Kecamatan Jatirejo, dan Kecamatan Pacet.

2. Hasil evaluasi menggunakan metode *V-Measure* pada subbab 4.7, diperoleh nilai sebesar 0.756, yang menunjukkan bahwa hasil pengelompokan data menggunakan metode *K-Means Clustering* yang berbasis titik pusat mampu menghasilkan kelompok dengan kategori akurat yang memiliki keseimbangan antara homogenitas dan kelengkapan. Dengan nilai homogenitas yang diperoleh sebesar 0.859 yang tergolong dalam kategori sangat baik menunjukkan bahwa sebagian besar anggota dalam masing-masing kelompok memiliki karakteristik yang serupa. Sedangkan, nilai kelengkapan yang diperoleh sebesar 0.672 menunjukkan bahwa hasil pengelompokan termasuk dalam kategori sudah baik dalam mencakup seluruh kelompok yang ada meskipun masih terdapat beberapa pencampuran.

5.2 Saran

Pada penelitian ini peneliti hanya mengkaji teknik pengelompokan *non-hierarki K-Means Clustering* dengan pemilihan *k* menggunakan metode *Elbow*. Sebagai bentuk upaya pengembangan dan penyempurnaan penelitian berkelanjutan, oleh karena itu saran yang dapat diberikan oleh peneliti sebagai berikut:

1. Pada penelitian selanjutnya dapat dilakukan metode pengelompokan dan evaluasi hasil pengelompokan yang berbeda agar dapat dilakukan sebagai perbandingan antar metode.
2. Pada penelitian selanjutnya dapat menggunakan data dengan variabel data yang meliputi sumber potensi kesejahteraan sosial yang berasal dari

manusia/masyarakat dan alam/lingkungan baik menggunakan metode yang sama maupun yang berbeda untuk memperoleh hasil yang lebih akurat.

3. Pada penelitian selanjutnya dapat menggunakan data kategorikal untuk memaksimalkan metode evaluasi hasil yang dilakukan oleh peneliti.
4. Hasil penelitian ini dapat menjadi bahan pertimbangan, rujukan, maupun pengembangan teori dan praktik dalam mata kuliah terkait.
5. Penelitian ini diharapkan dapat menjadi masukan strategis bagi Dinas Sosial Kabupaten dan Kota Mojokerto dalam merumuskan, mengevaluasi, atau menyempurnakan program-program sosial.

DAFTAR RUJUKAN

- Ardali, N. R., Zarghami, R., & Gharebagh, R. S. (2024). Optimized data driven fault detection and diagnosis in chemical processes. *Computers and Chemical Engineering*, 186(April), 108712. <https://doi.org/10.1016/j.compchemeng.2024.108712>
- Boutaba, R., Salahuddin, M. A., Limam, N., Ayoubi, S., Shahriar, N., Estrada-Solano, F., & Caicedo, O. M. (2018). A comprehensive survey on machine learning for networking: evolution, applications and research opportunities. *Journal of Internet Services and Applications*, 9(1). <https://doi.org/10.1186/s13174-018-0087-2>
- Dash, C. S. K., Behera, A. K., Dehuri, S., & Ghosh, A. (2023). An outliers detection and elimination framework in classification task of data mining. *Decision Analytics Journal*, 6(January), 100164. <https://doi.org/10.1016/j.dajour.2023.100164>
- Dewi, D. A. I. C., & Pramita, D. A. K. (2019). Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. *Matrix: Jurnal Manajemen Teknologi Dan Informatika*, 9(3), 102–109. <https://doi.org/10.31940/matrix.v9i3.1662>
- Ekasetya, V. A., & Jananto, A. (2020). Klusterisasi Optimal Dengan Elbow Method Untuk Pengelompokan Data Kecelakaan Lalu Lintas Di Kota Semarang. *Jurnal Dinamika Informatika*, 12(1), 20–28. <https://doi.org/10.35315/informatika.v12i1.8159>
- Fathi, H., & Rosmeida, M. (2022). Penerapan Algoritma Weighted Product Method (WPM) pada Sistem Pendukung Keputusan untuk Menentukan Calon Penerima Bantuan Program Keluarga Harapan. *Journal of Informatics and Advanced Computing (JIAC)*, 3(1), 46–51.
- Ghozali, I. (2011). *Ghozali_Imam_2011_Aplikasi_Analisis_Mult.pdf* (p. 129).
- Härdle, W. K., & Simar, L. (2015). *Cluster Analysis BT - Applied Multivariate Statistical Analysis* (W. K. Härdle & L. Simar (eds.); pp. 385–405). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-45171-7_13
- Ilyas, R. (2016). Etika Konsumsi dan Kesejahteraan dalam Perspektif Ekonomi Islam. *Jurnal At-Tawassuth*, 1, 164–165.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*.: Pearson Prentice Hall. In *Pearson Prentice Hall*.

Kemenag. (2024), Qur'an Kemenag

Khakim, E. N. R. (2022). Perbandingan Algoritma Klasifikasi Data Kesejahteraan Sosial Kabupaten Bantul. *Jurnal Processor*, 17(2), 91–100. <http://ejournal.stikom-db.ac.id/index.php/processor/article/view/1222%0Ahttp://ejournal.stikom-db.ac.id/index.php/processor/article/download/1222/896>

Kusuma, D. T., & Agani, N. (2015). Prototipe Komparasi Model Clustering Menggunakan Metode K-Means Dan FCM Untuk Menentukan Strategi Promosi : Study Kasus Sekolah Tinggi Teknik-PLN Jakarta. *Jurnal Ticom*, 3(3), 93460. <https://www.neliti.com/id/publications/93460/>

Lancia, C., Anninga, J., Spitoni, C., Sydes, M. R., Whelan, J., Hogendoorn, P. C. W., Gelderblom, H., & Fiocco, M. (2019). Method to measure the mismatch between target and achieved received dose intensity of chemotherapy in cancer trials: A retrospective analysis of the MRC BO06 trial in osteosarcoma. *BMJ Open*, 9(5), 1–8. <https://doi.org/10.1136/bmjopen-2018-022980>

Lawuna, M. P., Zega, Y., & Mendrofa, R. N. (2024). Analisis Cluster Mahasiswa Pendidikan Matematika Universitas Nias. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, 8(2), 946–961. <https://doi.org/10.31004/cendekia.v8i2.2940>

Maori, N. A., & Evanita, E. (2023). Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 14(2), 277–288. <https://doi.org/10.24176/simet.v14i2.9630>

Maspaitella, M. J., & Rahakbauwi, N. (2014). Pembangunan Kesejahteraan Sosial: Pemberdayaan Masyarakat Dalam Pendekatan Pekerja Sosial. *Aspirasi*, 5(2), 157–164. <https://jurnal.dpr.go.id/index.php/aspirasi/article/view/457/354>

Nella, N. I., Setiawan, N. Y., & Dian, D. E. (2022). Klasifikasi Penerima Bantuan Program Keluarga Harapan menggunakan Algoritme Decision Tree C4. 5 (Studi Kasus: Desa Mlirip Kabupaten Mojokerto). ... *Teknologi Informasi Dan ...*, 6(3), 1332–1339. <https://jptiik.multi.web.id/index.php/j-ptiik/article/view/10801%0Ahttps://jptiik.multi.web.id/index.php/j-ptiik/article/download/10801/4769>

Nufi Mu'tamar Almahmudi. (2019). Konsep Kesejahteraan dan Implementasinya dalam Perspektif Hukum Islam. *Khuluqiyya: Jurnal Kajian Hukum Dan Studi Islam*, 1–19. <https://doi.org/10.56593/khuluqiyya.v1i2.35>

Orleans, Bryan E. P. P. (2022). Clustering Algoritma (K-Means). <https://sis.binus.ac.id/2022/01/31/Clustering-Algoritma-K-Means/>, 1–5.

Purnamasari. (2023). Analisis Nilai-Nilai Pendidikan Dalam QS. Al-Baqarah Ayat 177 pada Mata Pelajaran PAI Kelas X. *Jurnal Pendidikan Profesi Guru*

Agama Islam, 3(4).

- Puspaningtyas, K. (2017). *Pengelompokan Kecamatan Di Wilayah Kabupaten Mojokerto Berdasarkan Jenis Potensi Dan Sumber Kesejahteraan Sosial Menggunakan Metode Cluster*. 66. <http://repository.its.ac.id/41953/>
- Putri, F. I., Damayanti, R., & Kismiantini. (2022). Penerapan Algoritma K-Means Untuk Mengelompokan Kecamatan Di Kabupaten Gunungkidul Berdasarkan Program Keluarga Harapan. *Prosiding Seminar Nasional Matematika, Statistika, Dan Aplikasinya*, 2, 408–418.
- Rahardja, U., Aini, Q., & Iqbal, M. (2020). Analisis Cluster dalam Pengelompokan Provinsi di Indonesia Berdasarkan Variabel Penyakit Menular Menggunakan Metode Complete Linkage, Average Linkage dan Ward. *InfoTekJar : Jurnal Nasional Informatika Dan Teknologi Jaringan*, 5(1), 40–43. <https://doi.org/10.30743/infotekjar.v5i1.2464%0Ahttps://doi.org/10.30743/infotekjar.v5i1.1997>
- Rahma Meilaniwati, E., & Fauzan, M. (2022). Klasifikasi Penduduk Miskin Penerima PKH Menggunakan Metode Naïve Bayes dan KNN. *Jurnal Kajian Dan Terapan Matematika*, 8(2), 75–84. <http://journal.student.uny.ac.id/ojs/index.php/jktm>:
- Rochimah, S., Zain, C., & Murti, D. H. (2006). Informasi Akademik Ftif Its Surabaya. *Juti*, 5(1), 19–25.
- Rohman, A. (2010). *Ekonomi Al-Ghazali : Menelusuri Konsep Ekonomi Islam dalam Ihya Ulum Al-Din*.
- Rosenberg, A., & Hirschberg, J. (2007). V-Measure: A conditional entropy-based external cluster evaluation measure. *EMNLP-CoNLL 2007 - Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, June, 410–420.
- Santoso, S. (2015). *Menguasai statistik multivariat*. Jakarta: PT Elex Media Komputindo.
- Sihombing, P. R., Suryadiningrat, S., Sunarjo, D. A., & Yuda, Y. P. A. C. (2023). Identifikasi Data Outlier (Pencilan) dan Kenormalan Data Pada Data Univariat serta Alternatif Penyelesaiannya. *Jurnal Ekonomi Dan Statistik Indonesia*, 2(3), 307–316. <https://doi.org/10.11594/jesi.02.03.07>
- Silvi, R. (2018). Analisis Cluster dengan Data Outlier Menggunakan Centroid Linkage dan K-Means Clustering untuk Pengelompokkan Indikator HIV/AIDS di Indonesia. *Jurnal Matematika "MANTIK,"* 4(1), 22–31. <https://doi.org/10.15642/mantik.2018.4.1.22-31>
- Solikhun, S., Yasin, V., & Donni Nasution. (2022). Optimization of the Number of

Clusters of the K-Means Method in Grouping Egg Production Data in Indonesia. *International Journal of Artificial Intelligence & Robotics (IJAIR)*, 4(1), 39–47. <https://doi.org/10.25139/ijair.v4i1.4328>

Suarna, N., & Wijaya, Y. A. (2023). *Analisa Penerapan Metode Clustering K-Means Untuk Pengelompokan Data Trransaksi Konsumen (Studi Kasus : Cv . Mitra Indexindo Pratama)*. 7(2), 1322–1328.

Suhartini, S., & Yuliani, R. (2021). Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur. *Infotek : Jurnal Informatika Dan Teknologi*, 4(1), 39–50. <https://doi.org/10.29408/jit.v4i1.2986>

Wijaya, W. T. (2022). *The Analysis Of Relationship Between Height Of Body And Gain Of Medals In Olympics*

Yahya bin Syaraf al-Nawawi. (1984). *Syarh Matan al-Arba'in al-Nawawiyyah fi al-Ahadits al-Shahihah al-Nabawiyyah* (pp. 1–116).

Yuhenny, Dwi. (2024). Profil Kemiskinan di Kabupeten Mojokerto Maret 2024. 43, 1-12.

Zaki, A., Irwan, I., & Sembe, I. A. (2022). Penerapan K-Means Clustering dalam Pengelompokan Data (Studi Kasus Profil Mahasiswa Matematika FMIPA UNM). *Journal of Mathematics Computations and Statistics*, 5(2), 163. <https://doi.org/10.35580/jmathcos.v5i2.38820>

LAMPIRAN

Lampiran 1 Data Potensi Sumber Kesejahteraan Sosial Per Kecamatan di Kabupaten dan Kota Mojokerto

No	Kecamatan	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
1	Kranggan	0	0	0	0	7	0	0	0	0	0	0	0
2	Magersari	0	0	0	0	7	0	0	0	0	0	0	0
3	Prajurit Kulon	0	0	0	0	7	0	0	0	0	0	0	0
4	Dawarblandong	18	41	3	80	18	0	0	128	88	3	1	25
5	Gedeg	0	41	0	183	14	0	28	42	63	4	1	9
6	Jetis	3	26	0	27	14	0	8	97	60	3	1	15
7	Pungging	0	4	5	220	19	0	3	327	19	4	1	4
8	Jatirejo	7	4	3	351	19	1	0	165	29	6	0	7
9	Trowulan	3	17	3	168	16	1	20	65	43	4	1	18
10	Ngoro	0	3	1	79	19	0	0	42	26	2	1	0
11	Kemlagi	0	9	1	174	20	1	0	61	17	0	1	0
12	Pacet	0	4	4	247	20	0	0	563	30	1	0	0
13	Dlanggu	0	30	0	24	16	0	21	91	36	6	1	3
14	Bangsals	4	12	9	52	17	0	0	185	44	0	1	4
15	Sooko	0	15	8	156	15	2	0	60	17	0	1	0
16	Mojoanyar	0	68	9	361	12	0	12	113	59	2	0	0
17	Puri	8	36	0	84	16	0	14	158	17	0	1	19
18	Gondang	0	2	2	0	18	0	0	93	18	0	1	0
19	Mojosari	2	13	2	219	19	0	3	0	9	2	1	4
20	Trawas	3	3	2	81	13	0	0	124	33	0	1	2
21	Kutorejo	1	35	1	181	17	0	4	95	20	1	1	0

Lampiran 2 Hasil Standardisasi Data Potensi Sumber Kesejahteraan Sosial Per Kecamatan di Kabupaten dan Kota Mojokerto

No	Kecamatan	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
1	Kranggan	-0.542	-0.938	-0.853	-1.153	-2.012	-0.442	-0.631	-0.894	-1.297	-0.887	-1.543	-0.691
2	Magersari	-0.542	-0.938	-0.853	-1.153	-2.012	-0.442	-0.631	-0.894	-1.297	-0.887	-1.543	-0.691
3	Prajurit Kulon	-0.542	-0.938	-0.853	-1.153	-2.012	-0.442	-0.631	-0.894	-1.297	-0.887	-1.543	-0.691
4	Dawarblandong	3.639	1.287	0.161	-0.432	0.629	-0.442	-0.631	0.104	2.519	0.584	0.617	2.606
5	Gedeg	-0.542	1.287	-0.853	0.496	-0.332	-0.442	2.652	-0.567	1.435	1.074	0.617	0.496
6	Jetis	0.155	0.473	-0.853	-0.909	-0.332	-0.442	0.307	-0.138	1.305	0.584	0.617	1.287
7	Pungging	-0.542	-0.721	0.837	0.829	0.869	-0.442	-0.279	1.655	-0.473	1.074	0.617	-0.163
8	Jatirejo	1.084	-0.721	0.161	2.009	0.869	1.414	-0.631	0.392	-0.039	2.054	-1.543	0.232
9	Trowulan	0.155	-0.016	0.161	0.361	0.149	1.414	1.714	-0.388	0.568	1.074	0.617	1.683
10	Ngoro	-0.542	-0.776	-0.515	-0.441	0.869	-0.442	-0.631	-0.567	-0.169	0.093	0.617	-0.691
11	Kemlagi	-0.542	-0.450	-0.515	0.415	1.109	1.414	-0.631	-0.419	-0.560	-0.887	0.617	-0.691
12	Pacet	-0.542	-0.721	0.499	1.072	1.109	-0.442	-0.631	3.496	0.004	-0.397	-1.543	-0.691
13	Dlanggu	-0.542	0.690	-0.853	-0.936	0.149	-0.442	1.831	-0.185	0.264	2.054	0.617	-0.295
14	Bangsals	0.387	-0.287	2.188	-0.684	0.389	-0.442	-0.631	0.548	0.611	-0.887	0.617	-0.163
15	Sooko	-0.542	-0.124	1.850	0.253	-0.091	3.269	-0.631	-0.427	-0.560	-0.887	0.617	-0.691
16	Mojoanyar	-0.542	2.753	2.188	2.099	-0.812	-0.442	0.776	-0.013	1.262	0.093	-1.543	-0.691
17	Puri	1.316	1.016	-0.853	-0.396	0.149	-0.442	1.011	0.338	-0.560	-0.887	0.617	1.815
18	Gondang	-0.542	-0.830	-0.177	-1.153	0.629	-0.442	-0.631	-0.169	-0.516	-0.887	0.617	-0.691
19	Mojosari	-0.077	-0.233	-0.177	0.820	0.869	-0.442	-0.279	-0.894	-0.907	0.093	0.617	-0.163
20	Trawas	0.155	-0.776	-0.177	-0.423	-0.572	-0.442	-0.631	0.072	0.134	-0.887	0.617	-0.427
21	Kutorejo	-0.310	0.962	-0.515	0.478	0.389	-0.442	-0.162	-0.154	-0.430	-0.397	0.617	-0.691

Lampiran 3 Hasil Perhitungan *Centroid* Baru dari Setiap Iterasi

Iterasi 1

Titik Pusat	x_1	x_2	x_3	x_4	x_5	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
$C1$	-0.542	-0.938	-0.853	-1.153	-2.012	-0.631	-0.894	-1.297	-0.887	-1.543	-0.691
$C2$	-0.542	1.287	-0.853	0.496	-0.332	2.652	-0.567	1.435	1.074	0.617	0.496
$C3$	-0.542	0.690	-0.853	-0.936	0.149	1.831	-0.185	0.264	2.054	0.617	-0.295
$C4$	-0.542	-0.830	-0.177	-1.153	0.629	-0.631	-0.169	-0.516	-0.887	0.617	-0.691
$C5$	-0.310	0.962	-0.515	0.478	0.389	-0.162	-0.154	-0.430	-0.397	0.617	-0.691

Iterasi 2

Titik Pusat	x_1	x_2	x_3	x_4	x_5	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
$C1$	-0.542	-0.938	-0.853	-1.153	-2.012	-0.631	-0.894	-1.297	-0.887	-1.543	-0.691
$C2$	1.084	0.853	-0.177	0.142	0.149	1.245	-0.284	1.508	0.910	0.617	1.595
$C3$	-0.194	0.582	-0.853	-0.923	-0.091	1.069	-0.162	0.785	1.319	0.617	0.496
$C4$	-0.044	-0.318	0.257	-0.347	0.354	-0.396	-0.089	-0.231	-0.747	0.617	0.220
$C5$	-0.155	0.220	0.499	1.218	0.549	-0.201	0.747	-0.097	0.420	-0.463	-0.361

Iterasi 3

Titik Pusat	x_1	x_2	x_3	x_4	x_5	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
$C1$	-0.542	-0.938	-0.853	-1.153	-2.012	-0.631	-0.894	-1.297	-0.887	-1.543	-0.691
$C2$	1.703	0.763	-0.177	-0.156	0.309	0.698	0.018	0.843	0.257	0.617	2.035
$C3$	-0.310	0.817	-0.853	-0.450	-0.171	1.597	-0.297	1.002	1.237	0.617	0.496
$C4$	-0.252	-0.314	0.245	-0.092	0.449	-0.528	-0.251	-0.299	-0.581	0.617	-0.526
$C5$	-0.136	0.147	0.921	1.503	0.509	-0.191	1.382	0.188	0.706	-1.003	-0.328

Iterasi 4

Titik Pusat	x_1	x_2	x_3	x_4	x_5	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
$C1$	-0.542	-0.938	-0.853	-1.153	-2.012	-0.631	-0.894	-1.297	-0.887	-1.543	-0.691
$C2$	2.478	1.152	-0.346	-0.414	0.389	0.190	0.221	0.980	-0.152	0.617	2.211
$C3$	-0.263	1.038	-0.042	0.222	-0.236	1.456	-0.258	0.967	0.976	0.185	0.496
$C4$	-0.252	-0.314	0.245	-0.092	0.449	-0.528	-0.251	-0.299	-0.581	0.617	-0.526
$C5$	0.000	-0.721	0.499	1.304	0.949	-0.514	1.848	-0.169	0.910	-0.823	-0.207

Lampiran 4 *Syntax Code RStudio* untuk Pengolahan Data Penelitian

```
# Instal Packages
install.packages("factoextra")
install.packages("ggplot")
install.packages("clusterCrit")

# Import Load Library
library(car)
library(factoextra)
library(clusterCrit)
library(ggplot2)
library(cluster)
library(ggrepel)

#Input Data
data <- read_excel("Data_PSKS.xlsx", sheet = "Sheet1")
data_sd <- scale(data)
View(data)

# Uji multikolinieritas
model <- lm(x1 ~ ., data = as.data.frame(data_sd)) #dilakukan berulang sesuai
dengan banyak variabel

# Menghitung nilai VIF untuk setiap variabel
nilai_VIF <- vif(model)
print(nilai_VIF)

# Fungsi untuk mencari titik optimal
fviz_nbclust(data_scaled, kmeans, method = "sse") +
ggtitle("Metode Elbow untuk k")

# Jalankan K-Means dengan k=5
set.seed(42)
kmeans <- kmeans(data_scd, centers = 5, nstart = 25)

# Tampilkan hasil cluster
print(kmeans$cluster)
```

Lanjutan Lampiran 4 *Syntax Code* RStudio untuk Pengolahan Data Penelitian

```
#Syntax Code Python untuk Evaluasi V-Measure
from sklearn.metrics import homogeneity_completeness_v_measure

# Data hasil kelompok dan cluster
kelompok = [1, 1, 1, 2, 3, 3, 5, 5, 3, 4, 4, 5, 5, 4, 4, 5, 4, 4, 4, 4, 4]
cluster = [1, 1, 3, 2, 2, 2, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 2, 2, 4, 2, 5]

# Menghitung homogeneity, completeness, dan V-Measure
homogeneity, completeness, v_measure =
homogeneity_completeness_v_measure(cluster, kelompok)

# Menampilkan hasil
print(f'Homogeneity: {homogeneity:.3f}')
print(f'Completeness: {completeness:.3f}')
print(f'V-Measure Score: {v_measure:.3f}')
```

RIWAYAT HIDUP



Riza Mar'atus Sholihah, lahir di Sidoarjo pada 28 Juli 2002, biasa dipanggil Riza. Peneliti merupakan anak kedua dari Bapak Kusnadi Ridlo dan Ibu Nurul Hidayah serta adik dari Fadlilatul Ilmillah. Peneliti telah menempuh pendidikan mulai dari RA Perwanida Krembung yang lulus pada tahun 2008, dilanjutkan menempuh pendidikan sekolah dasar di SDN Krembung 1 dan lulus pada tahun 2014. Kemudian peneliti melanjutkan pendidikan sekolah menengah pertama di MTs. Salafiyah Mojogeneng Mojokerto dan lulus pada tahun 2017. Setelah itu melanjutkan pendidikan sekolah menengah atas di MAS Al Ma'arif Singosari Malang dan lulus pada tahun 2020. Pada tahun yang sama, peneliti melanjutkan pendidikan di Universitas Islam Negeri Maulana Malik Ibrahim Malang pada Program Studi Matematika, Fakultas Sains dan Teknologi. Selain fokus pada bidang akademik, peneliti juga aktif dalam kegiatan pengajaran dengan memberikan bimbingan belajar privat kepada siswa sekolah sebagai bentuk kontribusi di bidang pendidikan.



KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayana No.50 Dinoyo Malang Telp. / Fax. (0341)558933

BUKTI KONSULTASI SKRIPSI

Nama : Riza Mar'atus Sholihah
NIM : 200601110054
Fakultas / Jurusan : Sains dan Teknologi / Matematika
Judul Skripsi : Penerapan *K-Means Clustering* dengan Evaluasi *F-Measure*
untuk Pengelompokan Wilayah Berdasarkan Data Potensi
Sumber Kesejahteraan Sosial
Pembimbing I : Dr. Fachrur Rozi, M.Si.
Pembimbing II : Dr. Abdussakir, M.Pd

No	Tanggal	Hal	Tanda Tangan
1.	29 Mei 2024	Konsultasi Topik dan Data	1.
2.	04 Juni 2024	Konsultasi Bab I, II, dan III	2.
3.	06 Juni 2024	Konsultasi Bab I, II, dan III	3.
4.	07 Agustus 2024	Konsultasi Kajian Agama Bab I dan II	4.
5.	09 Agustus 2024	Konsultasi Bab I, II, dan III	5.
6.	13 Agustus 2024	ACC Kajian Agama Bab I dan II	6.
7.	19 Agustus 2024	Konsultasi Bab I, II, dan III	7.
8.	29 Agustus 2024	ACC Bab I, II, dan III	8.
9.	30 Agustus 2024	ACC Seminar Proposal	9.
10.	03 Desember 2024	Konsultasi Revisi Seminar Proposal	10.
11.	18 Desember 2024	Konsultasi Bab IV dan V	11.
12.	10 Februari 2025	Konsultasi Kajian Agama Bab IV	12.
13.	14 Februari 2025	Konsultasi Bab IV dan V	13.
14.	21 Maret 2025	Konsultasi Bab IV dan V	14.



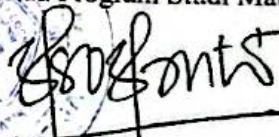
KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayana No.50 Dinoyo Malang Telp. / Fax. (0341)558933

15.	17 April 2025	ACC Bab IV dan V	15. <i>FP</i>
16.	18 April 2025	ACC Seminar Hasil	16. <i>FP</i>
17.	23 April 2025	ACC Kajian Agama Bab IV	17. <i>Q</i>
18.	27 Mei 2025	Konsultasi Revisi Seminar Hasil	18. <i>FP</i>
19.	03 Juni 2025	Konsultasi Revisi Seminar Hasil	19. <i>FP</i>
20.	10 Juni 2025	Konsultasi Revisi Seminar Hasil Kajian Agama	20. <i>Q</i>
21.	13 Juni 2025	ACC Sidang Skripsi	21. <i>FP</i>
22.	18 Juni 2025	ACC Keseluruhan	22. <i>FP</i>

Malang, 18 Juni 2025

Mengetahui,

Ketua Program Studi Matematika




Dr. Elly Susanti, M.Sc.

NIP. 19741129 200012 2 005