

PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI JAWA TENGAH BERDASARKAN INDIKATOR KESEJAHTERAAN DENGAN METODE *K- MEANS CLUSTER*

Wardono¹⁾, Sunarmi²⁾, Mutia Reflianti Wirawan³⁾

¹Dosen, Jurusan Matematika, Universitas Negeri Semarang, Indonesia
email: wardono@mail.unnes.ac.id

²Dosen, Jurusan Matematika, Universitas Negeri Semarang, Indonesia
email: sunarmi.mat@gmail.com

³Mahasiswa, Jurusan Matematika, Universitas Negeri Semarang, Indonesia
email: mutiareflianti08@gmail.com

Abstract

Welfare is the thing that determines a development in an area. The government always tries to implement people's welfare as a program goal to be achieved. For this reason, this research will discuss the grouping of regencies / cities based on indicators of people's welfare using the K-Means cluster analysis. The aim of this research is to apply the K-Means cluster method to classify districts or cities based on the indicators of people's welfare in Central Java Province in 2017. The data in this study use secondary data from BPS publications in Central Java Province. The data used in this study are population density data (X1), open unemployment (X2), adjusted per capita expenditure (X3), life expectancy (X4), average length of schooling (X5), and the percentage of poor population (X6). From the results of the K-Means cluster analysis, the regencies / cities in Central Java are divided into 3 groups. Cluster 1 is a Regency / City group whose welfare is low (low). Cluster 2 is a regency / city group whose welfare is high (good). Cluster 3 is a Regency / City group whose welfare is quite good (moderate).

Keywords: Kesejahteraan Rakyat, K-Means, Analisis Cluster.

1. PENDAHULUAN

Kesejahteraan rakyat selalu menjadi topik yang menarik untuk dibahas. Sebagaimana di setiap negara yang menjadi tujuan utama dalam pembangunan adalah peningkatan kesejahteraan rakyat. Seperti halnya di Indonesia, kesejahteraan rakyat adalah salah satu tujuan negara yang tertuang dalam Pembukaan UUD 1945 alinea IV [1]. Kesejahteraan rakyat pada suatu wilayah merupakan suatu manifestasi yang diraih oleh masyarakat tersebut yang diperoleh dari berbagai upaya, termasuk upaya dan kegiatan aktifitas ekonomi masyarakat tersebut. Untuk mencapai kondisi ideal ini diperlukan perencanaan dan aktivitas selain oleh rakyat, peran pemerintah juga sangat dominan, baik menyangkut perencanaan, kegiatan dan bantuan biaya berupa permodalan [2]. Secara umum kesejahteraan dapat diartikan sebagai suatu keadaan dimana segenap warga negara selalu berada dalam kondisi serba kecukupan segala kebutuhannya, baik material maupun spiritual [3]. Indikator kesejahteraan rakyat menyajikan data tentang tingkat perkembangan kesejahteraan rakyat dari waktu ke waktu. Istilah kesejahteraan mencakup berbagai aspek kehidupan yang sangat luas dan tidak semuanya dapat diukur. Adanya indikator kesejahteraan rakyat diharapkan dapat digunakan sebagai dasar perencanaan maupun evaluasi terhadap upaya peningkatan kualitas hidup masyarakat. Taraf yang dicakup antara lain meliputi aspek kependudukan, kesehatan, pendidikan, ketenagakerjaan, taraf dan pola konsumsi, perumahan, kemiskinan, serta sosial lainnya.

Jawa Tengah merupakan salah satu provinsi besar yang berada di Pulau Jawa. Jumlah penduduk Jawa Tengah pada tahun 2017 berdasarkan Proyeksi Penduduk Indonesia 2010-

2020 tercatat sebanyak 34.257.865 jiwa. Memperhatikan laju pertumbuhan penduduk per tahun periode 2013-2017 maka Pemerintah Provinsi Jawa Tengah patut berbangga karena data laju pertumbuhan penduduk pertahun mengalami penurunan. Ketika jumlah penduduk terus bertambah, berarti pemerintah juga harus terus menambah jumlah fasilitas hidup layak bagi masyarakatnya. Selain masalah tentang kependudukan, permasalahan di bidang ketenagakerjaan harus terus menjadi perhatian pemerintah agar dapat cepat diselesaikan. Permasalahan tersebut diantaranya tingginya tingkat pengangguran. Perkembangan tingkat pengangguran terbuka di Jawa Tengah selama periode 2015-2017 terus menurun dari 4,99 persen turun menjadi 4,63 persen dan turun lagi menjadi 4,57 persen [4]. Namun meskipun pengangguran terus menurun, bukan berarti persoalan ketenagakerjaan sudah teratasi. Masih harus diteliti lebih lanjut, mereka yang sudah bekerja apakah sudah memiliki produktivitas yang tinggi yang salah satunya dapat dilihat dari jam kerja yang digunakan. Salah satu prasyarat keberhasilan program-program pembangunan sangat tergantung pada ketepatan pengidentifikasian target grup dan target area.

Untuk menganalisis pengidentifikasian diatas diperlukan suatu teknik analisis multivariat. Analisis multivariat merupakan salah satu solusi dalam hal mengelompokkan objek-objek yang mempunyai kesamaan karakteristi [1]. Salah satu analisis multivariat yang dapat digunakan untuk memahami dan mempermudah menginterpretasi data adalah analisis *cluster*.

Analisis *Cluster* merupakan teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Analisis Cluster mengklasifikasi objek sehingga objek-objek yang paling dekat kesamaanya dengan objek lain berada dalam *cluster* yang sama [5]. Ciri-ciri pengelompokan yang baik adalah mempunyai kesamaan (homogenitas) yang tinggi antar anggota dalam satu *cluster* (*within cluster*), dan mempunyai ketaksamaan (heterogenitas) yang tinggi antar *cluster* (*between cluster*) [6].

Dalam penelitian ini akan dianalisis kasus pengelompokan *K-Means cluster* Kabupaten atau Kota berdasarkan indikator kesejahteraan rakyat. *K-Means Cluster* adalah salah satu metode dan *clustering* non-hierarki yang berusaha mengelompokkan data ke dalam suatu *cluster* sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu *cluster* yang sama [5].

Untuk mengetahui tingkat kesejahteraan rakyat pada suatu daerah, ada beberapa aspek yang bisa diukur seperti kepadatan penduduk, pengangguran terbuka, pengeluaran perkapita disesuaikan, angka harapan hidup, rata-rata lama sekolah, dan jumlah presentase penduduk miskin. Berdasarkan indikator kesejahteraan rakyat akan dilakukan *clustering* pada Kabupaten atau Kota di Jawa Tengah tahun 2017 dengan metode *K-Means Cluster*.

2. KAJIAN LITERATUR

2.1 Indikator Kesejahteraan Rakyat

Menurut Undang-undang No 11 Tahun 2009, tentang Kesejahteraan Masyarakat, kesejahteraan masyarakat adalah kondisi terpenuhinya kebutuhan material, spiritual, dan sosial warga negara agar dapat hidup layak dan mampu mengembangkan diri, sehingga dapat melaksanakan fungsi sosialnya. Dari Undang-Undang di atas dapat kita cermati bahwa ukuran tingkat kesejahteraan dapat dinilai dari kemampuan seorang individu atau kelompok dalam usaha nya memenuhi kebutuhan material dan spiritualnya. Tingkat kesejahteraan rakyat pada suatu daerah bisa dilihat dari berbagai aspek, seperti kepadatan penduduk, pengangguran terbuka, pengeluaran perkapita disesuaikan, angka harapan hidup, rata-rata lama sekolah, dan jumlah presentase penduduk miskin. Indikator-indikator tersebut ada yang menjelaskan tingkat kesejahteraan masyarakat berdasarkan tingkat kepuasannya. Salah satu indikator kesejahteraan masyarakat berdasarkan tingkat kepuasan adalah keterbebasan dari kemiskinan. Artinya kesejahteraan masyarakat dapat tercapai karena kebutuhan dasar terpenuhi yang didukung oleh lingkungannya agar keluar dari kerentanan tersebut [7].

2.1.1 Kepadatan Penduduk

Kepadatan penduduk adalah perbandingan antara jumlah penduduk dengan luas wilayah yang dihuni [8]. Analisis kepadatan penduduk penting untuk mengetahui persebaran penduduk suatu wilayah dan penataan ruang khususnya distribusi permukiman [9]. Pada daerah dengan kepadatan yang tinggi, usaha peningkatan kualitas penduduk akan lebih sulit dilakukan. Hal ini menimbulkan permasalahan sosial ekonomi, kesejahteraan, dll [10].

2.1.2 Pengangguran Terbuka

Pengangguran terbuka didefinisikan sebagai orang yang sedang mencari pekerjaan atau yang sedang mempersiapkan usaha atau juga yang tidak mencari pekerjaan karena merasa tidak mungkin lagi mendapatkan pekerjaan, termasuk juga mereka yang baru mendapatkan pekerjaan tetapi belum mulai bekerja[4].

2.1.3 Pengeluaran Perkapita disesuaikan

Pengeluaran per kapita yang disesuaikan ditentukan dari nilai pengeluaran per kapita dan paritas daya beli (*Purchasing Power Parity-PPP*). Kemampuan daya beli masyarakat terhadap sejumlah kebutuhan pokok dilihat dari besarnya rata-rata besarnya pengeluaran perkapita sebagai pendekatan yang mewakili pencapaian pembangunan untuk hidup layak. Tingkat kesejahteraan dapat dikatakan meningkat apabila terjadi peningkatan konsumsi riil perkapita, yaitu peningkatan nominal pengeluaran rumah tangga lebih tinggi dari tingkat inflasi pada periode yang sama[11].

2.1.4 Angka Harapan Hidup

Angka Harapan Hidup (AHH) merupakan salah satu indikator yang digunakan untuk menilai derajat kesehatan penduduk. Angka Harapan Hidup di suatu wilayah berbeda dengan wilayah lainnya tergantung dari kualitas hidup yang mampu dicapai oleh penduduk[12]. Tinggi rendahnya angka harapan hidup dapat menjadi pertimbangan dalam menggambarkan kemajuan sosial ekonomi masyarakat[13]

2.1.5 Rata-Rata Lama Sekolah

Rata-rata lama sekolah mencerminkan tingkat pendidikan yang dicapai oleh masyarakat suatu daerah. Semakin tinggi angka rata-rata lama sekolah berarti semakin tinggi jenjang pendidikan yang dijalani[11]. Tobing berpendapat bahwa seseorang yang pendidikannya tinggi yang di ukur dengan lamanya waktu bersekolah akan memiliki pekerjaan dan upah yang lebih baik dibandingkan dengan orang yang pendidikannya lebih rendah[14].

2.1.6 Jumlah Presentase Penduduk Miskin

Kemiskinan adalah suatu kondisi ketidakmampuan secara ekonomi untuk memenuhi standar hidup rata-rata masyarakat di suatu daerah. Kemiskinan di suatu wilayah dapat dijelaskan melalui beberapa indikator antara lain jumlah penduduk miskin, persentase penduduk miskin, garis kemiskinan, indeks kedalaman kemiskinan, indeks keparahan kemiskinan, serta karakteristik penduduk miskin itu sendiri berdasarkan fasilitas tempat tinggal, tingkat pendidikan, status pekerjaan, status kesehatan dan lainnya [4].

2.2 Analisis Statistika

Berdasarkan jumlah variabel yang dianalisis, analisis statistika dapat dikelompokkan menjadi analisis univariat, bivariat, dan multivariat. Analisis univariat berasal dari kata uni dan variate yang berarti analisis satu variabel. Analisis bivariat berasal dari kata bi dan variate berarti analisis dua variabel. Apabila analisis lebih dari dua variabel dinamakan analisis multivariat [15].

2.2 Statistika Multivariat

Secara umum, analisis multivariat atau metode multivariat berhubungan dengan metode-metode statistik yang secara bersama-sama (simultan) melakukan analisis terhadap lebih dari dua variabel pada setiap obyek atau orang . Beberapa ahli lain mengatakan bahwa tujuan analisis multivariat adalah mengukur, menerangkan, dan memprediksi tingkat relasi diantara variat-variati [16].

2.1.1 Jenis Data dalam Analisis Multivariat

Menurut Sukestiyarno, berdasarkan jenis perolehannya atau pengumpulannya data diklasifikasikan menjadi 5 macam tipe atau skala. Hal tersebut adalah: skala nominal, skala ordinal, skala kardinal, skala interval, dan skala ratio [17].

2.3.2 Klasifikasi Metode Statistik Multivariat

Dalam analisis multivariat terdapat dua jenis atau klasifikasi, yaitu metode dependensi dan metode interdependensi yang masing-masing metode memiliki beberapa tipe atau analisis. Berikut penjelasannya.

1. Metode Dependensi / Ketergantungan (*Dependence Method*)

Pengelompokan jenis analisis dengan metode dependensi dapat didasarkan oleh dua hal, yaitu variabel terikat (*dependent variabel*) dan tipe data pada variabel bebas dan terikat [17]. Apabila variabel terikat berjumlah satu, terdapat tiga analisis multivariat yaitu: regresi berganda, analisis konjoin, dan analisis diskriminan. Untuk variabel terikat yang berjumlah dua atau lebih, terdapat dua analisis multivariat yaitu : analisis korelasi kanonikal dan analisis varian multivariat [19].

2. Metode Interdependensi / Tidak saling Ketergantungan (*Independence Method*)

Apabila dalam suatu analisis kesulitan untuk memisahkan antara variabel bebas dengan variabel terikat, maka metode ini disebut dengan metode interdependensi, karena semua variabel dianggap independen. Analisis interdependensi bertujuan untuk memberikan arti (*meaning*) kepada suatu set variabel(kelompok variabel) atau mengelompokkan suatu set variabel menjadi kelompok yang lebih sedikit jumlahnya dan masing-masing kelompok membentuk variabel baru yang disebut faktor. Beberapa contoh analisis interdependensi antara lain analisis faktor, analisis *cluster*, penskalaan multidimensional [19].

2.3.3 Uji Data

Uji Data pada prinsipnya bertujuan untuk memastikan bahwa berbagai metode multivariat (*cluster analysis*, *factor analysis* dan lainnya) bisa digunakan pada data tertentu [20]. Uji yang biasanya dilakukan pada berbagai metode statistik multivariat seperti missing data, data outlier, uji normalitas data, dan uji homoskedastisitas data.

2.4 Analisis Klaster

Analisis *Cluster* adalah suatu teknik yang baru dikembangkan untuk menyelesaikan masalah analisis data. Teknik ini, akan mencari kategori-kategori atau pola sample data (data sets) berdasarkan proses pembentukan grup-grup data homogen yang disebut *cluster* [21]. Secara logika, *cluster* yang baik adalah *cluster* yang terbentuk memiliki:

1. Homogenitas (kesamaan) internal yang tinggi antar anggota dalam satu *cluster* (*Within-Cluster*).
2. Heterogenitas (perbedaan) eksternal yang tinggi antar anggota dalam satu *cluster* (*Between-Cluster*).

Dari dua hal diatas dapat disimpulkan bahwa sebuah *cluster* yang baik adalah *cluster* yang mempunyai anggota-anggota yang semirip mungkin antara satu dengan yang lain, namun sangat tidak mirip dengan anggota-anggota *cluster* yang lain. Disini, “mirip” diartikan sebagai tingkat kesamaan karakteristik antara dua data. Lebih besar kemiripannya dalam kelompok dan lebih besar perbedaannya diantara kelompok lain [22].

2.4.3 Proses Analisis Cluster

Dibawah ini merupakan tahapan proses analisis *cluster*:

2.4.3.1 Merumuskan Masalah

Hal yang paling penting di dalam perumusan masalah analisis klaster adalah pemilihan variabel-variabel yang akan dipergunakan untuk pengklasteran (pembentukan klaster) [19].

2.4.3.2 Proses Standardisasi

Proses standarisasi dilakukan apabila diantara variabel-variabel yang diteliti terdapat perbedaan ukuran satuan yang besar.

2.4.3.3 Mengukur Kesamaan Antar Objek (*Similarity*)

Oleh karena tujuan pengklasteran ialah untuk mengelompokkan objek yang mirip dalam klaster yang sama, maka beberapa ukuran diperlukan untuk mengakses seberapa mirip atau berbeda objek-objek tersebut. Berikut ada tiga metode yang dapat diterapkan, yaitu ukuran korelasi, ukuran jarak, dan ukuran asosiasi [19].

1. Ukuran Korelasi

Ukuran ini dapat diterapkan pada data dengan skala metrik, namun jarang digunakan karena titik beratnya pada nilai suatu pola tertentu, padahal fisik berat analisis klaster adalah besarnya objek.

2. Ukuran Berbentuk Jarak

Ukuran berbentuk jarak merupakan ukuran yang sering digunakan. Metode yang dapat digunakan untuk mengukur kesamaan adalah dengan mengukur jarak (*distance*) antar dua objek [23]. Beberapa *distance space* telah diimplementasikan dalam menghitung jarak (*distance*) antara data dan *centroid* termasuk diantaranya *Manhattan/City Block*, *Euclidean*, dan *Minkowski*. *Minkowski distance space* [24].

3. Ukuran Berbentuk Asosiasi

Pada dasarnya, cara ini akan mengasosiasikan sebuah data dengan klaster tertentu; dalam praktek, cara ini tidak sepopuler kedua cara sebelumnya [20].

2.4.3.4 Melakukan *Clustering*

Setelah data yang dianggap mempunyai satuan yang sangat berbeda diseragamkan, dan metode *cluster* ditentukan (misal dipilih *Euclidean*), langkah selanjutnya adalah membuat *cluster* atau bisa disebut pengelompokan data [21].

2.4.3.5 Melakukan Interpretasi Terhadap *Cluster* Yang Telah Terbentuk

Setelah jumlah *cluster* terbentuk dengan metode hierarki atau non-hierarki, langkah selanjutnya adalah melakukan interpretasi terhadap *cluster* yang telah terbentuk [21].

2.4.3.6 Melakukan Validasi dan *Profiling Cluster*

Proses validasi bertujuan menjamin bahwa solusi yang dihasilkan dari *Cluster Analysis* dapat mewakili populasi dan dapat digeneralisasi untuk objek lain. Kemudian dilakukan proses *profiling* untuk menjelaskan karakteristik setiap *cluster* berdasar profil tertentu [21].

2.4.4 Metode Analisis *Cluster*

Dalam analisis *cluster*, terdapat ada dua metode untuk mengelompokkan observasi ke dalam *cluster*, yaitu:

2.4.4.1 Metode Hierarki

Metode hierarki (*hierarchical method*) yaitu metode yang memulai pengelompokannya dengan dua atau lebih objek yang mempunyai kesamaan paling dekat, kemudian proses dilanjutkan ke obyek lain yang mempunyai kedekatan kedua. Tahap-tahap pengelompokan data dengan menggunakan metode hierarki adalah

1. Tentukan k sebagai jumlah *cluster* yang ingin dibentuk.
2. Setiap data obyek dianggap sebagai *cluster* sehingga $n = N$.
3. Menghitung jarak antar *cluster*.
4. Mencari dua *cluster* yang mempunyai jarak antar *cluster* paling minimal dan menggabungkannya (berarti $N = n-1$).
5. Jika $n > k$, maka kembali ke langkah 3. [26]

2.4.4.2 Metode Non-Hierarki

Berbeda dengan metode hierarki, metode ini justru dimulai dengan menentukan terlebih dahulu jumlah *cluster* yang diinginkan (dua *cluster*, tiga *cluster*, atau yang lain). Setelah jumlah *cluster* diketahui, baru proses *cluster* dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut *K-Means Cluster* [27].

2.5 K-Means Cluster

K-Means clustering merupakan metode non-hierarkial yang digunakan untuk mengelompokkan sebuah data ke dalam dua atau lebih variasi atau kelompok. Fungsinya

untuk meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok [28]. Metode *K-Means Cluster* ini bersifat *unsupervised* (tanpa arahan) karena data yang dianalisis tidak mempunyai label kelas yang berarti dalam proses pengelompokannya.

2.6 Software SPSS

Menurut Priyanto, SPSS adalah program atau *software* yang digunakan untuk mengolah data statistik. Dari berbagai program olah data statistik lainnya, SPSS merupakan program yang paling banyak digunakan. Umumnya, SPSS dulu digunakan untuk mengolah data statistik pada ilmu sosial saja. Oleh karena itu, kepanjangan SPSS adalah *Statistical Package for the Social Science* [29].

3. METODE PENELITIAN

Pada penelitian untuk mengelompokkan kabupaten/kota berdasarkan indikator kesejahteraan rakyat dengan menggunakan data sekunder dari publikasi BPS Jawa Tengah tahun 2017. Dalam penelitian ini ada sebanyak 35 Kabupaten/Kota di Provinsi Jawa Tengah.

Variabel yang digunakan pada penelitian ini adalah kepadatan penduduk (X_1), pengangguran terbuka (X_2), pengeluaran perkapita disesuaikan (X_3), angka harapan hidup (X_4), rata-rata lama sekolah (X_5), dan jumlah presentase penduduk miskin (X_6).

3.1 Metode Pengumpulan data

Metode yang digunakan pada penelitian ini yaitu metode dokumentasi dan metode literatur. Metode dokumentasi yaitu metode pengumpulan data dengan cara mengambil data-data tertulis yang berhubungan dengan tugas akhir. Metode ini digunakan untuk mendapatkan data tentang indikator kesejahteraan rakyat yang diperoleh dari Badan Pusat Statistika Jawa Tengah.

Metode literatur adalah mengumpulkan, memilih, dan menganalisis beberapa sumber bacaan yang berkaitan dengan rumusan masalah dalam penyusunan tugas akhir. Dengan metode ini, penulis mengumpulkan dan memilih sumber bacaan (buku-buku) yang berkaitan dengan penyusunan laporan tugas akhir ini.

3.2 Analisis Data

Dalam penelitian analisis *K-Means Cluster* ini dilakukan tahap-tahap untuk menganalisis data yang diperoleh, antara lain:

1. Uji asumsi yaitu uji multikolinearitas.
2. Melakukan standarisasi data.
3. Menentukan banyaknya kluster yang akan dibentuk, dalam penelitian ini kluster yang akan dibentuk adalah 3 kluster.
4. Menentukan pusat kluster awal secara acak.
5. Menghitung jarak antar setiap objek dengan pusat kluster.
6. Masukkan tiap objek ke satu kluster yang memiliki jarak terdekat dengan pusat klusternya.

4. HASIL PENELITIAN

Analisis kluster akan mengelompokkan kabupaten atau kota berdasarkan tingkat kesejahteraan rakyat di Jawa Tengah tahun 2017. Setelah memperoleh semua data dari Badan Pusat Statistika, maka data tersebut akan dianalisis menggunakan *software* SPSS 20. Hasil analisis yang diperoleh adalah sebagai berikut:

1. Uji Multikolinearitas

Uji multikolinearitas diperlukan untuk mengetahui ada tidaknya variabel independen yang memiliki kemiripan antar variabel independen dalam suatu model [30].

Cara mengetahui adanya gejala multikolinieritas adalah sebagai berikut:

1. Melihat nilai *tolerance*: terjadi multikolinieritas, jika nilai *tolerance* lebih besar dari 0,1.
2. Melihat nilai VIF (*Variance Inflation Factor*): tidak terjadi multikolinieritas, jika nilai VIF kurang dari 10.

Tabel 4. 1 Hasil Uji Multikolinieritas

Coefficients ^a			
Model	Collinearity Statistics		
	Tolerance	VIF	
1	X1	,331	3,018
	X2	,652	1,533
	X3	,312	3,207
	X4	,285	3,505
	X5	,150	6,656
	X6	,491	2,037

Tabel 4.1 adalah hasil pengujian multikolinieritas. Dari 7 variabel yang diteliti, semua variabel memiliki nilai *tolerance* lebih besar dari 0,1 serta nilai VIF nya kurang dari 10. Maka dapat disimpulkan bahwa semua variabel tidak terjadi multikolinieritas sehingga dapat diuji lanjut dengan analisis *K-Means Cluster*.

2. Analisis Deskriptif

Tabel 4. 2 Deskriptif

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
X1	35	479,00	11722,00	2043,6000	2454,49500
X2	35	4155,00	72051,00	23541,0857	15856,77597
X3	35	7785,00	14921,00	10414,0571	1700,62556
X4	35	68,61	77,49	74,6314	1,95120
X5	35	618,00	1050,00	758,2000	121,55797
X6	35	4,62	20,32	12,4914	4,13079
Valid N (listwise)	35				

Pada tabel diatas terdapat 6 variabel yang diamati atau dianalisis dengan informasi banyaknya data atau N, minimum, maksimum, mean atau rata-rata, dan standar deviasi. Untuk variabel kepadatan penduduk (X_1) data berjumlah 35 dengan angka minimum 479 , angka maksimum 11.722, nilai rata-rata 2.043,6 dan standar deviasi 2.454,495. Untuk variabel pengangguran terbuka (X_2) data berjumlah 35 dengan angka minimum 4.155, angka maksimum 72.051, nilai rata-rata 23.541,0857 dan standar deviasi 15.856,77597. Untuk variabel pengeluaran perkapita disesuaikan (X_3) data berjumlah 35 dengan angka minimum 7.785 , angka maksimum 14.921, nilai rata-rata 10.414,0571 dan standar deviasi 1.700,62556. Untuk variabel angka harapan hidup (X_4) data berjumlah 35 dengan angka minimum 68,61 , angka maksimum 77,49, nilai rata-rata 74,6314 dan standar deviasi 1,95120. Untuk variabel rata-rata lama sekolah (X_5) data berjumlah 35 dengan angka minimum 618 , angka maksimum 1.050, nilai rata-rata 758,2 dan standar deviasi 121,55797. Untuk variabel jumlah presentase penduduk miskin (X_6) data berjumlah 35 dengan angka minimum 4,62 , angka maksimum 20,32, nilai rata-rata 12,4914 dan standar deviasi 4,13079.

3. Analisis *K-Means Cluster*

Adapun proses hasil akhir klastering data setelah proses awal klaster dan dilakukan iterasi adalah sebagai berikut:

Tabel 4. 3 Final Cluster Centers

	Final Cluster Centers		
	Cluster		
	1	2	3
Zscore(X1)	-,35551	1,92452	-,40310
Zscore(X2)	2,21608	-,40879	-,16137
Zscore(X3)	-,52063	1,60379	-,31003
Zscore(X4)	-,186284	,72959	,04658
Zscore(X5)	-,85172	1,73964	-,30318
Zscore(X6)	,44428	-,122166	,23066

Tabel 4.5 adalah hasil akhir dari proses *clustering*, dengan analisis sebagai berikut.

Arti angka:

Angka di atas terkait dengan proses standardisasi data sebelumnya, yang mengacu pada angka z, dengan ketentuan:

1. Angka negatif berarti data di bawah rata-rata total.
2. Angka positif berarti data di atas rata-rata total.

Berikut merupakan karakteristik yang terbentuk pada tiap *cluster*:

1. *Cluster 1*

Cluster 1 berisi kepadatan penduduk di bawah rata-rata, jumlah pengangguran terbuka di atas rata-rata, pengeluaran perkapita disesuaikan di bawah rata-rata, angka harapan hidup di bawah rata-rata, rata-rata lama sekolah di bawah rata-rata, dan jumlah presentase penduduk miskin di atas rata-rata.

2. *Cluster 2*

Karakteristik *cluster* yang terbentuk pada *cluster 2* berisi kepadatan penduduk di atas rata-rata, jumlah pengangguran terbuka di bawah rata-rata, pengeluaran perkapita disesuaikan di atas rata-rata, angka harapan hidup di atas rata-rata, rata-rata lama sekolah di atas rata-rata, dan jumlah presentase penduduk miskin di bawah rata-rata.

3. *Cluster 3*

Karakteristik *cluster* yang terbentuk pada *cluster 3* berisi kepadatan penduduk di bawah rata-rata, jumlah pengangguran terbuka di bawah rata-rata, pengeluaran perkapita disesuaikan dibawah rata-rata, angka harapan hidup di atas rata-rata, rata-rata lama sekolah di bawah rata-rata, dan jumlah presentase penduduk miskin di atas rata-rata.

4. Perbedaan Variabel pada *Cluster* yang Terbentuk

Setelah mendapatkan hasil akhir *cluster*, langkah berikutnya adalah melihat apakah variabel-variabel yang telah membentuk *cluster* tersebut mempunyai perbedaan pada tiap *cluster*. Hal itu dilakukan dengan melihat tabel ANOVA seperti dibawah ini.

Tabel 4. 4 ANOVA

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Zscore(X1)	13,413	2	,224	32	59,836	,000
Zscore(X2)	8,206	2	,550	32	14,931	,000
Zscore(X3)	9,373	2	,477	32	19,661	,000
Zscore(X4)	6,830	2	,636	32	10,746	,000
Zscore(X5)	11,362	2	,352	32	32,245	,000
Zscore(X6)	5,465	2	,721	32	7,581	,002

Berdasarkan tabel 4.4 diperoleh hasil bahwa variabel kepadatan penduduk adalah variabel yang paling membedakan anggota dari ketiga *cluster* karena mempunyai nilai F terbesar diantara variabel yang lain yaitu 59,836 dan nilai signifikan 0,00.

5. Jumlah Anggota di Setiap *Cluster*

Setelah melihat variabel apa saja yang memberikan perbedaan pada *cluster* yang terbentuk, akan didapatkan jumlah anggota di setiap *cluster*.

Tabel 4. 5 Number of Cases in each Cluster

Number of Cases in each Cluster	
Cluster	
1	3,000
2	6,000
3	26,000
Valid	35,000
Missing	,000

Dari tabel 4.7 terlihat bahwa responden terbanyak ada di *cluster 3* dan tidak ada variabel yang hilang (*missing*). Jumlah anggota pada *cluster 1* sebanyak 3 kabupaten/kota, jumlah

anggota pada *cluster* 2 sebanyak 6 kabupaten/kota, jumlah anggota pada *cluster* 3 sebanyak 26 kabupaten/kota.

6. Komposisi *Cluster*

Untuk mengetahui Kabupaten/Kota mana sajakah yang termasuk ke dalam *cluster* 1, *cluster* 2, dan *cluster* 3 dapat dilihat pada QCL_1. Sedangkan untuk mengetahui jarak pada setiap *cluster* dapat dilihat pada QCL_2. Dari hasil *software* SPSS dihasilkan QCL_1 dan QCL_2 seperti di bawah ini.

Tabel 4. 6 Komposisi *Cluster*

No	Kabupaten/Kota	QCL_1	QCL_2
(1)	(2)	(3)	(4)
1	Kabupaten Cilacap	1	1,27208
2	Kabupaten Banyumas	3	1,64183
3	Kabupaten Purbalingga	3	1,68381
4	Kabupaten Banjarnegara	3	1,49374
5	Kabupaten Kebumen	3	2,08024

Penafsiran *cluster* berdasarkan nilai QCL_1 dan QCL_2 sebagai berikut:

1. Kabupaten Cilacap merupakan anggota kelompok *cluster* 1 dengan posisi nilai *cluster* berjarak 1,27208 dari pusat *cluster* 1.
2. Kabupaten Banyumas merupakan anggota kelompok *cluster* 3 dengan posisi nilai *cluster* berjarak 1,64183 dari pusat *cluster* 3.
3. Kabupaten Purbalingga merupakan anggota kelompok *cluster* 3 dengan posisi nilai *cluster* berjarak 1,68381 dari pusat *cluster* 3.
4. Kabupaten Banjarnegara merupakan anggota kelompok *cluster* 3 dengan posisi nilai *cluster* berjarak 1,49374 dari pusat *cluster* 3.
5. Kabupaten Kebumen merupakan anggota kelompok *cluster* 3 dengan posisi nilai *cluster* berjarak 2,08024 dari pusat *cluster* 3.

Catatan: tidak semua data ditampilkan

4.2 Pembahasan

Berdasarkan hasil analisis *K-Means Cluster* diatas, untuk menentukan mana kelompok *cluster* yang memiliki tingkat kesejahteraan tinggi, sedang, ataupun rendah dapat dilihat melalui karakteristik *cluster*. Pada *cluster* 1 dapat dilihat jumlah pengangguran terbuka dan jumlah presentase penduduk miskin memiliki nilai rata-rata yang lebih tinggi dibanding *cluster* 2 dan *cluster* 3. Namun pengeluaran perkapita disesuaikan, angka harapan hidup, dan rata-rata lama sekolah memiliki nilai rata-rata yang lebih rendah dibanding *cluster* 2 dan *cluster* 3. Karena jumlah pengangguran terbuka dan jumlah presentase penduduk miskin yang tinggi serta pengeluaran perkapita disesuaikan, angka harapan hidup, dan rata-rata lama sekolah yang rendah maka *cluster* 1 termasuk dalam kelompok kabupaten/kota dengan kesejahteraan rakyat yang masih kurang (rendah). Yang termasuk kelompok kabupaten/kota dengan kesejahteraan rakyat yang masih rendah adalah Kabupaten Cilacap, Kabupaten Tegal, dan Kabupaten Brebes.

Pada *cluster* 2 dapat dilihat bahwa jumlah kepadatan penduduk, pengeluaran pekapita disesuaikan, angka harapan hidup, dan rata-rata lama sekolah memiliki nilai rata-rata yang lebih tinggi dibanding dengan *cluster* lainnya. Sedangkan jumlah pengangguran terbuka dan jumlah presentase penduduk miskin memiliki nilai rata-rata yang lebih rendah dibanding dengan *cluster* lainnya. Karena jumlah pengangguran terbuka dan jumlah presentase penduduk miskinnya rendah serta pengeluaran perkapita disesuaikan, angka harapan hidup, dan rata-rata lama sekolah tinggi, maka *cluster* 2 termasuk dalam kelompok Kabupaten/Kota dengan tingkat kesejahteraan rakyat yang baik (tinggi). Yang termasuk kelompok kabupaten/kota dengan tingkat kesejahteraan rakyatnya tinggi adalah Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan, dan Kota Tegal.

Pada *cluster* 3 dapat dilihat bahwa jumlah pengangguran terbuka dan jumlah presentase penduduk miskin memiliki nilai rata-rata yang rendah, namun nilainya tetap masih rendah *cluster* 2. Sedangkan pada variabel pengeluaran perkapita, angka harapan hidup, dan rata-rata lama sekolah memiliki nilai rata-rata yang lebih tinggi dari *cluster* 1 dan lebih rendah dari *cluster* 2. Maka pada *cluster* 3 ini termasuk dalam kelompok kabupaten/kota dengan tingkat kesejahteraan rakyat yang cukup baik (sedang). Yang termasuk kelompok kabupaten/kota dengan tingkat kesejahteraan rakyatnya cukup baik (sedang) adalah Kabupaten Banyumas, Kabupaten Purbalingga, Kabupaten Kabupaten, Kabupaten Kebumen, Kabupaten Purworejo, Kabupaten Wonosobo, Kabupaten Magelang, Kabupaten Boyolali, Kabupaten Klaten, Kabupaten Sukoharjo, Kabupaten Wonogiri, Kabupaten Karanganyar, Kabupaten Sragen, Kabupaten Grobogan, Kabupaten, Kabupaten Blora, Kabupaten Rembang, Kabupaten Pati, Kabupaten Kudus, Kabupaten Jepara, Kabupaten Demak, Kabupaten Semarang, Kabupaten Temanggung, Kabupaten Kendal, Kabupaten Batang, Kabupaten Pekalongan, dan Kabupaten Pemalang.

Variabel yang memberikan perbedaan paling besar pada ketiga *cluster* yang terbentuk adalah variabel kepadatan penduduk dan rata-rata lama sekolah dengan nilai F sebesar 59,836 dan 32,245 dengan nilai signifikannya 0,00. Pembentukan *cluster* berdasarkan indikator kesejahteraan rakyat mengelompokkan 35 kabupaten/kota menjadi 3 kelompok *cluster*. Responden terbanyak ada di *cluster* 3 dengan tidak ada variabel yang hilang (*missing*). Jumlah anggota pada *cluster* 1 sebanyak 3 kabupaten/kota, jumlah anggota pada *cluster* 2 sebanyak 6 kabupaten/kota, jumlah anggota pada *cluster* 3 sebanyak 26 kabupaten/kota. Dengan demikian, semua kasus (responden) sejumlah 35 kabupaten/kota lengkap terpeta pada ketiga *cluster*.

5. SIMPULAN DAN SARATN

Penerapan metode *K-Means Cluster* digunakan untuk mengelompokkan Kabupaten/Kota di Provinsi Jawa Tengah berdasarkan indikator kesejahteraan rakyat. Sebelum dilakukan analisis *K-Means Cluster* perlu dilakukan uji asumsi multikolinieritas. Setelah asumsi terpenuhi, baru diuji lanjut analisis *K-Means Cluster*. Pengujian metode *K-Means Cluster* dalam pengelompokkan 35 Kabupaten/Kota di Provinsi Jawa Tengah berdasarkan indikator kesejahteraan rakyat dibentuk menjadi 3 kelompok *cluster*. *Cluster* 1 dengan anggota kelompok kabupaten/kota yang kesejahteraan rakyatnya kurang (rendah) terdiri dari 3 kabupaten. *Cluster* 2 dengan anggota kelompok kabupaten/kota yang kesejahteraan rakyatnya tinggi (baik) terdiri dari 6 kota. *Cluster* 3 dengan anggota kelompok kabupaten/kota yang kesejahteraan rakyatnya cukup baik (sedang) terdiri dari 26 kabupaten.

Karakteristik *cluster-cluster* yang terbentuk adalah *cluster* 1 berisi kelompok kabupaten/kota yang jumlah kepadatan penduduk, pengeluaran perkapita disesuaikan, angka harapan hidup, rata-rata lama sekolah di bawah rata-rata dan pengangguran terbuka, jumlah presentase penduduk miskin di atas rata-rata. *Cluster* 2 berisi kelompok kabupaten/kota yang jumlah kepadatan penduduk, pengeluaran perkapita disesuaikan, angka harapan hidup, rata-rata lama sekolah di atas rata-rata dan pengangguran terbuka, jumlah presentase penduduk miskin di bawah rata-rata. *Cluster* 3 berisi kelompok kabupaten/kota yang jumlah kepadatan penduduk, pengangguran terbuka, pengeluaran perkapita disesuaikan, rata-rata lama sekolah di bawah rata-rata dan angka harapan hidup, jumlah presentase penduduk miskin di atas rata-rata. Variabel yang memberikan perbedaan paling besar pada ketiga *cluster* yang terbentuk adalah variabel kepadatan penduduk dengan nilai F sebesar 59,836 dan nilai signifikannya 0,00.

Dari penjelasan simpulan di atas mengenai topik kesejahteraan rakyat diperlukan upaya peningkatan pada kesejahteraan masyarakat melalui beberapa aspek kesejahteraan seperti adanya program peningkatan pada bidang kesehatan dan pendidikan serta upaya dalam mengurangi jumlah penduduk.

6. REFERENSI

- Agusta, Y. (2007). K-Means - Penerapan, Pemasalahan dan Metode Terkait. *Jurnal Sistem dan Informatika Vol.3*, 47-60.
- Alwi, W., & Hasrul, M. (2018). Analisis Klaster Untuk Pengelompokkan Kabupaten/Kota Di Propinsi Sulawesi Selatan Berdasarkan Indikator Kesejahteraan Rakyat. *Jurnal MSA Vol. 6 No. 1*, 1.
- Awaliah, R. (2018). *Analisis Clustering Untuk Mengelompokkan Tingkat Kesejahteraan Kabupaten/Kota Berdasarkan Sosial Ekonomi Rumah Tangga Di Wilayah Provinsi Sulawesi Selatan*. Skripsi. Makassar: Universitas Islam Negeri (UIN) Alaudin Makassar.
- Azwan, R. (2013). Analisis Pengaruh Program Pengentasan Kemiskinan Desa (PPKD) Terhadap Pendapatan Usaha Kecil di Desa/Pekon Adiluhik Kabupaten Pringsewu Tahun 2012. 1.
- BPS Provinsi Jawa Tengah. (2017). *Indikator Kesejahteraan Rakyat Provinsi Jawa Tengah*. Semarang: Badan Pusat Statistika Provinsi Jawa tengah.
- Charis, C., Tedjo, P., & Martono, B. (2013). Analisis Dampak Kepadatan Penduduk Terhadap Kualitas Hidup Masyarakat Provinsi Jawa Tengah. *Serat Acitya-Jurnal Ilmiah*, 103.
- Danasari, L., & Wibowo, A. (2017). Analisis Angka Harapan Hidup di Jawa Timur Tahun 2015. *Jurnal Biometrika dan Kependudukan Vol.06 No.01*, 18.
- Ediyanto, M. (2013). pengklasifikasian karakteristik dengan metode k-means cluster analyssis. *Buletin ilmiah Mat, Stat, dan Terapan (Bimaster)*, 1.
- Fahamzah, A. (2017). K-Means Clustering Untuk Klasterisasi Kinerja Dosen Dalam Empat Kompetensi Dasar Berdasarkan Penilaian Mahasiswa. *Simki-Techsain Vol.01 No.01*.
- Gudono. (2011). *Analisis Data Multivariat Edisi Pertama*. Yogyakarta: BFE.
- Hastarini, & Dwi. (2005). Investasi Sumber Daya Manusia Melalui Pendidikan. *Jurnal Dinamika Pembangunan (JDP)*, 30-39.
- Mahmudi. (2013). *Clustering Mahasiswa Matematika Berdasarkan Karakteristik Menggunakan K-Means*. Tugas Akhir. Semarang: FMIPA UNNES.
- Mantra, I. (2007). *Demografi Umum*. Yogyakarta: Pustaka Pelajar Offset.
- Muta'ali, I. (2015). *Teknik Analisis Regional Untuk Perencanaan Wilayah, Tata Ruang, Dan Lingkungan*. Yogyakarta: Badan perbit Fakultas Geografi (BPFGe) Universitas Gajah Mada.
- Rahayu, G., & Mustakim. (2017). Principal Component Analysis untuk Dimensi Reduksi Data Clstering sebagai Pemetaan Presentase Sertifikasi Guru di Indonesia. *SNTIKI*.
- Rasbin. (2012). Faktor-FAktor Yang Mempengaruhi Tingkat Kesejahteraan Masyarakat Di Tujuh Provinsi Pemekaran Di Indonesia. *Jurnal SM*, 3.
- Rivanti, G. (2017). *Pengaruh Angka harapan hidup, Rata-Rata Lama Sekolah dan Pengeluaran Perkapita Terhadap Tingkat Kemiskinan Di Provinsi jawa Tengah Tahun 2010-2015* . Skripsi. yogyakarta: Fakultas Ekonomi Universitas Atma Jaya Yogyakarta.
- Santoso, S. (2014). *Statistik Multivariat*. Jakarta: PT Elex Media Komputtindo.
- Santoso, S. (2015). *Statistik Multivariat Konsep dasar dan Aplikasi dengan SPSS*. Jakarta: PT Elex Media Komputindo.
- Santoso, S. (2018). *Mahir Statistik Multivariat dengan SPSS*. Jakarta: PT Elex Media Komputindo.
- Sarjono, H., & Julianita, W. (2011). *SPSS vs LISREL: Sebuah Pengantar, Aplikasi untuk Riset*. Jakarta: Salemba Empat.
- Simamora, B. (2005). *Analisis Multivariat Pemasaran*. Jakarta: PT Gramedia Pustaka utama.
- Soemartini, & Supartini, E. (2017). Analisis K-Means Cluster Untuk Pengelompokkan. *Prosiding*, 144.

- Sugiantari, A., & Budiantara, I. (2013). Analisis Faktor-faktor yang Mempengaruhi Angka Harapan Hidup di Jawa Timur Menggunakan Regresi Semiparametrik Spline. *Jurnal sains dan Seni Pomits Vol.02 No.1*, D-37.
- Sujarweni, V. (2015). *SPSS untuk Penelitian*. Yogyakarta: Pustaka Baru .
- Sukestiyarno. (2011). *Statistika Dasar*. Semarang: Universitas Negeri Semarang.
- Supranto, J. (2004). *Analisis Multivariat Arti & Interpretasi*. Jakarta: PT Rineka Cipta.
- Usman, H., & Sobari, N. (2013). *Aplikasi Teknik Multivariate Untuk Riset Pemasaran*. Jakarta: PT Grafindo Persada.
- Widarjono, A. (2015). *Analisis Multivariat Terapan*. Yogyakarta: UPP STIM YKPN.
- Yulianto, S., & Hidayatullah, K. (2014). Analisis Klaster Untuk Pengelompokan Kabupaten/Kota Di Provinsi Jawa Tengah Berdasarkan Indikator Kesejahteraan Rakyat. Volume 02, pp. 56-63.

