

Implementasi *K-Means Cluster Analysis* untuk Mengelompokkan Kabupaten/Kota Berdasarkan Data Kemiskinan di Provinsi Jawa Tengah Tahun 2022

S N Mayasari¹, J Nugraha²

^{1,2} Program Studi Statistika, Universitas Islam Indonesia

E-mail: septiana.mayasari01@gmail.com¹, jaka.nugraha@uii.ac.id²

Abstrak. Salah satu masalah utama yang masih dihadapi oleh Indonesia adalah masalah kemiskinan. Jawa Tengah merupakan Provinsi termiskin kedua di Pulau Jawa, walaupun pada tahun 2022 masih berada di atas rata-rata tingkat kemiskinan nasional. Pemetaan secara lebih detail di tingkat Kabupaten/Kota diperlukan untuk memudahkan dalam mengenali tingkat kemiskinan suatu wilayah serta mengetahui daerah mana yang membutuhkan bantuan, sehingga pemerintah dapat secara tepat dan cepat mengambil kebijakan untuk mengatasi kemiskinan di daerah tersebut. Tujuan dari penelitian ini adalah untuk mengelompokkan Kabupaten/Kota di Provinsi Jawa Tengah berdasarkan data kemiskinan guna mengetahui karakteristik dari masing-masing wilayah. Penelitian ini dilakukan dengan menggunakan metode *K-Means Cluster Analysis* yaitu salah satu metode yang dapat digunakan untuk mengelompokkan kategori kemiskinan Kabupaten/Kota di Provinsi Jawa Tengah. Data yang digunakan yaitu jumlah penduduk miskin, laju pertumbuhan penduduk, dan tingkat pengangguran terbuka tahun 2022 yang diperoleh dari *Website* Badan Pusat Statistik Provinsi Jawa Tengah. Hasil dari penelitian dengan mengelompokkan Kabupaten/Kota ke dalam 3 *cluster*, yaitu didapatkan anggota dari *cluster* 1 terdiri dari 12 Kabupaten/Kota dengan kategori kemiskinan rendah, *cluster* 2 terdiri dari 7 Kabupaten/Kota dengan kategori kemiskinan tinggi, dan *cluster* 3 terdiri dari 16 Kabupaten/Kota dengan kategori kemiskinan sedang.

Kata kunci: kemiskinan; pengelompokan; analisis klaster K-Means

Abstract. One of the main problems that Indonesia is still facing is the problem of poverty. Central Java is the second poorest province on the island of Java, even though in 2022 it was still above the national average poverty rate. Mapping in more detail at the Regency/City level is needed to make it easier to identify the poverty level of an area and to know which areas need assistance, so that the government can make appropriate and fast policies to address poverty in that area. The purpose of this research is to classify districts/cities in Central Java province based on poverty data to determine the characteristics of each region. This research was conducted using the *K-Means Cluster Analysis* method, which is a method that can be used to classify district/city poverty categories in Central Java Province. The data used are the number of poor people, the rate of population growth, and the open unemployment rate in 2022 obtained from the Central Java Province Statistics Agency website. The results of the study grouped districts/cities into 3 clusters, namely the members of cluster 1 consisted of 12 regencies/cities with a low poverty category, cluster 2 consisted of 7 regencies/cities with a high poverty category, and cluster 3 consisted of 16 regencies/cities with moderate poverty category.

Keywords: poverty; grouping; *K-Means cluster analysis*

1. Pendahuluan

Kemiskinan merupakan masalah utama yang dihadapi oleh pemerintahan di hampir semua negara, salah satunya adalah Indonesia [1]. Indonesia adalah negara yang tergolong masih berkembang, di mana kemiskinan merupakan masalah yang cukup serius dan masih menjadi pusat perhatian. Maka dari itu, perlu adanya upaya untuk memecahkan masalah tersebut, di mana upaya pengentasan kemiskinan menjadi salah satu aspek penting dalam mencapai kesejahteraan rakyat. Pada tingkat nasional maupun regional, kemiskinan telah menjadi masalah yang rumit dan berkepanjangan, sehingga masalah tersebut menjadi prioritas utama pembangunan yang memerlukan strategi penyelesaian yang tepat dan berkelanjutan untuk mengatasinya, maka dari itu cara penanggulangan kemiskinan pun membutuhkan analisis yang tepat yang melibatkan semua komponen permasalahan [2]. Pemerintah Indonesia selama ini sudah memiliki banyak program-program pembangunan dan bantuan yang sudah terlaksana, di mana program-program tersebut telah memberikan kontribusi besar terhadap pengentasan kemiskinan [3]. Upaya pengentasan kemiskinan ini dilakukan untuk mewujudkan cita-cita bangsa Indonesia yaitu terciptanya masyarakat yang sejahtera, adil, dan makmur [4]. Meskipun demikian, masalah kemiskinan di Indonesia sampai saat ini masih menjadi masalah yang belum terselesaikan.

Kemiskinan merupakan salah satu indikator kesejahteraan, di mana suatu negara atau daerah dikatakan sejahtera apabila memiliki persentase penduduk miskin yang rendah [5]. Menurut data BPS pada tahun 2022 [6], dari 35 Kabupaten/Kota yang ada di Provinsi Jawa Tengah tercatat 17 kabupaten/kota memiliki persentase penduduk miskin di atas angka Provinsi yaitu 10,93% dan ada 19 kabupaten/kota memiliki persentase penduduk miskin di atas angka Nasional yaitu 9,57%. Jawa Tengah merupakan Provinsi termiskin kedua di pulau Jawa dengan persentase tingkat penduduk miskin pada bulan September 2022 sebesar 10,98%, naik sebesar 0,05% dibandingkan bulan Maret 2022.

Kondisi kemiskinan suatu negara atau wilayah merupakan cerminan dari tingkat kesejahteraan penduduk yang tinggal pada suatu negara atau wilayah tersebut [7]. Jumlah penduduk merupakan isu utama dalam pembangunan ekonomi suatu daerah. Menurut [8] tingkat kemiskinan di Jawa Tengah dipengaruhi secara positif dan signifikan oleh jumlah penduduk. Pertumbuhan penduduk akan menghasilkan perluasan angkatan kerja. Ketika terjadi pertumbuhan jumlah angkatan kerja tanpa diimbangi dengan peningkatan kesempatan kerja yang tersedia, maka sejumlah besar individu dalam angkatan kerja akan menganggur. Pengangguran ini kemudian dapat berkontribusi pada kemiskinan. Oleh karena itu, terdapat korelasi positif antara jumlah penduduk dan tingkat kemiskinan.[9] menyatakan bahwa apabila jumlah penduduk meningkat, maka jumlah penduduk miskin juga akan bertambah. Semakin banyak penduduk miskin di suatu daerah, maka semakin rendah tingkat kesejahteraan di daerah tersebut. Salah satu penyebab banyaknya penduduk miskin yaitu bertambahnya jumlah penduduk tiap tahun yang tidak diimbangi dengan pendapatan. Apabila laju pertumbuhan penduduk tidak terkendali, maka akan mengakibatkan tujuan pembangunan ekonomi yaitu kesejahteraan rakyat dan pengentasan kemiskinan tidak akan tercapai [8].

Semakin tinggi laju pertumbuhan penduduk, maka dapat menyebabkan kesenjangan yang berdampak terhadap meningkatnya kemiskinan [10]. Laju pertumbuhan penduduk yang cukup besar akan mempengaruhi kondisi sosial masyarakat, salah satunya adalah banyak munculnya pengangguran. Pengangguran saat ini masih menjadi masalah utama dalam hubungannya dengan kemiskinan, karena pengangguran merupakan pintu gerbang menuju kemiskinan. Menurut [11] tingkat pengangguran berpengaruh positif dan signifikan terhadap kemiskinan di Jawa Tengah. [12] menyatakan bahwa bukti-bukti empiris selalu menunjukkan bahwa pengangguran secara langsung akan berdampak kepada meningkatnya kemiskinan. Pengangguran itu sendiri juga disebabkan oleh banyak hal, antara lain rendahnya sumber daya manusia sehingga kalah bersaing dengan yang memiliki kemampuan, kecenderungan memperoleh gaji yang hanya cukup untuk memenuhi kehidupan sehari-hari sehingga tingkat kesejahteraannya rendah. Hal ini menyiratkan bahwa terdapat korelasi positif antara jumlah pengangguran dan prevalensi kemiskinan. Ketika memeriksa faktor-faktor yang berkontribusi terhadap tingkat kemiskinan, jelaslah bahwa pengangguran bukanlah satu-satunya penyebab. Sebaliknya, hal ini dapat dikaitkan dengan jumlah penduduk yang besar atau, dengan kata lain, anggapan bahwa kuantitas tidak selalu mengimbangi kualitas [13].

Mengingat tingginya angka kemiskinan di Provinsi Jawa Tengah, maka sangat penting untuk melakukan inisiatif penanggulangan kemiskinan. Ketersediaan data statistik kemiskinan yang akurat

dan pemahaman yang komprehensif mengenai karakteristik daerah merupakan elemen penting dalam memfasilitasi upaya pengentasan kemiskinan yang berhasil.

Pemerintah perlu untuk mengetahui tingkat kemiskinan yang berbeda-beda di berbagai daerah. Informasi ini sangat penting untuk mengalokasikan sumber daya secara efektif dan memfokuskan upaya-upaya untuk mengatasi masalah-masalah yang berkaitan dengan kemiskinan. Oleh karena itu, sangat penting untuk mengetahui atribut-atribut khas dari setiap kabupaten/kota di Provinsi Jawa Tengah untuk meningkatkan efektivitas inisiatif penanggulangan kemiskinan. Pengelompokan kemiskinan berfungsi sebagai metode untuk membedakan atribut-atribut yang berbeda terkait status kesejahteraan individu di berbagai daerah, sehingga memudahkan perumusan kebijakan dan inisiatif pembangunan yang terfokus. Beberapa penelitian telah dilakukan untuk mengusulkan pemanfaatan teknik pengelompokan untuk tujuan pengelompokan kabupaten atau kota di Provinsi Jawa Tengah, dengan fokus pada statistik kemiskinan. Peneliti [14] mengusulkan menggunakan metode *Hierarchical Clustering* (metode hirarki) yaitu metode *Single Linkage* dan *Ward*, serta metode *Non-Hierarchical Clustering* (tak hirarki) atau *Partitioning* (metode sekatan) yaitu metode *K-Means* dan *Fuzzy C-Means* (FCM) untuk mengelompokkan Kabupaten/Kota di Jawa Tengah Berdasarkan Faktor-Faktor yang Mempengaruhi Kemiskinan tahun 2016 yang menghasilkan 4 kelompok dan didapatkan hasil bahwa metode terbaik adalah metode *K-Means*. Peneliti [15] melakukan perbandingan antara metode *K-Means* dan *Fuzzy C-Means* untuk pengelompokan kemiskinan di Provinsi Jawa Tengah, dan hasil penelitian tersebut juga menyatakan bahwa metode *K-means* lebih baik daripada metode *Fuzzy C-Means*. Berdasarkan hasil penelitian terdahulu metode *K-Means* lebih unggul daripada metode yang lain untuk pengelompokan kemiskinan di Jawa Tengah.

Penelitian lain tentang pengelompokan kemiskinan di Jawa Tengah yaitu penelitian yang dilakukan oleh [16] tentang Pemetaan Kemiskinan Kabupaten/Kota Di Provinsi Jawa Tengah Menggunakan Algoritma *K-Means*. Penelitian ini menjelaskan pemanfaatan data indikator kemiskinan, yaitu proporsi individu yang hidup dalam kemiskinan, rumah tangga yang terdaftar dalam program Raskin, indeks keparahan kemiskinan, dan ambang batas kemiskinan, yang akan disusun ke dalam dua kelompok yang berbeda. Studi ini menghasilkan data yang berkaitan dengan pengelompokan kemiskinan di provinsi Jawa Tengah, di mana dua kelompok yang berbeda diidentifikasi. Klaster pertama terdiri dari 22 kabupaten/kota yang diklasifikasikan sebagai tidak miskin, sedangkan klaster kedua terdiri dari 13 kabupaten/kota yang diklasifikasikan sebagai miskin. Oleh karena itu, penelitian ini menggunakan algoritma *K-Means* untuk melakukan analisis pengelompokan tingkat kemiskinan di wilayah Jawa Tengah, dengan menggunakan data kemiskinan tahun 2022.

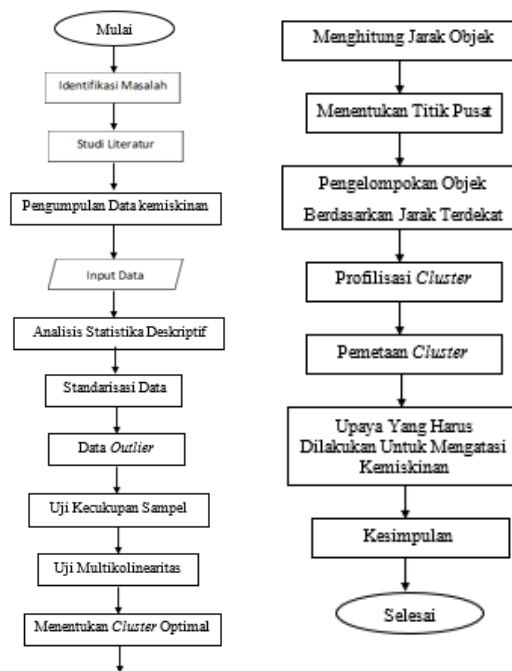
2. Metode

2.1. Data dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang berjenis data *cross section*, diperoleh dari Badan Pusat Statistik Provinsi Jawa Tengah yang bersumber dari *Website* resmi BPS Provinsi Jawa Tengah yang dapat diakses melalui laman <https://jateng.bps.go.id/>. Data yang digunakan yaitu data kemiskinan di Provinsi Jawa Tengah pada tahun 2022. Dengan menggunakan variabel jumlah penduduk miskin (JPM), laju pertumbuhan penduduk (LPP), dan tingkat pengangguran terbuka (TPT).

2.2. Diagram Alir Penelitian

Metode penelitian yang digunakan dalam penelitian ini yaitu menggunakan studi literatur dari jurnal, buku, maupun skripsi untuk menentukan variabel-variabel yang digunakan. Kemudian dilakukan pengelompokan Kabupaten/Kota di Jawa Tengah guna mengetahui kategori kemiskinan dengan menggunakan metode *K-Means Cluster*. Dengan demikian, data kemiskinan yang memiliki karakteristik data yang sama dikelompokkan dan dilihat berdasarkan pada indikator rata-rata nilai setiap variabel dengan demikian diharapkan pemerintah dapat segera mengatasi masalah kemiskinan yang ada agar tidak terjadi ketimpangan kemiskinan.



Gambar 1. Diagram Alir Penelitian

Gambar 1 di atas merupakan diagram alir dari penelitian yang dilakukan. Penelitian ini dimulai dengan mengidentifikasi masalah yang ada di sekitar, yang dilanjutkan dengan studi literatur untuk mendapatkan variabel data penelitian yang digunakan, setelah itu dilanjutkan dengan persiapan data penelitian yaitu berupa data kemiskinan di Provinsi Jawa Tengah tahun 2022. Kemudian menginputkan data ke dalam *Software RStudio* di mana *software* tersebut adalah alat yang digunakan untuk melakukan analisis (*K-Means Cluster Analysis*). Selanjutnya melakukan analisis statistika deskriptif untuk mengetahui gambaran umum atau ringkasan data dari data kemiskinan di Provinsi Jawa Tengah tahun 2022. Melakukan standarisasi data yang digunakan untuk standarisasi data karena pada data terdapat perbedaan satuan. Selanjutnya dari data yang telah distandarisasi, kemudian menguji apakah pada data terdapat data *outlier* atau tidak. Apakah data yang terdapat *outlier* akan dibuang atau akan tetap digunakan tetapi harus dengan ada alasan yang mendukung. Berikutnya yaitu melakukan pengujian asumsi analisis *K-Means Cluster*. Yang pertama yaitu melakukan pengecekan uji kecukupan sampel dengan menggunakan *Kaiser Meyer Olkin (KMO)*, asumsi tersebut terpenuhi maka dapat dikatakan sampel representatif dan dapat dilanjutkan untuk pengujian asumsi yang kedua. Untuk uji asumsi yang kedua yaitu uji multikolinearitas yang harus non multikolinearitas dengan menggunakan nilai koefisien korelasi. Kedua asumsi tersebut dipenuhi maka dapat dilakukan analisis menggunakan *K-Means Cluster*. Selanjutnya menentukan jumlah *cluster* yang optimum dengan membandingkan hasil dari metode *elbow* dan *library NbClust*, kemudian dipilih jumlah *cluster* yang optimal. Setelah itu menghitung jarak objek, menentukan titik pusat *cluster* dan mengelompokkan objek berdasarkan jarak terdekat. Setelah hasil kelompok sudah terbentuk maka dapat dilakukan profilisasi *cluster* untuk mengetahui karakteristik pada setiap *cluster* yang terbentuk serta melakukan pemetaan terkait penyebaran hasil dari analisis *K-Means Cluster* dengan menggunakan *Software QGis*. Dari hasil pengelompokan dan karakteristik dari tiap *cluster*-nya maka upaya apa yang harus dilakukan pemerintah dalam mengatasi kemiskinan. Yang terakhir yaitu memperoleh kesimpulan dari hasil analisis *clustering* tersebut dan memberikan saran perbaikan di masa depan sehingga penelitian telah dianggap selesai.

2.3. Metode Analisis

Metode yang digunakan dalam penelitian ini adalah metode *K-means* yang merupakan salah satu metode *clustering* (pengelompokan). *K-Means Clustering* adalah teknik pengelompokan data *non-hierarkis* yang mengelompokkan data ke dalam satu atau lebih *cluster*/kelompok. Data dengan

karakteristik yang sama dikelompokkan dalam satu kluster/kelompok, sedangkan data dengan karakteristik berbeda dikelompokkan dengan kluster/kelompok lain, sehingga data yang berada dalam satu *cluster*/kelompok memiliki tingkat variasi yang kecil [17].

2.3.1 Analisis Deskriptif

Analisis deskriptif merupakan metode analisis yang melibatkan proses pengumpulan, pengolahan, penyajian, dan interpretasi data kuantitatif atau persentase yang direpresentasikan dalam bentuk tabel atau grafik [18]. Pendekatan analisis deskriptif ini digunakan untuk menyelidiki data set dengan tujuan memberikan gambaran komprehensif mengenai data tersebut, sehingga data yang disajikan menjadi lebih mudah dipahami dan memberikan informasi yang berarti bagi pembaca.

2.3.2 Asumsi Analisis Cluster

Menurut [19], terdapat 2 asumsi untuk analisis *cluster* yaitu:

1. Sampel representatif

Kumpulan data yang digunakan dalam analisis kluster terdiri dari data populasi. Selain penggunaan data sampel, data yang digunakan haruslah representatif. Sampel yang representatif adalah sampel yang dipilih dengan cara yang memungkinkan mereka untuk secara akurat mewakili atau menggambarkan populasi yang lebih besar dari mana sampel tersebut diambil. Tidak adanya pedoman eksplisit tentang jumlah sampel yang representatif mengharuskan penggunaan ukuran sampel yang cukup besar untuk memastikan pelaksanaan prosedur pengelompokan yang akurat. Uji Kaiser-Mayer-Olkin (KMO) dapat digunakan untuk menilai keterwakilan sampel pengujian dalam kaitannya dengan populasi. Uji Kaiser-Mayer-Olkin (KMO) biasanya digunakan untuk menilai kesesuaian sampel yang diberikan dalam memenuhi standar yang diperlukan. Uji KMO menilai kecukupan pengambilan sampel secara keseluruhan dan kesesuaian pengambilan sampel untuk masing-masing indikator. Uji Kaiser-Mayer-Olkin (KMO) adalah ukuran statistik yang berkisar antara 0 hingga 1, yang menunjukkan tingkat kecukupan data untuk melakukan analisis faktor. Ketika nilai Kaiser-Meyer-Olkin (KMO) berada di kisaran 0,5 hingga 1. Hal ini menunjukkan bahwa sampel dapat dianggap mewakili populasi atau sampel yang representatif.

2. Non Multikolinearitas

Multikolinearitas merujuk pada keberadaan hubungan linear yang kuat atau sempurna antara beberapa atau bahkan semua variabel dalam suatu analisis. Untuk mengidentifikasi adanya multikolinearitas, salah satu metode yang dapat digunakan adalah dengan memeriksa koefisien korelasi di antara setiap variabel. Menurut Gujarati dalam [20] dikatakan terjadi multikolinearitas jika terdapat koefisien korelasi sederhana yang mencapai atau melebihi 0,8. Sebaiknya dalam analisis cluster tidak terjadi multikolinearitas. Jika terjadi multikolinearitas, maka dapat diatasi dengan menggunakan *Principal Component Analysis* (PCA).

2.3.3 Penentuan Jumlah Cluster

Untuk pemilihan jumlah *cluster*/kelompok yang optimal, penelitian ini menggunakan metode *elbow* di mana metode ini sangat praktis untuk memilih jumlah *cluster* k yang digunakan untuk pengelompokan data pada algoritma *K-Means*. Metode *Elbow* merupakan salah satu metode terbaik untuk menentukan jumlah *cluster* dengan cara melihat persentase hasil perbandingan antara jumlah *cluster* yang membentuk siku pada suatu titik [21]. Metode *elbow* ini dapat dihasilkan dari perbandingan hasil SSE (*Sum of Squared Error*) dengan rumus SSE seperti di bawah ini [22]:

$$SSE = \sum_{k=1}^k \sum_{x_i \in S_k} \|x_i - C_k\|^2 \quad (1)$$

di mana:

k adalah banyaknya kelompok yang digunakan pada algoritma *K-Means*

x_i adalah jumlah data

C_k adalah banyaknya clusteri pada cluster ke k .

2.3.4 Analisis Cluster K-Means

K-Means adalah suatu pendekatan dalam analisis pengelompokan data non-hirarkis yang berusaha untuk mengelompokkan data yang ada menjadi beberapa kelompok, dengan tujuan untuk menggabungkan data yang memiliki sifat serupa dalam satu kelompok yang sama, sementara data yang memiliki perbedaan sifat dikelompokkan dalam kelompok yang berbeda [23]. Menurut [24], analisis *cluster k-means* ini dimulai dengan menentukan jumlah *cluster* terlebih dahulu, sesuai yang diinginkan misalnya k *cluster*. Untuk mengelompokkan data dengan menggunakan metode *K-Means* dilakukan dengan cara sebagai berikut [25]:

1. Menentukan banyaknya k sebagai jumlah klaster yang ingin dibentuk.
 Penentuan jumlah cluster k dalam analisis K-Means dapat dilakukan melalui berbagai pertimbangan, termasuk pertimbangan teoritis dan konseptual yang dapat diusulkan untuk menentukan jumlah cluster yang tepat.
2. Menentukan k centroid (titik pusat *cluster*) awal secara *random*/acak.
 Penentuan awal centroid dilakukan secara acak dari objek-objek yang tersedia sebanyak k cluster. Selanjutnya, untuk menghitung centroid dari cluster ke- i selanjutnya, rumus yang digunakan adalah sebagai berikut:

$$C_{kj} = \frac{x_{1j} + x_{2j} + \dots + x_{nj}}{n} \quad (2)$$

di mana:

C_{kj} adalah pusat klaster ke- k pada variabel ke- j ($j = 1, 2, \dots, p$)

n adalah banyak data pada klaster ke- k .

3. Menentukan jarak dari setiap objek ke masing-masing centroid dari masing-masing *cluster*. Untuk menghitung jarak antara objek dengan centroid yaitu dengan menggunakan jarak euclidean.

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_p - y_p)^2} \quad (3)$$

4. Menghitung fungsi objektif

$$J = \sum_{i=1}^n \sum_{j=1}^k a_{ij} d(x_i, C_{kj})^2 \quad (4)$$

di mana:

n adalah jumlah data

k adalah jumlah *cluster*

a_{ij} adalah nilai keanggotaan titik data x_i ke pusat *cluster* C_{kj}

C_{kj} adalah pusat *cluster* ke- kj

$d(x_i, C_{kj})$ adalah jarak titik x_i ke *cluster* C_{kj} yang diikuti.

5. Mengalokasikan masing-masing data ke centroid/rata-rata terdekat, dengan menggunakan rumus seperti di bawah ini:

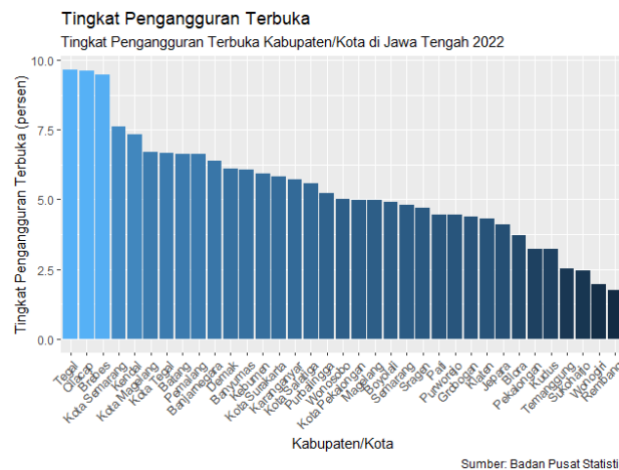
$$a_{ij} = \begin{cases} 1, & s = \min \{d(x_i, C_{kj})\} \\ 0, & \text{lainnya} \end{cases} \quad (5)$$

di mana:

a_{ij} adalah nilai keanggotaan titik x_i ke pusat *cluster* C_{kj}

s adalah jarak terpendek dari data x_i ke pusat *cluster*

Dari Gambar 3. dapat disimpulkan bahwa laju pertumbuhan penduduk pada tahun 2022 yang paling tinggi yaitu Kota Semarang sebesar 1.42% yang ditandai dengan batang tertinggi dan berwarna terang. Sedangkan laju pertumbuhan penduduk pada tahun 2022 paling rendah yaitu Kabupaten Kebumen, Kabupaten Wonogiri, dan Kabupaten Tegal sebesar 0.07% yang ditandai dengan batang terendah dan berwarna gelap.



Gambar 4. Grafik Variabel Tingkat Pengangguran Terbuka

Dari Gambar 4. dapat disimpulkan bahwa tingkat pengangguran terbuka pada tahun 2022 yang paling banyak yaitu Kabupaten Tegal sebesar 9.64% yang ditandai dengan batang tertinggi dan berwarna terang. Tingkat pengangguran terbuka Kabupaten Tegal sedikit di atas Kabupaten Cilacap dengan 9,62% kemudian diikuti Kabupaten Brebes sebesar 9,48%, sedangkan tingkat pengangguran terbuka pada tahun 2022 paling rendah yaitu Kabupaten Rembang sebesar 1.76% yang ditandai dengan batang terendah dan berwarna gelap.

3.2. Pengujian Asumsi Analisis Cluster

Sebelum melakukan analisis *cluster*, diperlukan pemenuhan dua asumsi yaitu yang pertama, jumlah populasi/sampel harus dapat merepresentasikan keadaan sesungguhnya (representatif). Kedua, tidak boleh adanya korelasi antar variabel penelitian (*non* multikolinearitas).

3.2.1 Uji Kecukupan Data dan Kelayakan Variabel

Asumsi pertama yaitu jumlah sampel harus dapat mewakili populasi (representatif). Uji asumsi ini dilakukan menggunakan nilai dari *Kaiser Meyer-Olkin* (KMO). Uji *Kaiser-Mayer-Olkin* (KMO) memiliki rentang nilai dari 0 hingga 1, apabila nilai KMO memiliki rentang dari 0,5 hingga 1 maka sampel data tersebut dapat dikatakan mewakili populasi atau sampel representatif atau sampel dapat mewakili data.

	MSA (<i>Measure Sampling Adequacy</i>)
<i>Overall</i>	0,56
Jumlah Penduduk Miskin	0,54
Laju Pertumbuhan Penduduk	0,59
Tingkat Pengangguran Terbuka	0,56

Dari Tabel 1. dapat diketahui bahwa nilai KMO dari ketiga variabel yaitu sebesar 0,56 dan nilai KMO pada setiap variabel di dalam rentang 0,5 hingga . Dapat disimpulkan bahwa sampel data tersebut dapat dikatakan mewakili populasi atau sampel representatif sehingga layak untuk dilakukan analisis *cluster*.

3.2.2 Hasil Uji Multikolinearitas

Pengaruh multikolinearitas antarvariabel merupakan suatu pelanggaran dalam analisis *cluster*, karena dapat mempengaruhi hasil *cluster*. Uji multikolinieritas berguna untuk mengetahui apakah antarvariabel saling berhubungan atau dengan kata lain mempunyai kemiripan antar variabel karena data yang baik adalah data yang bebas korelasi antara variabel independennya. Dikatakan terjadi multikolinearitas jika terdapat nilai koefisien korelasi yang mencapai atau melebihi dari 0,8. Hasil uji asumsi multikolinearitas dengan melihat nilai koefisien korelasi dari variabel-variabelnya.

Tabel 2. Uji Multikolinearitas

	JPM	LPP	TPT
JPM	1	-0,28	0,33
LPP		1	-0,12
TPT			1

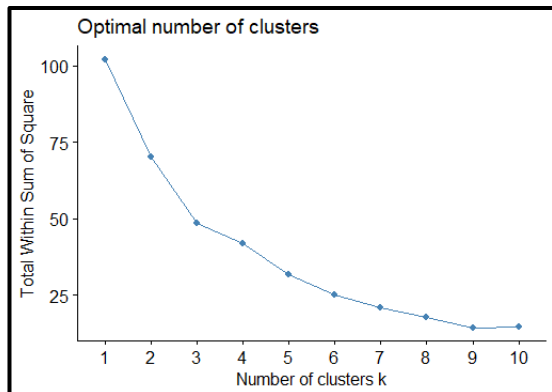
Dari Tabel 2. dapat diketahui bahwa nilai koefisien korelasi antarvariabel kurang dari 0,8, sehingga dapat disimpulkan bahwa tidak terjadi multikolinearitas antar variabelnya (*non* multikolinearitas).

3.3. Hasil Analisis Cluster

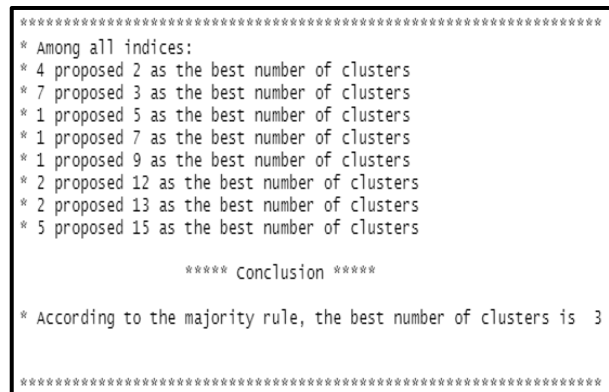
Karena pengujian asumsi *Kaiser-Mayer-Olkin* (KMO) dan multikolinearitas sudah terpenuhi yaitu sampel dapat mewakili data (*representatif*) dan tidak adanya multikolinearitas antar variabel satu dengan variabel lain maka dapat dilanjutkan ke tahap analisis *cluster*.

3.3.1 Penentuan Jumlah Cluster

Penentuan jumlah *cluster* atau banyaknya nilai *k*, yang dapat ditentukan langsung oleh peneliti atau dapat berdasarkan beberapa pendekatan untuk mendapatkan nilai *k* yang optimal. Salah satunya yaitu dengan menggunakan metode *elbow* atau *within cluster sum of square*. Dalam menentukan nilai *k* pada metode *elbow*, dilihat dari titik jumlah *k* cluster yang membentuk siku. Selain dengan menggunakan metode *elbow*, peneliti juga menggunakan metode lain yaitu dengan menggunakan *library NbClust*.



Gambar 5. Penentuan jumlah *cluster* dengan metode *elbow*



Gambar 6. Penentuan jumlah *cluster* dengan *library NbClust*

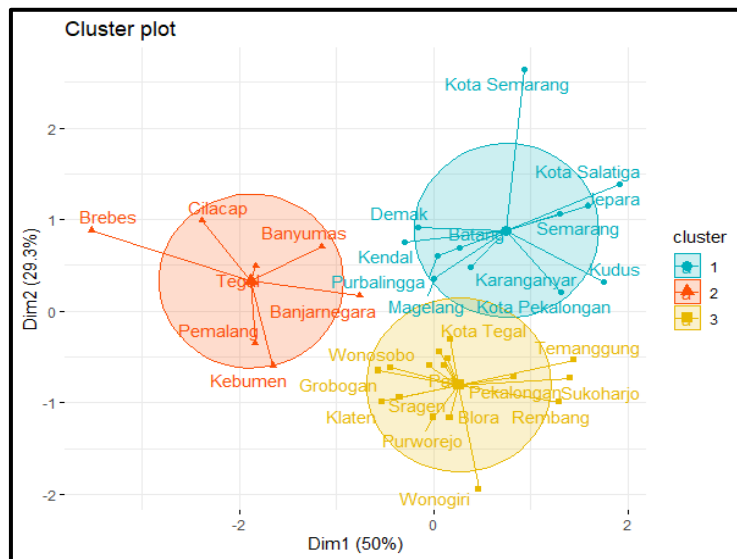
Gambar 5. Menunjukkan bahwa nilai *k* optimal yaitu berada pada angka 3 yang ditandai dengan titik *cluster* yang membentuk siku, sehingga diperoleh nilai *k* yang optimum pada metode *elbow* yaitu 3. Serta Gambar 6. menunjukkan bahwa jumlah *cluster* yang optimal dengan menggunakan *library NbClust* yaitu 3 *cluster*.

Pembentukan *cluster* ini sudah sesuai dengan keinginan peneliti yang membentuk *cluster* menjadi 3, karena dalam pembentukan *cluster* dengan metode *K-Means* ini dapat ditentukan clusternya terlebih dahulu yang sesuai dengan keinginan. Oleh karena itu, dalam penelitian ini penulis

menggunakan 3 *cluster* untuk mengelompokkan Kabupaten/Kota di Provinsi Jawa Tengah berdasarkan data kemiskinan.

3.3.2 Anggota Pada Setiap Cluster

Analisis *cluster* dengan metode *K-Means* harus menggunakan data yang telah distandarisasi. Dalam menentukan nilai *k* optimum dengan menggunakan metode *elbow* atau *within cluster sum of square* dan menggunakan *library NbClust* yang didapatkan hasil nilai *k* optimum adalah 3, maka akan terdapat 3 pengelompokan Kabupaten/Kota di Provinsi Jawa Tengah berdasarkan data kemiskinan seperti berikut.



Gambar 7. Anggota Cluster

Berdasarkan Gambar 7., hasil menunjukkan bahwa anggota pengelompokan kemiskinan Kabupaten/Kota di Jawa Tengah yang terdiri dari 3 *cluster*, dengan masing-masing *cluster* memiliki ciri khas warna yang berbeda-beda. Pada *cluster* 1 dilambangkan dengan warna hijau, *cluster* 2 dilambangkan dengan warna merah, dan *cluster* 3 dilambangkan dengan warna kuning. Untuk mempermudah dalam membacanya, berikut pemaparan tabel anggota klasterisasi kemiskinan di Provinsi Jawa Tengah.

Tabel 3. Anggota Klasterisasi Kemiskinan di Provinsi Jawa Tengah

Cluster	Anggota Cluster
1	Purbalingga, Magelang, Karanganyar, Kudus, Jepara, Demak, Semarang, Kendal, Batang, Kota Salatiga, Kota Semarang, dan Kota Pekalongan.
2	Cilacap, Banyumas, Banjarnegara, Kebumen, Pemalang, Tegal, dan Brebes.
3	Purworejo, Wonosobo, Boyolali, Klaten, Sukoharjo, Wonogiri, Sragen, Grobogan, Blora, Rembang, Pati, Temanggung, Pekalongan, Kota Magelang, Kota Surakarta, dan Kota Tegal.

Dari Tabel 3. di atas dapat diketahui bahwa pada *cluster* 1 beranggotakan 12 Kabupaten/Kota, *cluster* 2 beranggotakan 7 Kabupaten/Kota, dan *cluster* 3 beranggotakan 16 Kabupaten/Kota.

3.3.3 Profilisasi

Setelah menentukan jumlah *cluster* yang optimal serta masing-masing anggota *cluster*-nya, kemudian dilakukan profilisasi untuk mengetahui karakteristik dari masing-masing *cluster*. Untuk melakukan profilisasi kelompok, dilakukan dengan cara menghitung nilai rata-rata objek terhadap ketiga *cluster* yang telah terbentuk dengan menganalisis variabel yang ada yaitu variabel jumlah penduduk miskin, laju pertumbuhan penduduk, dan tingkat pengangguran terbuka. Ketiga variabel yang digunakan dalam penelitian ini menunjukkan data kemiskinan di Kabupaten/Kota Provinsi Jawa Tengah.

Tabel 4. Nilai rata-rata dari setiap variabel

Variabel	Cluster		
	1	2	3
Jumlah Penduduk Miskin	85.83917	192.70857	90.77625
Laju Pertumbuhan Penduduk	0.8891667	0.2471429	0.2818750
Tingkat Pengangguran Terbuka	5.522500	7.674286	4.193750

Berdasarkan Tabel 4. di atas, hasil profilisasi anggota pada setiap *cluster* untuk kabupaten/kota yang masuk dalam satu *cluster* yang sama, dapat diketahui nilai rata-rata kemiskinan di Kabupaten/Kota Provinsi Jawa Tengah pada setiap *cluster* kemudian diberi kategori berdasarkan nilai rata-rata tersebut. Nilai rata-rata tertinggi ditandai dengan warna gelap diberi kategori tinggi, nilai rata-rata tertinggi kedua ditandai dengan warna terang diberi kategori sedang, dan nilai rata-rata terendah diberi kategori rendah.

Berikut merupakan penjelasan *profiling* masing-masing *cluster* mengenai kemiskinan di Provinsi Jawa Tengah:

- *Cluster 1* : Merupakan kelompok Kabupaten/Kota yang memiliki nilai rata-rata kemiskinan paling rendah. Sehingga kelompok Kabupaten/Kota pada *cluster 1* dapat dikelompokkan menjadi kelompok Kabupaten/Kota dengan tingkat kemiskinan rendah, di mana karakteristiknya yaitu memiliki jumlah penduduk miskin rendah, laju pertumbuhan penduduk tinggi, dan tingkat pengangguran terbuka sedang.
- *Cluster 2* : Merupakan kelompok Kabupaten/Kota yang memiliki nilai rata-rata kemiskinan paling tinggi. Sehingga kelompok Kabupaten/Kota pada *cluster 2* dapat dikelompokkan menjadi kelompok Kabupaten/Kota dengan tingkat kemiskinan tinggi, di mana karakteristiknya yaitu memiliki jumlah penduduk miskin dan tingkat pengangguran terbuka tinggi, sedangkan laju pertumbuhan penduduk rendah.
- *Cluster 3* : Merupakan kelompok Kabupaten/Kota yang memiliki nilai rata-rata kemiskinan sedang. Sehingga kelompok Kabupaten/Kota pada *cluster 3* dapat dikelompokkan menjadi kelompok Kabupaten/Kota dengan tingkat kemiskinan sedang, di mana karakteristiknya yaitu memiliki jumlah penduduk miskin dan laju pertumbuhan penduduk sedang, serta tingkat pengangguran terbuka rendah.

Dari hasil klasterisasi dan profilisasi dengan menggunakan metode *K-Means* peneliti ingin memvisualisasikan hasil *clustering* dan *profiling* tersebut menjadi bentuk peta yang bertujuan agar para pembaca dapat memahami dengan cepat dan ringkas.



Gambar 8. Pemetaan Anggota Setiap *Cluster*

Gambar 8. menunjukkan penyebaran pemetaan untuk mengelompokkan kabupaten/kota yang ada di Provinsi Jawa Tengah berdasarkan data kemiskinan tahun 2022. Untuk kabupaten/kota dengan tingkat kemiskinan tertinggi yang masuk ke dalam *cluster 2* digambarkan dengan warna merah, kabupaten/kota dengan tingkat kemiskinan rendah yang masuk ke dalam *cluster 1* digambarkan dengan

warna hijau, dan kabupaten/kota dengan tingkat kemiskinan sedang yang masuk ke dalam *cluster* 3 digambarkan dengan warna kuning.

Dapat disimpulkan bahwa kabupaten/kota yang masuk ke dalam *cluster* 2 dijadikan sebagai kabupaten/kota yang menjadi prioritas dalam memaksimalkan upaya untuk mengatasi kemiskinan oleh pemerintah Provinsi Jawa Tengah. Dengan mengetahui data ini, pemerintah dapat mengurangi kemiskinan dan dapat meningkatkan kesejahteraan masyarakat.

4. Kesimpulan

Berdasarkan hasil pembahasan dapat disimpulkan bahwa kondisi kemiskinan pada setiap Kabupaten/Kota di Provinsi Jawa Tengah berbeda dan masih terjadi ketimpangan. Masih ada beberapa Kabupaten/Kota di Provinsi Jawa Tengah yang memiliki tingkat kemiskinan tinggi. Dengan menggunakan analisis *cluster* yaitu metode *K-Means Cluster*, mengelompokkan Kabupaten/Kota di Provinsi Jawa Tengah berdasarkan data kemiskinan yaitu data jumlah penduduk miskin, laju pertumbuhan penduduk, dan tingkat pengangguran terbuka yang dikelompokkan menjadi 3 *cluster*. Di mana untuk *cluster* 1 merupakan Kabupaten/Kota dengan kategori kemiskinan rendah yang terdiri dari 12 Kabupaten/Kota dengan karakteristiknya yaitu memiliki jumlah penduduk miskin rendah, laju pertumbuhan penduduk tinggi, dan tingkat pengangguran terbuka sedang. *Cluster* 2 merupakan Kabupaten/Kota dengan kategori kemiskinan tinggi yang terdiri dari 7 Kabupaten/Kota karakteristiknya yaitu memiliki jumlah penduduk miskin dan tingkat pengangguran terbuka tinggi, sedangkan laju pertumbuhan penduduk rendah. Sedangkan *cluster* 3 merupakan Kabupaten/Kota dengan kategori kemiskinan sedang yang terdiri dari 16 Kabupaten/Kota karakteristiknya yaitu memiliki jumlah penduduk miskin dan laju pertumbuhan penduduk sedang, serta tingkat pengangguran terbuka rendah.

Dengan karakteristik yang terdapat pada masing-masing daerah, tentunya diharapkan akan dijadikan sebagai bahan pertimbangan oleh pemerintah untuk memberikan kebijakan yang tepat dan cepat dalam mengatasi kemiskinan di wilayah yang memiliki tingkat kemiskinan tinggi.

Referensi

- [1] S. E. Pujiana, I. P. Sari, V. M. Mardika and M. Putri, "Analisis Algoritma Back Propagation Dalam Prediksi Angka Kemiskinan Di Indonesia," *Pendekar : Jurnal Pendidikan Berkarakter*, vol. III, pp. 11-17, 2020.
- [2] N. P. A. P. Margareni, I. K. Djayasta and I. G. W. M. Yasa, "Faktor-Faktor yang Mempengaruhi Kemiskinan di Provinsi Bali," *PIRAMIDA*, vol. XII, no. 1, pp. 101-110, 2016.
- [3] T. Alawiyah and F. Setiawan, "Pengentasan Kemiskinan Berbasis Kearifan Lokal pada Masyarakat Desa," *Jurnal Sosiologi USK*, vol. V, no. 2, pp. 131-154, 2021.
- [4] S. Royat, Kebijakan Pemerintah dalam Penanggulangan Kemiskinan., Jakarta: Menko Kesra Bidang Koordinasi Pengangulan Kemiskinan, 2015.
- [5] A. Alifah, H. Yozza and Y. Asdi, "Faktor-Faktor Yang Mempengaruhi Kemiskinan Di Kota/Kabupaten Provinsi Sumatera Barat Dengan Menggunakan Analisis Regresi Panel," *Jurnal Matematika UNAND*, pp. 53-61, 2020.
- [6] "Kemiskinan 2020-2022," Badan Pusat Statistik Jawa Tengah, [Online]. Available: <https://jateng.bps.go.id/indicator/23/34/1/kemiskinan.html>.
- [7] T. Christianto, "Determinan dan Karakteristik Kemiskinan di Provinsi Riau," *ISSN : 1978 – 3612*, vol. VII, no. 2, 2017.
- [8] W. A. Saputra, Analisis Pengaruh Jumlah Penduduk, PDRB, IPM, Pengangguran terhadap Tingkat Kemiskinan di Kabupaten / Kota Jawa Tengah, Semarang: Universitas Diponegoro , 2011.
- [9] H. Siregar and D. Wahyuniarti, "Dampak Pertumbuhan Ekonomi Terhadap Penurunan Jumlah Penduduk Miskin," *Prosiding Seminar Nasional Meningkatkan Peran Sektor Pertanian dan Penanggulangan Kemiskinan*, pp. 23-34, 2008.

- [10] N. V. Suryani and A. K. Putri, "Trend Laju Pertumbuhan Penduduk Dan Kemiskinan Di Provinsi Kepulauan Bangka Belitung," *Holistic Journal of Management Research*, vol. I, no. 1, pp. 27-38, 2020.
- [11] A. B. M. Bintang and N. Woyanti, "Pengaruh PDRB, Pendidikan, Kesehatan, Dan Pengangguran Terhadap Tingkat Kemiskinan Di Jawa Tengah (2011-2015)," *Media Ekonomi dan Manajemen*, vol. 33, no. 1, pp. 20-28, 2018.
- [12] P. O'Campo, A. Molnar, E. Ng, E. Renahy, C. Mitchell, K. Shankardass, A. S. John, C. Bambra and C. Muntane, "Social welfare matters: a realist review of when, how, and why unemployment insurance impacts poverty and health," *Social Science & Medicine*, pp. 88-94, 2015.
- [13] T. R. Akbar, Pengaruh Jumlah Penduduk, Tingkat Pengangguran, dan Tingkat Pendidikan Terhadap Kemiskinan, Jawa Timur: Universitas Pembangunan Nasional Veteran Jawa Timur, 2013.
- [14] R. Hidayat, Analisis Klaster dengan Metode Hierarchical Clustering dan NonHierarchical Clustering pada Pengelompokan Kabupaten/Kota di Jawa Tengah Berdasarkan Faktor-Faktor yang Mempengaruhi Kemiskinan, Semarang: Perpus UNIMUS, 2017.
- [15] R. Hidayat, R. Wasono and M. Y. Darsyah, "Pengelompokan Kabupaten/Kota Di Jawa Tengah Menggunakan Metode K-Means Dan Fuzzy C-Means," *Seminar Nasional Pendidikan, Sains dan Teknologi*, pp. 240-250, 2017.
- [16] A. Alifah, "Pemetaan Kemiskinan Kabupaten/Kota Di Provinsi Jawa Tengah Menggunakan Algoritma K-Means," *Seminar Nasional Edusainstek FMIPA UNIMUS 2019*, pp. 553-557, 2019.
- [17] Y. Agusta, "K-means - Penerapan, Permasalahan dan Metode Terkait," *Jurnal Sistem dan Informatika*, vol. III, no. 3, pp. 47-60, 2007.
- [18] R. E. Walpole, Introduction to Statistics, New York : Macmillan ; London : Collier Macmillan, 1982.
- [19] J. F. H. JR., W. C. Black, B. J. Babin and R. E. Anderson, Multivariate Data Analysis Sixth Edition, New Jersey: Pearson Education, Inc, 2006.
- [20] L. Rahmawati, A. and T. E. Lestari, Analisis Kelompok Dengan Menggunakan Metode Hierarki Untuk Pengelompokan Kabupaten/Kota Di Jawa Timur Berdasarkan Indikator Kesehatan, Jawa Barat: Universitas Negeri Malang, 2013.
- [21] T. S. Madhulatha, "An Overview On Clustering Methods," *IOSR Journal of Engineering*, pp. 719-725, 2012.
- [22] I. Y. Purwananto and R. Soelaiman, "Optimasi Kinerja Algoritma Klasterisasi K-Means Untuk Kuantisasi Warna Citra," *Jurnal Teknik ITS*, vol. I, no. 1, pp. 197-202, 2012.
- [23] E. Prasetyo, Data Mining Konsep dan Aplikasi Menggunakan Matlab, Yogyakarta: Andi Offset, 2012.
- [24] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. V, no. 1, pp. 281-297, 1967.
- [25] Ediyanto, M. N. Mara and N. Satyahadewi, "Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis," *Buletin Ilmiah Mat. Stat dan Terapannya (Bimaster)*, vol. II, no. 2, pp. 133-136, 2013.
- [26] J. Supranto, Analisis Multivariat Arti & Interpretasi, Jakarta: PT RINEKA CIPTA, 2004.