



# Pengelompokan wilayah menurut potensi ekonomi menggunakan modifikasi algoritme fuzzy k-prototypes untuk penentuan target pembangunan desa

## *Regional clustering based on economic potential with a modified fuzzy k-prototypes algorithm for village developing target determination*

Hermawan Prasetyo<sup>\*)</sup>

Badan Pusat Statistik Provinsi Jawa Tengah  
Jl. Pahlawan Nomor 6, Kota Semarang, Jawa Tengah 50241, Indonesia

**Cara sitasi:** H. Prasetyo, "Pengelompokan wilayah menurut potensi ekonomi menggunakan modifikasi algoritme fuzzy k-prototypes untuk penentuan target pembangunan desa," *Jurnal Teknologi dan Sistem Komputer*, vol. 10, no. 1, pp. 46-52, 2022. doi: [10.14710/jtsiskom.2022.14247](https://doi.org/10.14710/jtsiskom.2022.14247), [Online].

**Abstract** – The clustering algorithm can group regions based on economic potential with mixed attributes data, consisting of numeric and categorical data. This study aims to group villages according to their economic potential in determining village development targets in Demak Regency using the fuzzy k-prototypes algorithm and modified Eskin distance to measure the distance of categorical attributes. The data used are PODES2018 data and the 2019 Wilkerstat Mapping. Village clustering produces three village clusters according to their economic potential, namely low, medium, and high economic clusters. Clusters of high economic potential are located on the main transportation routes of Semarang–Kudus and Semarang–Grobogan. However, villages on the main transportation route are still included in the low economic cluster. Considering the status of the urban/rural village classification, most of these villages are included in the urban village category. The results of this clustering can be used to determine village development targets in increasing the Village Developing Index in Demak Regency.

**Keywords** – clustering; mix attributes; fuzzy k-prototypes; village potential; village developing index

**Abstrak** – Pengelompokan wilayah berdasarkan potensi ekonomi dapat dilakukan dengan klasterisasi data yang beratribut campuran, yaitu terdiri dari data numerik dan kategorik. Penelitian ini bertujuan untuk melakukan pengelompokan desa menurut potensi ekonomi dalam menentukan target pembangunan desa di Kabupaten Demak. Klasterisasi dilakukan dengan algoritme fuzzy k-prototypes dan modifikasi jarak Eskin untuk mengukur jarak atribut kategorik. Data yang digunakan adalah data PODES2018 dan Pemetaan Wilkerstat 2019. Klasterisasi desa menghasilkan tiga klaster desa menurut potensi ekonominya, yaitu klaster ekonomi rendah, sedang,

dan tinggi. Klaster potensi ekonomi tinggi berada pada jalur transportasi utama Semarang–Kudus dan Semarang–Grobogan. Namun, desa-desa yang berada pada jalur utama transportasi tersebut masih ada yang masuk dalam klaster ekonomi rendah. Dengan mempertimbangkan status klasifikasi perkotaan/perdesaan desa, sebagian besar desa tersebut termasuk dalam kategori desa perkotaan. Hasil klasterisasi ini dapat dijadikan pedoman dalam menentukan target pembangunan desa dalam meningkatkan Indeks Desa Membangun di Kabupaten Demak.

**Kata kunci** – klasterisasi; atribut campuran; fuzzy k-prototypes; potensi desa; indeks desa membangun

### I. PENDAHULUAN

Klasterisasi sebagai salah satu teknik dalam *data mining* dapat menghasilkan kelompok-kelompok objek sedemikian rupa sehingga objek-objek di dalam klaster sangat mirip dan objek-objek di luar klaster tidak mirip [1]. Klaster adalah sebuah himpunan dari objek data yang mirip satu sama lain dalam suatu klaster dan tidak mirip dengan objek-objek yang ada di luar klaster. Pemanfaatan klasterisasi sangat luas di berbagai bidang, antara lain untuk menentukan prioritas perbaikan jalan [2], pengolahan citra digital [3], segmentasi pelanggan universitas [4], dan pengelompokan wilayah berdasarkan kapasitas investasi [5].

Klasterisasi dapat dilakukan pada tipe data numerik, kategorik, dan campuran numerik dan kategorik. Penelitian terkait klasterisasi untuk data campuran numerik dan kategorik dilakukan antara lain dengan memodifikasi jarak dan mengaplikasikannya pada algoritme yang ada [6]. Ahmad dan Dey [7] melakukan modifikasi dari algoritme k-means dengan dengan fungsi objektif dan penghitungan jarak yang baru. Shih dkk. [6] melakukan klasterisasi dua tahap untuk klasterisasi data beratribut campuran. Huang [8] melakukan klasterisasi dengan algoritme k-prototypes, yaitu suatu turunan dari algoritme k-means yang mengintegrasikan antara algoritme k-means dan k-modes. Algoritme k-prototypes mempunyai kelemahan

<sup>\*)</sup> Penulis korespondensi (Hermawan Prasetyo)  
Email: herpas@bps.go.id

seperti algoritme k-means, yaitu antara lain dalam penghitungan jarak untuk atribut kategorik [6]. Untuk mengatasi kelemahan algoritme tersebut, beberapa kajian dilakukan dengan memodifikasi algoritme k-prototypes dengan menggunakan logika fuzzy [9] dan modifikasi jarak [10], [11].

Salah satu kegunaan dari klasterisasi data beratribut campuran adalah untuk pengelompokan wilayah berdasarkan potensinya seperti dalam [10]. Potensi wilayah dapat diketahui dari hasil Pendataan Potensi Desa (PODES) [12]. Data PODES terdiri dari atribut dengan tipe data numerik dan kategorik. Pengelompokan wilayah dengan data PODES dilakukan untuk mengetahui konsentrasi aktivitas sosial ekonomi wilayah. Lebih lanjut, Prasetyo dan Purwarianti [10] membandingkan ukuran jarak atribut kategorik pada pengelompokan wilayah menggunakan data PODES. Pengelompokan wilayah dengan data PODES juga dilakukan pada [13] menggunakan modifikasi algoritme fuzzy k-prototypes (FKP).

Kajian pengelompokan wilayah pada [10]-[13] hanya memanfaatkan data PODES. Pengelompokan wilayah dalam [10] dengan membandingkan ukuran jarak kategorik pada algoritme FKP dan mengaplikasikannya pada data PODES Kota Semarang. Pengelompokan data PODES pada [12] diaplikasikan hanya pada atribut numerik sehingga tidak dapat mewakili atribut kategorik yang ada pada data PODES. Sementara itu, potensi wilayah juga dapat diketahui dari hasil Pemetaan Wilayah Kerja Statistik (Pemetaan Wilkerstat). Dengan menggabungkan informasi dari data PODES dan Pemetaan Wilkerstat, data potensi wilayah menjadi semakin lengkap untuk dianalisis.

Pendataan PODES dan Pemetaan Wilkerstat dilakukan di seluruh wilayah Indonesia, salah satunya adalah di Kabupaten Demak. Dilihat dari Indeks Desa Membangun (IDM) 2020, kabupaten ini berada pada peringkat terbawah di Jawa Tengah dengan indeks kabupaten sebesar 0,6427 dan status IDM berkembang [14]. Kebijakan pemerintah dalam mengoptimalkan potensi di suatu desa diperlukan untuk meningkatkan IDM. Salah satu potensi yang dapat dioptimalkan untuk meningkatkan IDM adalah potensi ekonomi. Analisis potensi ekonomi wilayah dapat dilakukan dengan klasterisasi.

Penelitian ini bertujuan untuk mengelompokkan wilayah di Kabupaten Demak menurut potensi ekonominya dalam menentukan target pembangunan desa untuk meningkatkan IDM Kabupaten Demak. Pengelompokan wilayah dilakukan pada level desa dengan metode klasterisasi data PODES 2018 dan Pemetaan Wilkerstat 2019 menggunakan modifikasi algoritme FKP. Dengan pengelompokan desa menurut potensi ekonomi di Kabupaten Demak ini, penentu kebijakan dapat mengetahui kelompok-kelompok desa yang perlu diutamakan dalam pembangunan desa.

## II. METODE PENELITIAN

Penelitian ini dilakukan dengan beberapa tahapan, yaitu meliputi penentuan data, prapemrosesan data berupa seleksi dan agregasi atribut, pemrosesan data dengan algoritme FKP dan modifikasi jarak, evaluasi hasil klasterisasi, dan visualisasi hasil klasterisasi dengan peta digital. Dari hasil klasterisasi dan visualisasi data dilakukan analisis untuk penentuan target pembangunan desa dalam meningkatkan IDM.

### A. Penyiapan data

Data yang digunakan dalam penelitian ini berasal dari data PODES18-DESA dan SP2020-RD untuk wilayah Kabupaten Demak. Data PODES18-DESA terdiri dari 249 baris, yaitu sejumlah desa yang ada. Sementara itu, data SP2020-RD terdiri dari 22.745 baris, yaitu sejumlah Satuan Wilayah Setempat (SLS) terkecil dan non-SLS dalam satu Kabupaten Demak.

Atribut yang tersedia pada data PODES18-DESA sejumlah 780 atribut, sedangkan atribut yang tersedia pada data SP2020-RD sebanyak 72 atribut. Jumlah atribut menunjukkan jumlah identitas baris dan variabel yang dimiliki tiap baris.

### B. Prapemrosesan data

Pada tahap ini dilakukan seleksi terhadap atribut pada data Pendataan Potensi Desa/Kelurahan 2018 (PODES18-DESA)<sup>1</sup> dan Daftar Rekap Desa dalam Pemetaan dan Pemutakhiran Muatan Wilkerstat SP2020, 2019 (SP2020-RD)<sup>2</sup>. Atribut diseleksi berdasarkan variabel yang ada pada BLOK XII kuesioner PODES18-DESA, yaitu blok tentang ekonomi. Pemilihan variabel tersebut diharapkan dapat mencerminkan potensi ekonomi suatu wilayah. Untuk data SP2020-RD, variabel yang dipilih adalah jumlah bangunan sensus kegiatan ekonomi (BSKEKO). Selain seleksi atribut juga dilakukan agregasi nilai atribut, yaitu penjumlahan nilai dari beberapa atribut yang ada menjadi satu nilai atribut baru. Hasil dari prapemrosesan data diperoleh beberapa atribut seperti pada Tabel 1.

### C. Pemrosesan data

Pengembangan algoritme klasterisasi antara lain dilakukan dengan melakukan modifikasi pengukuran jarak [10], [11], [15]. Jarak pada klasterisasi digunakan untuk menghitung fungsi objektif. Pengembangan ini juga dilakukan pada algoritme k-prototypes, dengan menerapkan logika fuzzy dalam FKP [9] dan modifikasi jarak Eskin untuk data kategorik [13]. FKP adalah pengembangan dari algoritme k-prototypes untuk melakukan klasterisasi data dengan atribut campuran. Jika  $X$  adalah sebuah himpunan objek  $X_1, X_2, \dots, X_n$  yang mempunyai atribut  $A_1, A_2, \dots, A_p$  dan atribut kategorik  $A_1, A_2, \dots, A_q$  dan  $k$  adalah jumlah klaster,

<sup>1</sup> <https://sirusa.bps.go.id/sirusa/index.php/kuesioner/1987> (diakses: 19 Agustus 2021)

<sup>2</sup> <https://sirusa.bps.go.id/sirusa/index.php/kuesioner/2247> (diakses: 19 Agustus 2021)

**Tabel 1.** Daftar atribut terpilih pada data PODES18-DESA dan SP2020-RD

No	Rincian	Deskripsi	Type Data	Sumber Data
1.	R1201	Jumlah industri mikro dan kecil	Numerik	PODES
2.	R1202	Jumlah sentra industri, lingkungan industri kecil, dan perkampungan industri kecil	Numerik	PODES
3.	R1203a	Keberadaan pangkalan/agen/penjual minyak tanah	Kategorik	PODES
4.	R1203b	Keberadaan pangkalan/agen/penjual LPG	Kategorik	PODES
5.	R1204	Jumlah KUD	Numerik	PODES
6.	R1205a	Jumlah koperasi selain KUD	Numerik	PODES
7.	R1205b1	Keberadaan toko penjual sarana produksi pertanian milik KUD	Kategorik	PODES
8.	R1205b2	Keberadaan toko penjual sarana produksi pertanian milik BUM Desa	Kategorik	PODES
9.	R1205b3	Keberadaan toko penjual sarana produksi pertanian selain milik KUD dan BUM Desa	Kategorik	PODES
10.	R1206	Jumlah sarana dan prasarana ekonomi	Numerik	PODES
11.	R1207a	Fasilitas kredit KUR	Kategorik	PODES
12.	R1207b	Fasilitas kredit KKP-E	Kategorik	PODES
13.	R1207c	Fasilitas kredit KUK	Kategorik	PODES
14.	R1207d	Fasilitas kredit KUBE	Kategorik	PODES
15.	R1208	Jumlah sarana lembaga keuangan	Numerik	PODES
16.	R1209a	Keberadaan BMT	Kategorik	PODES
17.	R1209b	Keberadaan pegadaian	Kategorik	PODES
18.	R1209c	Keberadaan ATM	Kategorik	PODES
19.	R1209d	Keberadaan bengkel	Kategorik	PODES
20.	R1209e	Keberadaan salon	Kategorik	PODES
21.	R1209f	Keberadaan agen tiket	Kategorik	PODES
22.	R1209g	Keberadaan agen bank	Kategorik	PODES
23.	BSKEKO	Jumlah bangunan sensus kegiatan ekonomi	Numerik	RD

permasalahan pada algoritme FKP dapat diatasi dengan meminimalkan fungsi objektif  $F(W, Z)$ , yang didefinisikan dalam (1) di mana  $\alpha \geq 1$  adalah sebuah parameter fuzzy,  $W = [w_{li}]$  adalah sebuah matriks keanggotaan berukuran  $k \times n$ , dan  $Z = [Z_1, Z_2, \dots, Z_k]$  adalah sebuah himpunan sejumlah  $k$  prototipe berukuran  $m$  dimensi.

$$F(W, Z) = \sum_{l=1}^k \sum_{i=1}^n w_{li}^\alpha d(X_i, Z_l) \quad (1)$$

Langkah-langkah dalam algoritme FKP didefinisikan oleh Chen dkk. [9] sebagai berikut:

- (1) Memilih sebuah himpunan prototipe awal  $Z^{(l)}$  dan sebuah pengontrol iterasi  $\epsilon$ .
- (2) Menentukan  $W^{(l)}$  yang meminimalkan  $F(W, Z^{(l)})$ . Mengatur  $i = 1$ .
- (3) Menentukan  $Z^{(i+1)}$  yang meminimalkan  $F(W^{(i)}, Z)$ . Jika  $|F(W^{(i)}, Z^{(i+1)}) - F(W^{(i)}, Z^{(i)})| < \epsilon$ , maka berhenti. Jika tidak, maka mengatur  $(W^*, Z^*) = (W^{(i)}, Z^{(i+1)})$ .
- (4) Menentukan  $W^{(i+1)}$  yang meminimalkan  $F(W, Z^{(i+1)})$ . Jika  $|F(W^{(i+1)}, Z^{(i+1)}) - F(W^{(i)}, Z^{(i+1)})| < \epsilon$ , maka berhenti. Jika tidak, maka atur  $(W^*, Z^*) = (W^{(i+1)}, Z^{(i+1)})$ . Menambahkan  $i$  dengan 1 kemudian kembali ke Langkah 3.

Dua jenis ukuran yang digunakan untuk estimasi kemiripan (*similarity*) atau ketidakmiripan (*dissimilarity*) dari dua objek dalam klusterisasi adalah ukuran jarak dan kemiripan [16]. Salah satu metode pengukuran jarak untuk data kategorik adalah jarak Eskin [17], [18]. Kemiripan tiap atribut antara dua objek  $x$  dan  $y$  dengan ukuran Eskin dapat dihitung dengan (2).

$$s_k(x_k, y_k) = \begin{cases} 1, \wedge x_k = y_k \\ \frac{n_k^2}{n_k^2 + 2}, \wedge x_k \neq y_k \end{cases} \quad (2)$$

Modifikasi jarak Eskin dilakukan dengan menggunakan parameter peluang kejadian nilai suatu atribut [13]. Diasumsikan bahwa  $\hat{p}_k(x)$  adalah estimasi peluang atribut  $A_k$  bernilai  $x$  dan  $\hat{p}_k(y)$  adalah estimasi peluang atribut  $A_k$  bernilai  $y$ , dengan  $\hat{p}_k(x)$  dan  $\hat{p}_k(y)$  adalah dua peluang diskrit. Berdasarkan kemunculannya,  $\hat{p}_k(x)$  dan  $\hat{p}_k(y)$  adalah dua kejadian yang saling lepas, yaitu atribut  $A_k$  hanya dapat bernilai  $x$  atau  $y$  dalam satu objek data. Pada kejadian saling lepas  $A$  dan  $B$ , nilai peluang munculnya  $A$  atau  $B$  dapat dihitung dengan (3), dengan  $P(A \cup B)$  adalah peluang kejadian  $A$  atau  $B$ ,  $P(A)$  adalah peluang kejadian  $A$ , dan  $P(B)$  adalah peluang kejadian  $B$ .

$$P(A \cup B) = P(A) + P(B) \quad (3)$$

Dari teori tersebut dapat dinyatakan optimasi kemiripan dalam ukuran kemiripan per atribut (4). ukuran kemiripan per atribut dapat didefinisikan dalam (5).

$$s_k(x_k, y_k) = \begin{cases} 1, \wedge x_k = y_k \\ \frac{n_k^2}{n_k^2 + \hat{p}_k(x \cup y)}, \wedge x_k \neq y_k \end{cases} \quad (4)$$

$$s_k(x_k, y_k) = \begin{cases} 1, \wedge x_k = y_k \\ \frac{n_k^2}{n_k^2 + \hat{p}_k(x) + \hat{p}_k(y)}, \wedge x_k \neq y_k \end{cases} \quad (5)$$

#### D. Evaluasi

Evaluasi hasil klusterisasi dilakukan untuk mengetahui kualitas hasil klusterisasi. Metode evaluasi menggunakan indeks *partition coefficient* (PC) dan indeks *partition entropy* (PE) [1].

Diasumsikan  $W = (w_{li}), 1 \leq l \leq c, 1 \leq i \leq n$  merupakan matriks keanggotaan dari sebuah  $c$  partisi fuzzy dari data  $D$  dengan  $n$  objek. Penghitungan indeks PC dapat dilakukan dengan (6).

$$V_{PC} = \frac{1}{n} \sum_{l=1}^c \sum_{i=1}^n w_{li}^2 \quad (6)$$

Nilai dari indeks PC berada pada rentang  $[1/c, 1]$ . Klusterisasi semakin dapat memisahkan objek dalam kluster yang tepat jika nilai indeks PC yang semakin mendekati 1. Sementara itu, nilai indeks PC yang semakin mendekati  $1/c$  menunjukkan bahwa hasil klusterisasi masih belum dapat memisahkan objek dalam kluster yang tepat.

Indeks PE dapat dihitung dengan (7), dengan  $a$  adalah bilangan basis logaritma dan  $W = (w_{li})$  adalah matriks keanggotaan dari sebuah  $c$  partisi fuzzy. Nilai dari indeks PE berada pada rentang  $[0, \log_a c]$ . Klusterisasi semakin semakin baik, yaitu dapat memisahkan objek dalam kluster yang tepat jika nilai indeks mendekati 0.

$$V_{PE} = -\frac{1}{n} \sum_{l=1}^c \sum_{i=1}^n w_{li} \log_a(w_{li}) \quad (7)$$

#### E. Visualisasi data

Dari hasil klusterisasi, visualisasi data dilakukan berupa peta pengelompokan desa menurut kluster yang terbentuk. Peta digital yang digunakan adalah peta Wilkerstat desa hasil Pemetaan Wilkerstat 2019 dan Google Road. Piranti lunak yang digunakan adalah QGIS 3.10.14.

### III. HASIL DAN PEMBAHASAN

Dari data yang ada dilakukan percobaan klusterisasi dengan nilai koefisien kekaburan/fuzziness  $\lambda_1, \lambda_2, \dots, \lambda_c$  konstan sebesar 2 pada beberapa jumlah kluster. Pemilihan nilai konstan koefisien kekaburan ini berdasarkan pada nilai koefisien yang paling banyak digunakan dalam klusterisasi fuzzy, walaupun angka ini tidak selalu optimal terutama untuk data dengan jumlah kluster yang sangat bervariasi [19].

Percobaan menghasilkan nilai indeks PC dan PE seperti ditunjukkan pada Tabel 2. Percobaan dengan jumlah kluster dua menghasilkan hasil klusterisasi yang paling baik dibanding yang lain. Pada jumlah kluster dua, nilai indeks PC mendekati 1 dan indeks PE mendekati nol. Hal ini mengindikasikan bahwa klusterisasi semakin dapat mengelompokkan objek pada kluster yang tepat.

**Tabel 2.** Nilai indeks PC dan PE pada percobaan klusterisasi

Jumlah Kluster	Indeks PC	Indeks PE
2	0,8376	0,2733
3	0,7430	0,4682
4	0,6483	0,6662
5	0,5717	0,8280

Pada jumlah kluster tiga, nilai indeks PC masih mendekati satu tetapi nilainya di bawah indeks PC pada percobaan dengan jumlah kluster dua. Nilai indeks PE semakin menjauhi angka nol tetapi masih di bawah 0,5. Dari nilai kedua indeks tersebut dapat disimpulkan bahwa klusterisasi masih dapat mengelompokkan objek pada kluster yang tepat.

Pada jumlah kluster empat, nilai indeks PC dan PE hampir sama, bahkan nilai indeks PE melebihi indeks PC. Hal ini juga terjadi pada klusterisasi dengan jumlah kluster lima. Dengan demikian, pada jumlah kluster empat dan lima proses klusterisasi tidak dapat mengelompokkan objek-objek ke dalam kluster yang tepat dengan baik.

Dari hasil percobaan klusterisasi desa, studi kasus dilakukan untuk pengelompokan desa di Kabupaten Demak ke dalam tiga kelompok. Walaupun jumlah kluster terbaik adalah dua, studi kasus menggunakan jumlah kluster tiga dengan tujuan untuk mendapatkan hasil analisis yang lebih beragam. Jika dilihat dari hasil evaluasi klusterisasi pada Tabel 2, pada jumlah kluster tiga hasil klusterisasi masih dapat dikatakan baik karena nilai indeks PC sebesar 0,7430 masih mendekati 1 dan nilai indeks PE sebesar 0,4682 masih dapat dikatakan mendekati 0.

Klusterisasi desa ke dalam tiga kelompok di Kabupaten Demak menghasilkan nilai pusat seperti ditunjukkan pada Tabel 3. Variabel bertipe kategorik ketiga kluster memiliki nilai pusat yang hampir sama, kecuali nilai pusat variabel 'keberadaan agen bank', yang bernilai 'tidak ada' pada kluster 2. Untuk variabel bertipe numerik ketiga kluster memiliki nilai pusat yang berbeda. Jika dilihat dari nilai pusat variabel bertipe numerik, kluster 1 menunjukkan nilai pusat yang paling tinggi. Dari nilai pusat dapat disimpulkan bahwa desa-desa pada kluster 1 memiliki potensi ekonomi yang paling tinggi, yang disusul oleh Kluster 3 dan Kluster 2.

Hasil klusterisasi selanjutnya divisualisasikan dalam bentuk peta klusterisasi desa seperti pada Gambar 1. Desa-desa di Kabupaten Demak dibagi menjadi tiga kluster, yaitu Kluster 1 (hijau/ekonomi tinggi), Kluster 2 (merah/ekonomi rendah), dan Kluster 3 (kuning/ekonomi sedang). Jika dilihat berdasarkan lokasi, sebagian besar desa pada Kluster 1 terdapat di sekitar jalur transportasi utama Semarang – Kudus dan Semarang – Grobogan. Beberapa desa pada Kluster 3 juga terletak pada jalur utama Semarang – Kudus dan Semarang – Grobogan.

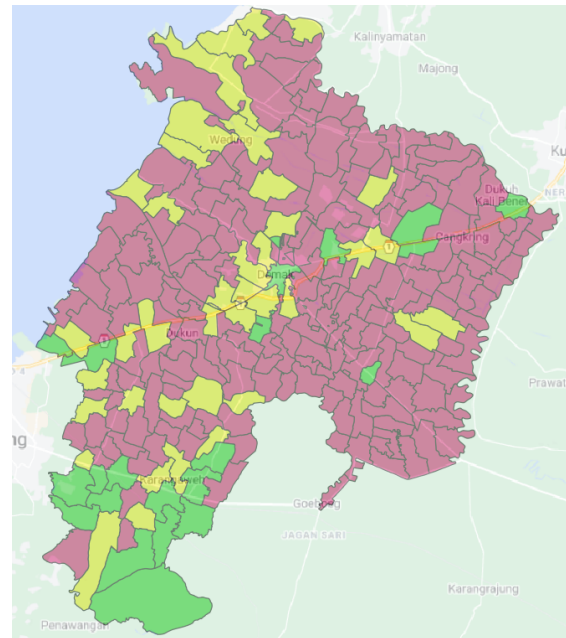
Dari hasil nilai pusat, desa-desa pada Kluster 1 dan 3 memiliki potensi ekonomi yang lebih tinggi dibanding desa-desa pada Kluster 2. Potensi ekonomi yang lebih

**Tabel 3.** Nilai pusat hasil pengelompokan desa di Kabupaten Demak

Variabel	Nilai Pusat		
	Klaster 1	Klaster 2	Klaster 3
Jumlah industri mikro dan kecil	203,88	23,86	69,40
Jumlah sentra industri, lingkungan industri kecil, dan perkampungan industri kecil	0,30	0,18	0,24
Keberadaan pangkalan/agen/penjual minyak tanah	Tidak ada	Tidak ada	Tidak ada
Keberadaan pangkalan/agen/penjual LPG	Ada	Ada	Ada
Jumlah KUD	0,42	0,10	0,12
Jumlah koperasi selain KUD	1,29	0,59	0,79
Keberadaan toko penjual sarana produksi pertanian milik KUD	Tidak ada	Tidak ada	Tidak ada
Keberadaan toko penjual sarana produksi pertanian milik BUM Desa	Tidak ada	Tidak ada	Tidak ada
Keberadaan toko penjual sarana produksi pertanian selain milik KUD dan BUM Desa	Ada	Ada	Ada
Jumlah sarana dan prasarana ekonomi	178,92	58,39	98,07
Fasilitas kredit KUR	Ada	Ada	Ada
Fasilitas kredit KKP-E	Tidak ada	Tidak ada	Tidak ada
Fasilitas kredit KUK	Tidak ada	Tidak ada	Tidak ada
Fasilitas kredit KUBE	Tidak ada	Tidak ada	Tidak ada
Jumlah sarana lembaga keuangan	1,16	0,16	0,54
Keberadaan BMT	Tidak ada	Tidak ada	Tidak ada
Keberadaan pegadaian	Tidak ada	Tidak ada	Tidak ada
Keberadaan ATM	Tidak ada	Tidak ada	Tidak ada
Keberadaan bengkel	Ada	Ada	Ada
Keberadaan salon	Ada	Ada	Ada
Keberadaan agen tiket	Tidak ada	Tidak ada	Tidak ada
Keberadaan agen bank	Ada	Tidak ada	Ada
Jumlah bangunan sensus kegiatan ekonomi	537,36	125,73	300,21

tinggi dimungkinkan karena lokasi desa yang strategis sehingga mendorong kegiatan ekonomi untuk mendukung kedua jalur transportasi tersebut. Selain pada lokasi jalur transportasi utama, desa-desa yang termasuk dalam Klaster 1 dan 3 terdapat pada beberapa wilayah yang lain. Perlu pendalaman lebih lanjut terkait penyebab desa-desa tersebut termasuk dalam klaster desa dengan potensi ekonomi tinggi.

Jika dilihat pada **Gambar 1**, desa-desa yang berlokasi di jalur utama transportasi juga ada yang masuk dalam Klaster 2, yaitu desa dengan potensi ekonomi rendah. Jika dilihat dari klasifikasi perkotaan/perdesaan Badan Pusat Statistik dalam Perka BPS Nomor 120 Tahun 2020 [20], klasifikasi desa-desa



**Gambar 1.** Hasil pengelompokan desa di Kabupaten Demak

**Tabel 4.** Klasifikasi perkotaan/perdesaan desa-desa pada Klaster 2 yang terletak di jalur transportasi utama

Kecamatan	Desa	Klasifikasi
Sayung	Sidogemah	Perkotaan
Sayung	Gemulak	Perkotaan
Karang Tengah	Wonokerto	Perkotaan
Karang Tengah	Dukun	Perkotaan
Karang Tengah	Karangtowo	Perkotaan
Karang Tengah	Pulosari	Perkotaan
Wonosalam	Mranak	Perkotaan
Demak	Cabean	Perkotaan
Demak	Bolo	Perkotaan
Gajah	Sari	Perkotaan
Karanganyar	Wonoketingal	Perkotaan
Karanganyar	Wonorejo	Perkotaan
Karanganyar	Cangkring	Perkotaan
Karangawen	Sido Rejo	Perdesaan

pada Klaster 2 yang berada pada jalur transportasi utama dapat dilihat pada **Tabel 4**.

Sebagian besar desa yang masuk dalam Klaster 2 dan berada pada jalur utama Semarang – Kudus dan Semarang – Grobogan berstatus perkotaan, kecuali Desa Sido Rejo. Walaupun desa-desa tersebut berstatus perkotaan dan berada pada jalur transportasi utama, desa-desa tersebut masih masuk dalam klaster potensi ekonomi rendah. Hal ini dapat disebabkan karena penentuan klasifikasi perkotaan/perdesaan suatu desa tidak hanya berdasarkan indikator ekonomi, tetapi juga indikator kepadatan penduduk, persentase keluarga pertanian, keberadaan fasilitas sekolah, keberadaan rumah sakit, persentase keluarga pengguna telepon kabel, dan persentase keluarga pengguna listrik [20]. Selain itu, indikator ekonomi yang digunakan untuk penentuan klasifikasi perkotaan/perdesaan juga tidak

sebanyak indikator ekonomi yang digunakan pada penelitian ini.

Lokasi wilayah yang dekat dengan infrastruktur transportasi seharusnya menopang ekonomi wilayah tersebut. Jalan sebagai infrastruktur transportasi memiliki peran penting di bidang ekonomi serta sebagai prasarana distribusi barang dan jasa [21]. Kartiasih [22] menjelaskan bahwa infrastruktur transportasi mempengaruhi pertumbuhan ekonomi. Palilu dan Suripatty [23] juga menjelaskan bahwa infrastruktur transportasi mempunyai pengaruh positif terhadap pertumbuhan ekonomi. Namun, pada penelitian ini masih terdapat desa-desa kategori perkotaan yang berada pada jalur utama transportasi yang masuk dalam kluster ekonomi rendah. Hal ini dapat dijadikan bahan penelitian lanjutan terkait penyebab desa-desa tersebut memiliki potensi ekonomi lebih rendah. Selain itu, desa-desa pada kluster tersebut juga dapat dijadikan desa prioritas untuk pembangunan perekonomian desa agar lebih meningkatkan indeks desa membangun di Kabupaten Demak.

#### IV. KESIMPULAN

Klasterisasi desa di Kabupaten Demak dengan algoritme FKP menghasilkan tiga kluster desa menurut potensi ekonominya. Desa pada kluster ekonomi tinggi sebagian besar berada pada wilayah jalur transportasi utama. Namun, ditemukan 14 desa pada jalur transportasi utama yang masuk dalam kluster ekonomi rendah. Jika dilihat dari klasifikasi perkotaan/perdesaan pada 14 desa tersebut, status perkotaan desa tidak selalu menunjukkan suatu desa masuk dalam kluster ekonomi tinggi. Informasi tentang desa perkotaan yang masuk dalam kluster ekonomi rendah dapat digunakan sebagai pedoman dalam penentuan target pembangunan desa untuk meningkatkan IDM.

#### UCAPAN TERIMA KASIH

Terima kasih diucapkan kepada Badan Pusat Statistik dan Pusat Pendidikan dan Pelatihan Badan Pusat Statistik yang telah mendukung dan membantu dalam penelitian ini.

#### DAFTAR PUSTAKA

- [1] G. Gan, C. Ma, and J. Wu, *Data clustering: Theory, algorithms, and applications, second edition*. 2020. doi: [10.1137/1.9781611976335](https://doi.org/10.1137/1.9781611976335)
- [2] N. Puspitasari, R. Rosmasari, and S. Stefanie, "Penentuan prioritas perbaikan jalan menggunakan fuzzy c-means: studi kasus perbaikan jalan di Kota Samarinda," *Jurnal Teknologi dan Sistem Komputer*, vol. 5, no. 1, pp. 7-14, 2017. doi: [10.14710/jtsiskom.5.1.2017.7-14](https://doi.org/10.14710/jtsiskom.5.1.2017.7-14)
- [3] L. Arsy, O. D. Nurhayati, and K. T. Martono, "Aplikasi pengolahan citra digital meat detection dengan metode segmentasi k-mean clustering berbasis Opencv dan Eclipse," *Jurnal Teknologi dan Sistem Komputer*, vol. 4, no. 2, pp. 322-332, 2016. doi: [10.14710/jtsiskom.4.2.2016.322-332](https://doi.org/10.14710/jtsiskom.4.2.2016.322-332)
- [4] S. Hidayat, R. Rismayati, M. Tajuddin, and N. L. P. Merawati, "Segmentation of university customers loyalty based on RFM analysis using fuzzy c-means clustering," *Jurnal Teknologi dan Sistem Komputer*, vol. 8, no. 2, pp. 133-139, 2020. doi: [10.14710/jtsiskom.8.2.2020.133-139](https://doi.org/10.14710/jtsiskom.8.2.2020.133-139)
- [5] T. Nabarian, S. Sutoto, N. Gusmawati, D. P. Trimaratus Sholehah, A. Nizar Hidayanto, and A. M. Sari, "Clustering of provinces in Indonesia based on regional investment capacity with density-based spatial clustering of applications with noise method," in *International Conference on Computing Engineering and Design*, Singapore, Apr. 2019, pp. 3-8. doi: [10.1109/ICCED46541.2019.9161110](https://doi.org/10.1109/ICCED46541.2019.9161110)
- [6] M. Y. Shih, J. W. Jheng, and L. F. Lai, "A two-step method for clustering mixed categorical and numeric data," *Tamkang Journal of Science and Engineering*, vol. 13, no. 1, pp. 11-19, 2010. doi: [10.6180/jase.2010.13.1.02](https://doi.org/10.6180/jase.2010.13.1.02)
- [7] A. Ahmad and L. Dey, "A k-mean clustering algorithm for mixed numeric and categorical data," *Data and Knowledge Engineering*, vol. 63, no. 2, pp. 503-527, 2007. doi: [10.1016/j.datak.2007.03.016](https://doi.org/10.1016/j.datak.2007.03.016)
- [8] Z. Huang, "Clustering large data sets with mixed numeric and categorical values," in *the Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Singapore, Feb. 1997, pp. 21-34.
- [9] N. Chen, A. Chen, and L. X. Zhou, "Fuzzy K-prototypes algorithm for clustering mixed numeric and categorical valued data," *Ruan Jian Xue Bao/Journal of Software*, vol. 12, no. 8. pp. 1107-1119, 2001.
- [10] H. Prasetyo and A. Purwarianti, "Comparison of distance measures for clustering data with mix attribute types for Indonesian potential-based regional grouping," in *International Conference on Information Technology Systems and Innovation*, Bandung, Indonesia, Nov. 2014, pp. 13-18. doi: [10.1109/ICITSI.2014.7048230](https://doi.org/10.1109/ICITSI.2014.7048230)
- [11] H. Prasetyo and A. Purwarianti, "Comparison of distance and dissimilarity measures for clustering data with mix attribute types," in *the International Conference on Information Technology, Computer, and Electrical Engineering: Green Technology and Its Applications for a Better Future*, Semarang, Indonesia, Nov. 2014, pp. 276-280. doi: [10.1109/ICITACEE.2014.7065756](https://doi.org/10.1109/ICITACEE.2014.7065756)
- [12] Badan Pusat Statistik, *Peta Tematik Lokasi Pemusatan Kegiatan dan Sarana Sosial Ekonomi di Indonesia*. Jakarta: Badan Pusat Statistik, 2012.
- [13] H. Prasetyo, "Perbaikan algoritma fuzzy k-prototypes untuk pengelompokan data beratribut campuran (studi kasus: pengelompokan desa menurut potensinya di Provinsi Jawa Tengah)," Institut Teknologi Bandung, Bandung, 2015.
- [14] Kementerian Desa Pembangunan Daerah Tertinggal dan Transmigrasi, *Peringkat Status Indeks Desa*

*Membangun (IDM) Provinsi Jawa Tengah*. Jakarta, 2020.

- [15] A. Aditya, B. N. Sari, and T. N. Padilah, "Perbandingan pengukuran jarak Euclidean dan Gower pada kluster k-medoids Comparison analysis of Euclidean and Gower distance measures on k-medoids," *Jurnal Teknologi dan Sistem Komputer*, vol. 9, no. 1, pp. 1–7, 2021. doi: [10.14710/jtsiskom.2021.13747](https://doi.org/10.14710/jtsiskom.2021.13747)
- [16] S. Wang and W. Shi, *Data mining and knowledge discovery*. 2012.
- [17] E. Eskin, A. Arnold, M. Prerau, L. Portnoy, and S. Stolfo, "A geometric framework for unsupervised anomaly detection," *Applications of Data Mining in Computer Security*, pp. 77–101, 2002. doi: [10.1007/978-1-4615-0953-0\\_4](https://doi.org/10.1007/978-1-4615-0953-0_4)
- [18] S. Boriah, V. Chandola, and V. Kumar, "Similarity measures for categorical data: A comparative evaluation," in *Proceedings of the 2008 SIAM International Conference on Data Mining*, vol. 1, pp. 243–254, 2008. doi: [10.1137/1.9781611972788.22](https://doi.org/10.1137/1.9781611972788.22)
- [19] K. Zhou and S. Yang, "Fuzzifier selection in fuzzy c-means from cluster size distribution perspective," *Informatica*, vol. 30, no. 3, pp. 613–628, 2019. doi: [10.15388/informatica.2019.221](https://doi.org/10.15388/informatica.2019.221)
- [20] Badan Pusat Statistik, "Klasifikasi desa perkotaan dan pedesaan di Indonesia tahun 2020," Perka BPS Nomor 120 Tahun 2020, 2020. Available: <https://www.bps.go.id/website/fileMenu/Perka-Nomor-120-Tahun-2020.pdf> [Accessed: Aug. 18, 2021]
- [21] A. Prasetyo, "Efisiensi transportasi untuk pertumbuhan ekonomi di Jawa Tengah melalui pendekatan interaksi spasial," *Jurnal Penelitian Transportasi Darat*, vol. 17, no. 3, pp. 157–170, 2015.
- [22] F. Kartiasih, "Dampak infrastruktur transportasi terhadap pertumbuhan ekonomi di Indonesia menggunakan regresi data panel," *Jurnal Ilmiah Ekonomi dan Bisnis*, vol. 16, no. 1, pp. 67–77, 2019. doi: [10.31849/jieb.v16i1.2306](https://doi.org/10.31849/jieb.v16i1.2306)
- [23] A. Palilu and R. Suripatty, "Pengaruh infrastruktur transportasi terhadap pertumbuhan ekonomi Kota Sorong Provinsi Papua Barat," *Ekuivalensi, Jurnal Ekonomi Bisnis*, vol. 4, no. 2, pp. 238–257, 2016.



©2022. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).