EVALUASI DAERAH TERTINGGAL DI JAWA TIMUR BERDASARKAN INDIKATOR KEMENTERIAN NEGARA PEMBANGUNAN DAERAH TERTINGGAL (KPDT) MENGGUNAKAN *ENSEMBLE ROBUST CLUSTERING USING LINK (ROCK)*

SKRIPSI

Diajukan untuk memenuhi salah satu persyaratan dalam memperoleh
gelar Sarjana Matematika (S.Mat) pada program studi Matematika

Disusun Oleh
LULUK WULANDARI
NIM.H72215017

PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
SURABAYA
2019

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama	: Luluk Wulandari
NIM	: H72215017
Program Studi	: Matematika
Angkatan	: 2015

Menyatakan bahwa saya tidak melakukan plagiat dalam tulisan saya yang berjudul : **"EVALUASI DAERAH TERTINGGAL DI JAWA TIMUR BERDASARKAN INDIKATOR KEMENTERIAN NEGARA PEMBANGUNAN DAERAH TERTINGGAL (KPDT) MENGGUNAKAN *ENSEMBLE ROBUST CLUSTERING USING LINK* (ROCK)"**. Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 02 Juli 2019

Yang menyatakan,

Luluk Wulandari

NIM.H72215017

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

NAMA : Luluk Wulandari

NIM : H72215017

JUDUL : Evaluasi Daerah Tertinggal Di Jawa Timur Berdasarkan Indikator Kementerian Negara Pembangunan Daerah Tertinggal (KPDT) Menggunakan *Ensemble Robust Clustering Using Link* (ROCK) .

Ini telah diperiksa dan disetujui untuk diujikan

Surabaya, 29 Maret 2019

Dosen Pembimbing 1	Dosen Pembimbing 2
(Yuniar Farida, M.T)	(Aris Fatani, M.Kom)
NIP : 197905272014032002	NIP : 198701272014031002

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh

Nama : Luluk Wulandari

NIM : H72215017

Judul : Evaluasi Daerah Tertinggal Di Jawa Timur Berdasarkan Indikator Kementerian Negara Pembangunan Daerah Tertinggal (KPDT) Menggunakan *Ensemble Robust Clustering Using Link* (ROCK)

Telah dipertahankan di depan Tim Penguji Skripsi pada

Tanggal 02 Juli 2019

Mengesahkan,
Tim Penguji

Penguji 1

(Yuniar Farida, M.T)
NIP : 197905272014032002

Penguji II

(Aris Fanani, M.Kom)
NIP : 198701272014031002

Penguji 1II

(Nurissaidah Ulinnuha, M.Kom)
NIP : 199901102201403200 4

Penguji IV

(Putroue Keumala Intan,M.Si)
NIP : 198805282018012001

Mengetahui,
Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya

Dr. Eni Purwati, M.Ag
NIP.196512211990022001

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN
Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Luluk Wulandari
NIM : H72215017
Fakultas/Jurusan : Sains dan Teknologi / Matematika
E-mail address : lulukwulandari618@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :
☑ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :

Evaluasi Daerah Tertinggal Di Jawa Timur Berdasarkan Indikator Kementerian Negara Pembangunan Daerah Tertinggal (KPDT) Menggunakan Ensemble K-obust Clustering Using Link (K-OCK)

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Ekslusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara ***fulltext*** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 02 Juli 2019

Penulis

(Luluk Wulandari)
nama terang dan tanda tangan

ABSTRAK
EVALUASI DAERAH TERTINGGAL DI JAWA TIMUR BERDASARKAN INDIKATOR KEMENTERIAN NEGARA PEMBANGUNAN DAERAH TERTINGGAL (KPDT) MENGGUNAKAN *ENSEMBLE ROBUST CLUSTERING USING LINK (ROCK)*

Jawa Timur merupakan salah satu Provinsi yang memiliki permasalahan dalam hal pemerataan pembangunan. Berbagai permasalahan ketimpangan dan kesenjangan menyebabkan Provinsi Jawa Timur menjadi salah satu dari 21 Provinsi yang berstatus tertinggal di Indonesia. Hal ini tercantum dalam keputusan yang terdapat dalam Peraturan Presiden Nomor 131 Tahun 2015 yang menunjukkan bahwa Provinsi Jawa Timur menjadi satu-satunya Provinsi di Pulau Jawa yang berstatus tertinggal. Penetapan daerah tertinggal dilakukan 5 tahun sekali menggunakan 6 kriteria dan 33 indikator yang telah ditetapkan oleh KPDT. Namun dari beberapa penelitian yang telah dilakukan tentang penetapan daerah tertinggal, belum ada penelitian yang menerapkan 33 indikator tersebut secara keseluruhan. Sehingga dalam penelitian ini akan dilakukan penetapan daerah tertinggal dengan menggunakan 33 indikator yang telah ditetapkan oleh KPDT. Data kriteria KPDT yang digunakan adalah hasil survey tahun 2014 dan tahun 2018. Hasil penetapan dengan data survey tahun 2014 akan digunakan sebagai pembanding dengan keputusan Pemerintah tahun 2015, dan hasil penetapan dengan data survey tahun 2018 akan digunakan sebagai prediksi penetapan daerah tertinggal tahun 2020 mendatang. Hasil dari penelitian ini diharapkan dapat memberikan gambaran terkait daerah yang ditetapkan sebagai daerah tertinggal untuk kedepannya. Adapun metode yang digunakan adalah *ensemble* ROCK, karena metode ini memiliki akurasi yang lebih baik jika dibandingkan dengan metode hierarki *agglomerative* (pemusatan) dengan sifat skalabilitas yang baik. Hasil klaster penetapan daerah tertinggal tahun 2015 terdiri dari Trenggalek, Jember, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Bangkalan, Sampang, Pamekasan, Sumenep, dan Kota Probolinggo. Hasil klaster terbaik terdiri dari 4 klaster dengan nilai $ratio\ sw$ dan sb terkecil sebesar $0{,}3917473$ dan nilai *threshold* optimum sebesar $0{,}02$. Sedangkan untuk prediksi penetapan daerah tertinggal untuk tahun 2020 terdiri dari Trenggalek, Bondowoso, Situbondo, Probolinggo, Tuban, Pamekasan, Sumenep, Bangkalan dan Sampang. Hasil klaster terbaik untuk prediksi penetapan daerah tertinggal tahun 2020 terdiri dari 4 klaster dengan nilai $ratio\ sw$ dan sb terkecil sebesar 0,3873984 dan nilai *threshold* optimum sebesar 0,04.

Kata Kunci : daerah tertinggal, indikator KPDT, *ensemble* ROCK

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

ABSTRACT

EVALUATION OF UNDERDEVELOPED AREAS IN EAST JAVA BASED ON INDICATORS OF STATE MINISTRY KPDT USING ENSEMBLE ROBUST CLUSTERING USING LINK (ROCK)

East Java is one of the provinces that has problems in terms of equitable development. Various problems of inequality have caused East Java Province to become one of the 21 provinces left behind in Indonesia. This is stated in the decision contained in Presidential Regulation No. 131 of 2015 which shows that East Java Province is the only Province in Java that is left behind. Determination of uderdeveloped areas is carried out for 5 years using 6 criteria and 33 indicators that have been determined by KPDT. However, several studies have been conducted on the determination of underdeveloped areas, but no studies that have applied 33 of these indicators as a whole. So in this study the determination of underdeveloped areas will be carried out using 33 criteria set by KPDT. The KPDT criteria data used are the results of the 2014 and 2018 surveys. The results of the 2014 survey data will be used as a comparison with the Government's decision, and the results of determination with survey data in 2018 will be used as predictions for the designation of 2020 regions. The results of this study are expected to provide an overview of the regions designated as lagging regions for the future. In this research, the Ensemble ROCK method is used, because this method has better accuracy compared to the agglomerative hierarchy method with good scalability. The results of the cluster determination of underdeveloped areas in 2015 consisted of Trenggalek, Jember, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Bangkalan, Sampang, Pamekasan, Sumenep, and Kota Probolinggo. The best cluster results consist of 4 clusters with the smallest sw and $sb\ ratio$ values of $0{,}3917473$ and the optimum Threshold value of $0{,}02$. As for the prediction of the determination of underdeveloped areas for 2020 consists of Trenggalek, Bondowoso, Situbondo, Probolinggo, Tuban, Pamekasan, Sumenep, Bangkalan and Sampang. The best cluster results for the prediction of the determination of underdeveloped areas in 2020 consists of 4 clusters with the smallest sw and sb ratio of $0{,}3873984$ and the optimum Threshold value of $0{,}04$.

Keywords: underdeveloped areas, indicators of KPDT, Ensemble ROCK

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR ISI

HALAMAN JUDUL i

LEMBAR PERSETUJUAN PEMBIMBING ii

PENGESAHAN TIM PENGUJI SKRIPSI iii

PERNYATAAN KEASLIAN iv

MOTTO v

HALAMAN PERSEMBAHAN vi

KATA PENGANTAR vii

ABSTRAK ix

ABSTRACT x

DAFTAR ISI xi

DAFTAR TABEL xiii

DAFTAR GAMBAR xvii

DAFTAR LAMPIRAN xviii

BAB 1 PENDAHULUAN 1

A. Latar Belakang 1

B. Rumusan Masalah 6

C. Tujuan Penelitian 7

D. Manfaat Penelitian 7

E. Batasan Masalah 8

BAB II KAJIAN PUSTAKA 10

A. Gambaran Umum Provinsi Jawa Timur 10

B. Evaluasi Daerah Tertinggal 12

C. Kriteria dan Indikator Daerah tertinggal 15

D. Jenis Data 25

E. Analisis Data 25

F. Analisis Faktor 27

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

G. Analisis Kelompok (Klaster) 35
H. Metode dan Validasi Analisis *Clustering* 37
I. *Ensemble Clustering* 43
J. *Robust Clustering Using Links* (ROCK) 45
K. Validitas Pengelompokan Metode ROCK 48
L. Pengujian Signifikansi Hasil Klaster 49
BAB III METODE PENELITIAN 54
A. Sumber Data 54
B. Variabel Penelitian 54
C. Tahap Analisis Data Penelitian 57
BAB V HASIL DAN PEMBAHASAN 62
A. Analisis Faktor Data Numerik 62
B. Analisis Klaster Data Numerik Menggunakan *K-Means* 65
C. Analisis Faktor Data Kategorik 81
D. Analisis Klaster Data Kategorik Menggunakan *K-Modes* 84
E. Analisis Klaster Data Campuran Menggunakan ROCK 94
F. Karakteristik Hasil Klaster Metode ROCK 104
G. Evaluasi Penetapan Daerah Tertinggal Tahun 2015 110
H. Prediksi Penetapan Daerah Tertinggal Tahun 2020 114
I. Evaluasi Prediksi Penetapan Daerah Tertinggal Tahun 2020 123
BAB V PENUTUP 126
A. Simpulan 126
B. Saran 127
DAFTAR PUSTAKA 129
LAMPIRAN

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR TABEL

Tabel 2.1 Wilayah Administratif Provinsi Jawa Timur Beserta luas wilayah, jumlah Kecamatan dan jumlah Kelurahan/Desa 11

Tabel 2.2 Gambaran Data *Multivariate* 26

Tabel 2.3 Kriteria Kesesuaian Penggunaan Analisis Faktor Menggunakan Nilai KMO 32

Tabel 2.3 Tabel Perhitungan Uji Barlett 33

Tabel 2.4 Distribusi *wilk's lambda* 51

Tabel 3.1 Variabel Alternatif Penelitian 55

Tabel 3.2 Variabel Kriteria Data Numerik 55

Tabel 3.3 Variabel Kriteria Data Kategorik 57

Tabel 4.1 Hasil uji KMO dan Uji Barlett 63

Tabel 4.2 Hasil *factoring* menggunakan PCA 63

Tabel 4.3 Nilai *Loading factor* masing-masing variabel 64

Tabel 4.4 Delapan Variabel Terpilih 65

Tabel 4.5 Data sampel perhitungan manual *K-Means* 66

Tabel 4.6 *Centroid* iterasi ke-1 66

Tabel 4.7 Hasil perhitungan jarak *euclidean* dan hasil klaster iterasi ke-1 68

Tabel 4.8 Pengelompokan data berdasarkan klaster 1 69

Tabel 4.9 Pengelompokan data berdasarkan klaster 2 69

Tabel 4.10 Pengelompokan data berdasarkan klaster 3 70

Tabel 4.11 Pengelompokan data berdasarkan klaster 4 70

Tabel 4.12 *Centroid* untul iterasi ke-2 71

Tabel 4.13 Hasil perhitungan jarak *euclidean* dan hasil klaster iterasi ke-2 71

Tabel 4.14 Perbandingan hasil klaster iterasi ke-1 dan iterasi ke-2 72

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.15 *Centroid* iterasi ke-3 73

Tabel 4.16 Hasil perhitungan jarak *euclidean* dan hasil klaster iterasi ke-3 73

Tabel 4.17 Perbandingan hasil klaster iterasi ke-2 dan iterasi ke-3 74

Tabel 4.18 *Centroid* iterasi ke-4 75

Tabel 4.19 Hasil perhitungan jarak *euclidean* dan hasil klaster iterasi ke-4 75

Tabel 4.20 Perbandingan hasil klaster iterasi ke-3 dan iterasi ke-4 76

Tabel 4.21 Hasil klaster 2 kelompok menggunakan *K-Means* 79

Tabel 4.22 Hasil klaster 3 kelompok menggunakan *K-Means* 80

Tabel 4.23 Hasil klaster 4 kelompok menggunakan *K-Means* 80

Tabel 4.24 Perbandingan nilai *R-Square* dan *Pseudo F-Statistic* masing-masing pengelompokan 81

Tabel 4.25 Hasil Uji Asumsi KMO dan *Barlett Test* 82

Tabel 2.26 Hasil *Factoring* Menggunakan PCA 82

Tabel 4.27 Nilai *Loading factor* masing-masing variabel 83

Tabel 4.28 Tiga Variabel Terpilih 83

Tabel 4.29 Data sampel perhitungan manual *K-Modes* 84

Tabel 4.30 *Centroid* iterasi ke-1 85

Tabel 4.31 Hasil perhitungan jarak dan hasil klaster iterasi ke-1 86

Tabel 4.32 Pengelompokan data dan modus klaster berdasarkan klaster 1 86

Tabel 4.33 Pengelompokan data dan modus klaster berdasarkan klaster 3 87

Tabel 4.34 Pengelompokan data dan modus klaster berdasarkan klaster 4 87

Tabel 4.35 *Centroid* iterasi ke-2 87

Tabel 4.36 Hasil perhitungan jarak dan hasil klaster iterasi ke-2 88

Tabel 4.37 Perbandingan hasil klaster iterasi ke-1 dan iterasi ke-2 89

Tabel 4.38 *Centroid* iterasi ke-3 89

Tabel 4.39 Hasil perhitungan jarak dan hasil klaster iterasi ke-3 90

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.40 Perbandingan hasil klaster iterasi ke-2 dan iterasi ke-3 91

Tabel 4.41 Perhitungan nilai r 92

Tabel 4.42 Hasil klaster 2 kelompok menggunakan *K-Modes* 92

Tabel 4.43 Hasil klaster 3 kelompok menggunakan *K-Modes* 93

Tabel 4.44 Hasil klaster 4 kelompok menggunakan *K-Modes* 93

Tabel 4.45 Perbandingan nilai r masing-masing pengelompokan..... 94

Tabel 4.46 Struktur data campuran 95

Tabel 4.47 Nilai $ratio\ sw$ dan sb dari masing-masing *Threshold* untuk pengelompokan 3 dan 4 klaster 95

Tabel 4.48 Data sampel perhitungan manual ROCK..... 96

Tabel 4.49 Hasil klaster metode ROCK ($\theta = 0{,}02$) 104

Tabel 4.50 Hasil Uji *Kruskal Wallis* 105

Tabel 4.51 Statistik Deskripstif (Rata-rata) Indikator Skala Numerik Anggota Tiap Klaster..... 105

Tabel 4.52 Statistik Deskriptif (Modus) Indikator Skala Kategorik Anggota Tiap Klaster..... 107

Tabel 4.53 Anggota klaster daerah tertinggal, berkembang, mandiri dan maju . 109

Tabel 4.54 Hasil analisis faktor data numerik..... 114

Tabel 4.55 Hasil analisis faktor data kategorik 114

Tabel 4.56 Perbandingan nilai *R-Square* dan *Pseudo F-Statistic* masing-masing pengelompokan..... 115

Tabel 4.57 Hasil klaster 4 kelompok menggunakan *K-Means* 115

Tabel 4.58 Perbandingan nilai r masing-masing pengelompokan..... 116

Tabel 4.59 Hasil klaster 4 kelompok menggunakan *K-Modes* 116

Tabel 4.61 Nilai $ratio\ sw$ dan sb dari masing-masing *Threshold* untuk pengelompokan 3 dan 4 klaster 117

Tabel 4.62 Hasil klaster metode ROCK ($\theta = 0{,}04$) 117

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.63 Hasil Uji Kruskal Wallis 118

Tabel 4.64 Statistik Deskriptif (Rata-rata) Indikator Skala Numerik Anggota Tiap Klaster 118

Tabel 4.65 Statistik Deskriptif (Modus) Indikator Skala Kategorik Anggota Tiap Klaster 120

Tabel 4.66 Anggota klaster daerah tertinggal, berkembang, mandiri dan maju . 122

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR GAMBAR

Gambar 2.1 Peta Provinsi JawaTimur 12

Gambar 2.2 Bagan Kriteria dan Indikator Daerah Tertinggal 24

Gambar 2.3 Tahapan umum pengelompokan Ensemble 44

Gambar 2.4 Tahapan Algoritma CEBMDC 44

Gambar 3.1 Diagram Alir Tahapan Analisis Data 58

Gambar 3.2 TahapanAnalisis Clustering Menggunakan K-Means 59

Gambar 4.1 Pemetaan Daerah Tertinggal Provinsi Jawa Timur Tahun 2015..... 110

Gambar 4.3 Grafik 5 Peringkat teratas Kabupaten/Kota Miskin di Jawa Timur 124

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR LAMPIRAN

Lampiran 1 Data Penelitian.....I-1

Lampiran 2 Syntax Program R-Studio..... II-1

Lampiran 3 Output Program Syntax R-Studio..... III-1

Lampiran 4 Hasil Perhitungan Akurasi Nilai r Hasil Pengelompokan K-Modes 2,3 dan 4 Klaster.....IV-1

Lampiran 5 Hasil Pengelompokan Penetapan Daerah Tertinggal Tahun 2015 Menggunakan Metode Ensemble ROCK V-1

Lampiran 6 Hasil Pengelompokan Prediksi Penetapan Daerah Tertinggal Tahun 2020 Menggunakan Metode Ensemble ROCK VI-1

Lampiran 7 Output SPSS Analisis Faktor Data Tahun 2018.....VII-1

Lampiran 8 Output SPSS Uji Signifikansi Hasil Klaster..... VIII-1

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB I
PENDAHULUAN

A. Latar Belakang

Kesenjangan dan ketimpangan yang terjadi di beberapa daerah terus menjadi permasalahan krusial yang perlu untuk diatasi. Hal tersebut sejalan dengan tujuan pembangunan nasional yang tercantum dalam pembukaan UUD 1945 Alenia ke-4 yaitu memajukan kesejahteraan umum. Kesenjangan dan ketimpangan yang terjadi di beberapa daerah disebabkan oleh permasalahan yang cukup beragam, diantaranya adalah tingkat kemajuan ekonomi, letak geografis, aspek sosial budaya, kualitas sumber daya manusia, ketersediaan sumber daya alam dan lain sebagainya.

Berdasarkan Peraturan Pemerintah nomor 78 tahun 2014 pasal 1 ayat 3, daerah tertinggal adalah daerah Kabupaten/Kota yang wilayah dan masyarakatnya sebagian besar cenderung kurang berkembang jika dibandingkan dengan daerah lain dalam skala nasional (Republik Indonesia, 2014). Adanya beberapa daerah yang memiliki tingkat perkembangan tertinggal tersebut menunjukkan terjadinya kesenjangan dalam hal pembangunan. Pembangunan daerah tertinggal adalah bentuk dari upaya Pemerintah dalam mengatasi berbagai permasalahan yang terjadi di suatu daerah mulai dari permasalahan ekonomi, sosial, lingkungan maupun keterbatasan fisik, sehingga mampu menjadikan daerah tersebut memiliki kualitas hidup yang tidak tertinggal jauh dengan daerah lain yang ada di Indonesia. Pembangunan daerah tertinggal sangat perlu untuk diperhatikan dengan tujuan terciptanya pembangunan yang merata. Seperti tujuan dari

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

pembangunan daerah tertinggal yang tercantum dalam Peraturan Pemerintah nomor 78 tahun 2014 Pasal 2 ayat 1 yaitu mengurangi kesenjangan antar wilayah daerah untuk menjamin terwujudnya keadilan pemerataan dari pembangunan nasional.

Jawa Timur merupakan salah satu Provinsi yang memiliki permasalahan dalam hal pemerataan pembangunan. Berbagai permasalahan ketimpangan dan kesenjangan menyebabkan Provinsi Jawa Timur menjadi salah satu dari 21 Provinsi yang berstatus tertinggal di Indonesia. Hal ini tercantum dalam keputusan yang terdapat dalam Peraturan Presiden Nomor 131 Tahun 2015 yang menunjukkan bahwa Provinsi Jawa Timur menjadi satu-satunya Provinsi di Pulau Jawa yang berstatus tertinggal (Republik Indonesia, 2015). Berdasarkan pernyataan tersebut, perlu dilakukan pengkajian khusus untuk mengatasi ketimpangan dan ketidakmerataan pembangunan di Jawa timur.

Peraturan Pemerintah nomor 78 tahun 2014 pasal 6 ayat 1 menyebutkan bahwa penetapan daerah tertinggal dilakukan selama 5 tahun sekali berdasarkan kriteria dan indikator yang telah ditetapkan oleh KPDT dengan sumber data yang didapatkan dari Badan Pusat Statistika berupa data potensi desa, statistik kesejahteraan rakyat dan profil masing-masing provinsi dalam angka tahun tertentu (Republik Indonesia, 2014). Dalam kasus ini, daerah tertinggal terakhir kali ditetapkan pada tahun 2015 yang tercantum dalam Peraturan Presiden nomor 131 tahun 2015 dan akan dilakukan penetapan kembali pada tahun 2020 mendatang. Dalam penelitian ini data kriteria KPDT yang digunakan adalah hasil survey tahun 2014 dan tahun 2018. Hasil penetapan dengan data survey tahun

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2014 akan digunakan sebagai pembanding dengan keputusan Pemerintah, dan hasil penetapan dengan data survey tahun 2018 akan digunakan sebagai prediksi penetapan daerah tertinggal tahun 2020 mendatang. Sehingga hasil dari penelitian ini dapat memberikan gambaran terkait daerah yang ditetapkan sebagai daerah tertinggal untuk kedepannya. Dengan demikian, pemerintah Kabupaten/Kota dapat mengambil kebijakan terhadap daerahnya yang disesuaikan dengan karakteristik masing-masing daerah dengan tujuan untuk mengentaskan daerah tersebut dari status tertinggal.

Beberapa penelitian terkait evaluasi daerah tertinggal telah banyak dilakukan, diantaranya yaitu Anik Djuridah dalam penelitiannya yang mengevaluasi status daerah tertinggal menggunakan analisis *Diskriminan* (Djuraidah, 2009) Dalam penelitian tersebut, peneliti hanya menentukan beberapa indikator yang berpengaruh dalam penentuan status tertinggal dari suatu daerah. Tanpa diketahui secara pasti daerah mana yang masuk kedalam kelompok daerah tertinggal dan tidak. Sama halnya dengan penelitian yang dilakukan oleh Satria, Herman dan Fajar yang menganalisis pembangunan wilayah tertinggal di Jawa Timur menggunakan analisis *Location Quotient dan Shift Share Esteban Marquillas*(Wiratama, Diartho, & Prianto, 2018). Dalam penelitian tersebut, peneliti hanya menggunakan variabel PDRB saja. Selanjutnya Dwi Hariadi Setiadi dalam tugas akhirnya yang melakukan pemetaan Kabupaten/Kota daerah tertinggal menggunakaan Metode *ensemble Simmilarity Weight And Filter Method* (SWFM) dan *Robust Clustering Using Link* (ROCK) (Setiadi, 2018). Dalam penelitian tersebut, peneliti hanya menggunakan 5 kriteria dan 13

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

indikator. Lima indikator tersebut yaitu infrastruktur, karakteristik daerah, perekonomian, sumber daya manusia (SDM), dan kemampuan keuangan daerah, tanpa menyertakan kriteria aksesibilitas. Sedangkan dalam Peraturan Pemerintah Kementrian Pembangunan Daerah Tertinggal (KPDT) yang tercantum dalam UU No 78 Tahun 2014 dan dijelaskan dalam Peraturan Presiden Nomor 131 Tahun 2015 pasal 2 ayat 1, 2, dan 3 yang menyebutkan bahwa penetapan daerah tertinggal menggunakan 6 kriteria (perekonomian masyarakat, sumber daya manusia, sarana dan prasarana, kemampuan keuangan daerah, aksesibiltas dan karakteristik daerah) yang terdiri dari 33 indikator yang digunakan untuk menentukan status daerah tertinggal. Berdasarkan beberapa penelitian terkait yang telah disebutkan diatas, belum ada penelitian yang melakukan evaluasi daerah tertinggal menggunakan semua kriteria dan indikator yang telah ditetapkan oleh Kementerian Negara Pembangunan Daerah Tertinggal (KPDT) secara keseluruhan. Sehingga dalam penelitian ini akan dilakukan evaluasi daerah tertinggal berdasarkan semua kriteria dan indikator yang telah ditetapkan oleh KPDT.

Dalam praktiknya, Pemerintah melakukan penetapan daerah tertinggal didasarkan pada Peraturan Presiden Nomor 131 Tahun 2015 Pasal 6 Ayat 2 yaitu dengan menggunakan aspek komposit dan nilai selang (*range*) (Republik Indonesia, 2015). Secara statistik kedua metode tersebut hanya cocok untuk melakukan analisis data yang bersifat numerik. Sedangkan pada kenyataannya, indikator untuk menentukan status daerah tertinggal tidak hanya mengacu kepada data yang bersifat numerik saja. Melainkan terdapat beberapa indikator yang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

bersifat kategorik. Sehingga, jika aspek komposit dan nilai selang yang digunakan sebagai analisis, maka tidak akan mampu mengakomodasi 6 kriteria yang terdiri dari 33 indikator tersebut.

Dalam mengatasi hal tersebut, diperlukan suatu metode khusus yang mampu mengakomodasi semua jenis data, baik kategorik maupun numerik. Metode statistik yang dapat digunakan untuk melakukan *clustering* pada data campuran adalah metode *ensemble*. Dalam penelitian ini, metode *ensemble* yang digunakan adalah *Robust Clustering Using Link* (ROCK). Metode *ensemble* ROCK merupakan metode *clustering* yang menggunakan konsep jarak untuk mengukur kesamaan atau kedekatan antara sepasang titik data (Sudipo Guha, Rajeev Rastogi, 1999). Kelebihan metode ROCK adalah memiliki akurasi yang lebih baik jika dibandingkan dengan metode hierarki *agglomerative* (pemusatan) dengan sifat skalabilitas yang baik (Reddy & B.Kavitha, 2012).

Penggunaan metode *ensemble* ROCK telah terbukti optimal untuk melakukan *clustering* data campuran (*ensemble*) dalam menyelesaikan berbagai kasus. Seperti penelitian yang dilakukan Shashi Sharma dan Ram Lal Yadav, dalam penelitian tersebut terbukti bahwa metode ROCK lebih optimal jika dibandingkan dengan metode *K-Means* untuk proses analisis klaster (Shashi Sharma & Yadav, 2013). Sama halnya dengan penelitian yang dilakukan oleh Dwi Harid Setiadi, dalam penerapan metode *ensemble* ROCK untuk pemetaan daerah tertinggal, terbukti lebih optimal jika dibandingkan dengan metode *ensemble* SWFM (Setiadi, 2018). Kemudian Alvionita dalam thesisnya yang membandingkan metode SWFM dan ROCK untuk pengelompokan aksesi jeruk.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Dalam penelitian tersebut, diketahui bahwa metode ROCK memiliki kinerja pengelompokan yang lebih baik daripada metode SWFM (Alvionita, Sutikno, & Suharsono, 2017).

Dalam penelitian ini, peneliti akan melakukan evaluasi daerah tertinggal di Jawa Timur berdasarkan indikator Kementerian Negara Pembangunan Daerah Tertinggal (KPDT) menggunakan *ensemble Robust Clustering Using Link* (ROCK). Dengan demikian, peneliti mengambil penelitian dengan judul "Evaluasi Daerah Tertinggal di Jawa Timur Berdasarkan Indikator Kementerian Negara Pembangunan Daerah Tertinggal (KPDT) Menggunakan *ensemble Robust Clustering Using Link (ROCK)*".

B. Rumusan Masalah

Berdasarkan latar belakang masalah yang telah dijelaskan pada sub bab A, pokok permasalahan yang akan dikaji dalam penelitian ini adalah sebagai berikut.

1. Bagaimana hasil Klaster Kabupaten/Kota Tertinggal di Jawa Timur berdasarkan Indikator Kementerian Negara Pembangunan Daerah Tertinggal (KPDT) menggunakan Metode *ensemble* ROCK ?
2. Bagaimana hasil validasi kinerja pengelompokan hasil klaster menggunakan Metode *ensemble* ROCK ?
3. Bagaimana karakteristik dari masing-masing daerah yang terindikasi tertinggal berdasarkan Indikator Kementerian Negara Pembangunan Daerah Tertinggal (KPDT) ?

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

C. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijelaskan pada sub bab B, tujuan dari penelitian ini adalah sebagai berikut.

1. Mengetahui hasil evaluasi daerah Kabupaten/Kota Tertinggal di Jawa Timur berdasarkan Indikator Kementerian Negara Pembangunan Daerah Tertinggal (KPDT) menggunakan metode *ensemble* ROCK.
2. Mengetahui validasi kinerja pengelompokan hasil evaluasi menggunakan metode *ensemble* ROCK.
3. Mengetahui karakteristik dari masing-masing daerah yang terindikasi tertinggal berdasarkan Indikator Kementerian Negara Pembangunan Daerah Tertinggal (KPDT).

D. Manfaat Penelitian

Manfaat yang akan didapatkan dalam penelitian ini terdiri dibagi menjadi dua, yaitu manfaat secara teoritis dan secara praktik.

Secara teoritis, penelitian ini bermanfaat untuk memberikan tambahan pengetahuan dan informasi terkait implementasi metode *ensemble* ROCK sebagai evaluasi daerah tertinggal Kabupaten/Kota di Jawa Timur. Sehingga, dapat menambah *literature* untuk penelitian selanjutnya.

Secara praktik, penelitian ini bermanfaat untuk memberikan gambaran bagi Pemerintah Kabupaten/Kota di Jawa Timur terkait karakteristik daerahnya masing-masing. Hasil penelitian ini mampu membantu Pemerintah

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Kabupaten/Kota dalam menentukan kebijakan sebagai bahan pertimbangan dalam mengatasi daerah tertinggal yang ada di wilayahnya masing-masing.

E. Batasan Masalah

Pembahasan permasalahan dalam penelitian ini dibatasi oleh beberapa hal sebagai berikut :

1. Evaluasi daerah tertinggal Kabupaten/Kota hanya dilakukan untuk Provinsi Jawa Timur yang terdiri dari 38 daerah Kabupaten/Kota.
2. Evaluasi daerah Kabupaten/Kota Tertinggal di Jawa Timur ditinjau dari enam kriteria yaitu perekonomian masyarakat, sumber daya manusia, aksesibilitas, karakteristik daerah, sarana prasarana serta kemampuan keuangan daerah. Enam kriteria tersebut terdiri dari 33 indikator dengan memperhatikan beberapa karakteristik dari daerah tertentu.
3. Data yang digunakan adalah data Potensi Desa tahun 2014 dan 2018, Statistik Kesejahteraan Rakyat tahun 2014 dan 2018, dan Profil Provinsi Jawa Timur dalam Angka 2014 dan 2018, Badan Perencanaan dan Pembangunan Daerah dan Data Informasi Perkembangan Daerah tertentu (Daerah Rawan Pangan) tahun 2014 dan 2018.
4. Indikator jumlah desa dengan akses ke pelayanan kesehatan $> 5km$ tidak disertakan, karena data tersebut tidak tersedia dalam potensi desa.
5. Metode yang digunakan untuk menyelesaikan permasalahan dalam penelitian ini adalah metode *K-Means* untuk analisis klaster data numerik,

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

metode *K-Modes* untuk analisis klaster data kategorik dan metode *ensemble* ROCK untuk analisis data campuran .

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB II

KAJIAN PUSTAKA

A. Gambaran Umum Provinsi Jawa Timur

Berdasarkan Badan Pembangunan dan Pendapatan Daerah, Provinsi Jawa Timur memiliki wilayah seluas $47.157{,}72\ km^2$ yang terdiri dari daratan seluas $47.130{,}15\ km^3$ dan lautan seluas $110.764{,}28\ km^2$. Provinsi ini terletak diantara $111{,}0'BT$ hingga $114{,}4'BT$ serta terletak pada Garis Lintang $7{,}12''LS$ dan $8{,}48'LS$ (BAPPEDA, 2018). Provinsi Jawa Timur memiliki lokasi di sekitar garis khatulistiwa, sehingga Jawa Timur memiliki dua perubahan musim (kemarau dan penghujan) seperti wilayah Provinsi lainnya yang ada di Indonesia. Wilayah Jawa Timur berbatasan dengan Pulau Bali di sebelah Timur.Sebelah Barat berbatasan langsung dengan Provinsi Jawa Tengah dan sebelah Utara berbatasan langsung dengan Provinsi Kalimantan Selatan. Serta sebelah Selatan berbatasan dengan Samudera Hindia. Provinsi Jawa Timur terbagi menjadi dua bagian besar. Diantaranya yaitu Jawa Timur bagian daratan dan Pulau Madura.

Provinsi Jawa Timur terdiri dari 38 Kabupaten/Kota (29 Kabupaten dan 9 Kota), 664 Kecamatan, 777 Kelurahan serta 7724 desa (BAPPEDA, 2018). Luas wilayah Kabupaten terbesar dimiliki oleh Banyuwangi, yang memiliki luas 5.782,40 Km^2 (BPS Kabupaten Banyuwangi, 2018). Sementara Kabupaten yang memiliki wilayah terkecil adalah Kota Mojokerto, dengan luas sebesar 16,47 Km^2 (BPS Kabupaten Mojokerto, 2018). Selanjutnya, Kabupaten yang memiliki jumlah kecamatan terbesar adalah Kabupaten Malang yang terdiri dari 33

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

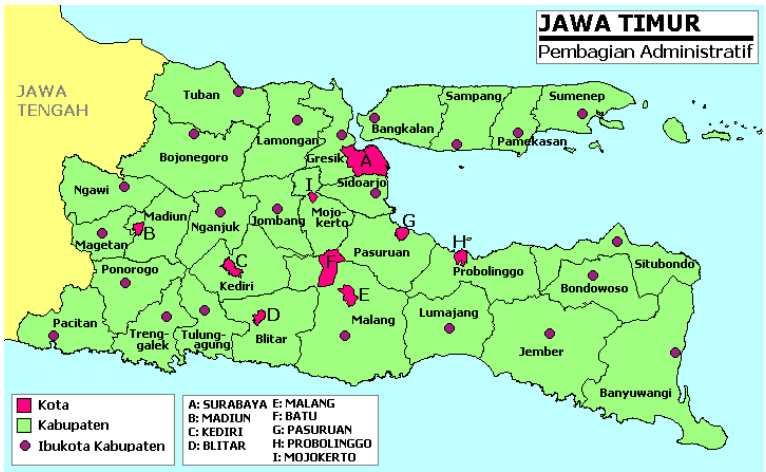
Kecamatan (BPS Kabupaten Malang, 2018) dan Kabupaten yang memiliki jumlah Kelurahan/Desa terbesar adalah Kabupaten Lamongan yang terdiri dari 474 Kelurahan/Desa (BPS Kabupaten Lamongan, 2018). Tabel 2.1 menunjukkan wilayah administratif Provinsi Jawa Timur beserta luas wilayah, jumlah Kecamatan dan jumlah Kelurahan/Desa (BAPPEDA, 2018).

Tabel 2.1 Wilayah Administratif Provinsi Jawa Timur Beserta luas wilayah, jumlah Kecamatan dan jumlah Kelurahan/Desa

No	Kabupaten/Kota	Luas Wilayah	Jumlah Kecamatan	Jumlah Kelurahan/Desa
01	Kab.Pacitan	1389,92 Km^2	12	171
02	Kab.Ponorogo	1305,70 Km^2	21	307
03	Kab. Trenggalek	1147,22 Km^2	14	157
04	Kab.Tulungagung	1055,65 Km^2	19	271
05	Kab.Blitar	1336,48 Km^2	22	248
06	Kab.Kediri	1386,05 Km^2	26	344
07	Kab.Malang	3530,65 Km^2	33	390
08	Kab.Lumajang	1790,90 Km^2	21	205
09	Kab. Jember	3092,34 Km^2	31	248
10	Kab. Banyuwangi	5782,40 Km^2	24	217
11	Kab. Bondowoso	1525,97 Km^2	23	219
12	Kab.Situbondo	1669,21 Km^2	17	136
13	Kab. Probolinggo	1696,21 Km^2	24	330
14	Kab.Pasuruan	1474,02 Km^2	24	365
15	Kab. Sidoarjo	634,38 Km^2	18	353
16	Kab.Mojokerto	717,83 Km^2	18	304
17	Kab.Jombang	1115,09 Km^2	21	306
18	Kab.Nganjuk	1224,25 Km^2	20	284
19	Kab.Madiun	1037,58 Km^2	15	206
20	Kab.Magetan	688,84 Km^2	18	235
21	Kab.Ngawi	1295,98 Km^2	19	217
22	Kab.Bojonegoro	2198,79 Km^2	28	430
23	Kab. Tuban	1834,15 Km^2	20	328
24	Kab. Lamongan	1782,05 Km^2	27	474
25	Kab. Gresik	1191,25 Km^2	18	356
26	Kab.Bangkalan	1233,08 Km^2	18	281
27	Kab. Sampang	792,24 Km^2	14	186
28	Kab.Pamekasan	1998,54 Km^2	13	189
29	Kab. Sumenep	1001,44 Km^2	27	334
30	Kota Kediri	63,40 Km^2	3	46
31	Kota Blitar	32,57 Km^2	3	21
32	Kota Malang	145,28 Km^2	5	57
33	Kota Probolinggo	56,67 Km^2	5	29
34	Kota Pasuruan	35,29 Km^2	4	34
35	Kota Mojokerto	16,47 Km^2	2	18
36	Kota Madiun	33,92 Km^2	3	27
37	Kota Surabaya	350,54 Km^2	31	154
38	Kota Batu	136,74 Km^2	3	24

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Pembagian wilayah administratif Provinsi Jawa Timur ditunjukkan pada Gambar 2.1.



Gambar 2.1 Peta Provinsi JawaTimur

(Sumber : https://www.Eastjava.com/Peta-Jawa-Timur-Lengkap-Dengan-Daftar-29-Nama-Kabupaten-dan-9-Kota.html)

B. Evaluasi Daerah Tertinggal

Evaluasi adalah bentuk dari upaya untuk melakukan pengukuran terhadap hasil serta dampak dari suatu proyek, aktivitas atau program tertentu dengan cara meninjau ulang tujuan yang telah ditetapkan sebelumnya. Menurut Calongesi (1995), evaluasi merupakan hasil dari sebuah keputusan berupa nilai yang didapatkan dari hasil pengukuran (Wulan, 2016).

Dalam Al-Qur'an evaluasi adalah hal yang selalu melekat dalam setiap kehidupan manusia. Hal tersebut karena dalam setiap aspek kehidupan, manusia melakukan beragam aktivitas. Setiap aktivitas yang dilakukan baik secara kelompok maupun individu, selalu membutuhkan evaluasi agar mampu memperbaiki kinerja yang telah dilakukan sebelumnya menjadi lebih baik lagi. Evaluasi merupakan komponen yang penting, karena mampu meningkatkan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kuantitas dan kualitas hidup. Konsep mengenai evaluasi dalam Al-Qur'an telah dijelaskan dengan berbagai makna dan kalimat yang bermacam-macam. Seperti yang tercantum dalam Q.S Al-Baqarah ayat 284.

لِلَّهِ مَا فِي السَّمَاوَاتِ وَمَا فِي الْأَرْضِ ۗ وَإِنْ تُبْدُوا مَا فِي أَنْفُسِكُمْ أَوْ تُخْفُوهُ
يُحَاسِبْكُمْ بِهِ اللَّهُ ۖ فَيَغْفِرُ لِمَنْ يَشَاءُ وَيُعَذِّبُ مَنْ يَشَاءُ ۗ وَاللَّهُ عَلَىٰ كُلِّ شَيْءٍ قَدِيرٌ

Artinya: :

"Kepunyaan Allah-lah segala apa yang ada di langit dan apa yang ada di bumi. Dan jika kamu menampakkan apa yang ada di dalam hatimu atau kamu menyembunyikan, niscaya Allah akan membuat perhitungan dengan kamu tentang perbuatanmu itu. Maka Allah mengampuni siapa yang dikehandakiNya dan menyiksa siapa yang dikehendakiNya; dan Allah Mahakuasa atas segala sesuatu." (Al-Baqarah: 284).

Dalam ayat Al Qur'an tersebut, diketahui bahwa Allah SWT akan senantiasa melakukan *hisab* atau perhitungan terhadap segala hal yang dilakukan oleh setiap makhluknya. Dengan demikian, Allah akan mengambil tindakan terhadap apa yang dilakukan oleh hamba-Nya sebagai bentuk pertanggung jawaban. Sama halnya dalam konteks evalusi dalam hal ini, setiap manusia dianjurkan untuk melakukan evaluasi diri dengan melakukan pengukuran menggunakan kriteria tertentu. Sehingga, akan mampu mendapatkan hasil yang dapat mempermudah dalam pengambilan keputusan untuk kedepannya. Dalam hal ini sejalan dengan pengevaluasian daerah tertinggal yang didasarkan pada kriteria yang telah ditetapkan oleh KPDT yang berorientasi untuk memberikan solusi dalam mengentaskan daerah yang memiliki status tertinggal.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Daerah tertinggal adalah daerah Kabupaten/Kota yang wilayah dan masyarakatnya sebagian besar cenderung kurang berkembang jika dibandingkan dengan daerah lain dalam skala nasional. Sehingga, dalam hal ini diperlukan suatu langkah yang tepat guna mengentaskan beberapa daerah terkait yang memiliki status sebagai daerah tertinggal tersebut. Sebagaimana Q.S An Nisa' ayat 9.

وَلْيَخْشَ الَّذِينَ لَوْ تَرَكُوا مِنْ خَلْفِهِمْ ذُرِّيَّةً ضِعَافًا خَافُوا عَلَيْهِمْ فَلْيَتَّقُوا اللَّهَ وَلْيَقُولُوا قَوْلًا سَدِيدًا

Artinya :

"Dan hendaklah takut kepada Allah orang-orang yang seandainya meninggalkan dibelakang mereka anak-anak yang lemah, yang mereka khawatir terhadap (kesejahteraan) mereka. Oleh sebab itu hendaklah mereka bertakwa kepada Allah dan hendaklah mereka mengucapkan perkataan yang benar." (QS:An-Nisaa Ayat: 9).

Ayat tersebut mengajarkan kita untuk peduli terhadap kaum yang lemah. Dengan demikian, akan tercipta kesejahteraan antar sesama. Seperti halnya konsep dari evaluasi daerah tertinggal. Hasil evaluasi ini, akan mampu membantu mengentaskan daerah yang berstatus tertinggal karena mampu memberikan masukan bagi Pemerintah untuk menentukan kebijakan yang akan diambil berdasarkan karakteristik dari daerahnya masing-masing.

Dalam konteks penelitian ini, evaluasi daerah tertinggal merupakan penilaian terhadap status ketertinggalan suatu daerah sebagai tolak ukur keberhasilan Pemerintah dalam upaya meningkatkan kesejahteraan daerah atau wilayahnya masing-masing. Penilaian ini didasarkan pada kriteria dan indikator

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

yang telah ditetapkan oleh KPDT dengan tujuan untuk mengurangi kesenjangan antar wilayah daerah guna untuk menjamin terwujudnya keadilan pemerataan dari pembangunan nasional. Selain itu, bertujuan pula untuk memenuhi berbagai sarana prasarana dan kebutuhan dasar yang dibutuhkan oleh daerah tertinggal tersebut.

C. Kriteria dan Indikator Daerah tertinggal

Menurut Peraturan Pemerintah Nomor 78 tahun 2014, penentuan daerah tertinggal ini dilakukan selama 5 tahun sekali dengan memperhatikan kriteria yang terdiri dari beberapa indikator dan karakteristik daerah tertentu yang tercantum dalam Peraturan Menteri (Republik Indonesia, 2014). Dalam proses identifikasi status ketertinggalan suatu daerah, dilaksanakan dengan melakukan pengukuran berdasarkan standar yang tercantum dalam Peraturan Menteri Desa, Pembangunan Daerah Tertinggal dan Transmigrasi Nomor 3 Tahun 2016 . Kriteria penentuan daerah tertinggal terdiri dari 6 kriteria yang terdiri dari 27 indikator dengan memperhatikan 6 indikator karakteristik daerah tertentu, sehingga total keseluruhan terdapat 33 indikator. Selanjutnya, penjelasan terkait kriteria dan masing-masing indikatornya akan dijelaskan dalam sub bab selanjutnya.

1. Perekonomian Masyarakat

Kriteria dari perekonomian masyarakat merupakan gambaran dari kondisi perekonomian penduduk dalam suatu wilayah tertentu. Perekonomian

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

masyarakat ini memiliki 2 indikator yang harus dipenuhi untuk melakukan identifikasi daerah tertinggal, yaitu :

a. Presentase Penduduk Miskin

Menurut Badan Pusat Statistika (BPS), penduduk miskin merupakan penduduk yang mempunyai rata-rata pengeluaran perkapita setiap bulannya dibawah garis kemiskinan. Dengan demikian, penduduk miskin adalah kelompok penduduk yang tidak memiliki kemampuan berbagai akses, seperti sosial, ekonomi, sarana dan prasarana.

b. Pengeluaran Konsumsi Per Kapita

Menurut (BPS) pengeluaran konsumsi per kapita merupakan biaya yang harus dikeluarkan untuk memenuhi kebutuhan konsumsi seluruh anggota keluarga dalam rumah tangga selama satu bulan dibagi dengan jumlah anggota keluarga. Data ini dapat menggambarkan pola konsumsi dari kelompok rumah tangga sehingga dapat dijadikan sebagai pengukuran tingkat kesejahteraan penduduk.

2. Sumber Daya Manusia

Menurut Nawawi (2003), sumber daya manusia meupakan seluruh penduduk suatu negara yang berada dalam batas wilayah tertentu yang telah memenuhi usia kerja baik yang sudah memiliki kerjaan atau belum(Farchan, Fauzi, 2016). Kriteria sumber daya manusia terdiri dari tiga indikator, yaitu :

a. Angka Harapan Hidup

Menurut (BPS) Angka harapan hidup merupakan perkiraan rata-rata banyaknya tahun yang dapat ditempuh oleh seseorang selama hidupnya.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Angka harapan hidup adalah rata-rata dari tahun hidup yang akan dijalani oleh individu tertentu yang telah mencapai umur dalam tahun tertentu. Perhitungan angka harapan hidup didasarkan pada angka kematian. Data ini didapatkan dari catatan registrasi kematian.

b. Rata-rata Lama Sekolah

Menurut (BPS) rata-rata lama sekolah merupakan jumlah tahun yang dihabiskan oleh penduduk usia 15 tahun keatas untuk menjalani pendidikan formal. Dalam menghitung indikator ini dibutuhkan beberapa informasi terkait. Diantaranya yaitu partisipasi sekolah, ijasah sekolah terakhir, kelas tertinggi yang sedang atau pernah diduduki dan jenis pendidikan yang sedang atau pernah diduduki. Penduduk yang tamat SD terhitung sekolah selama 6 tahun. Penduduk yang tamat SMP terhitung sekolah selama 9 tahun dan seterusnya. Perhitngan rata-rata lama sekolah mengabaikan apakah penduduk tersebut naik kelas atau tidak.

c. Angka Melek Huruf

Angka melek huruf atau sekarang yang sering dikenal dengan angka harapan lama sekolah merupakan jumlah tahun yang diharapkan untuk dapat dirasakan oleh anak dengan usia tertentu dimasa yang akan datang. Menurut (BPS), angka melek huruf (AMH) adalah proporsi antara penduduk yang berusia 15 tahun keatas yang memiliki kemampuan menulis dan membaca kalimat dalam bentuk huruf latin, arab dan lainnya dengan anak yang berusia 15 tahun keatas secara keseluruhan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

3. Infrastruktur

Infrastruktur merupakan pokok penting yang harus diperhatikan untuk mendukung keberhasilan dari suatu program. Dalam menentukan status daerah tertinggal diperlukan kriteria sarana dan prasarana. Kriteria ini ditinjau dari beberapa aspek, yaitu:

a. Banyaknya desa yang menggunakan transportasi darat berdasarkan jenis permukaan jalan terluas.

Dalam aspek ini, jenis permukaan jalan terluas terdiri dari aspal atau beton, diperkeras berupa kerikil atau batu, tanah dan lain-lain (kayu atau papan yang digunakan didaerah rawa, hutan dan sejenisnya). Indikator yang dinilai dalam aspek ini terdiri dari empat indikator, yaitu :

1) Persentase desa atau kelurahan dengan jenis permukaan jalan utama terluas aspal atau beton
2) Persentase desa atau kelurahan dengan jenis permukaan jalan utama terluas diperkeras.
3) Persentase desa atau kelurahan dengan jenis permukaan jalan utama terluas tanah.
4) Persentase desa atau kelurahan dengan jenis permukaan jalan utama terluas lainnya

b. Banyaknya desa atau kelurahan yang memiliki sekolah.

Penilaian terhadap aspek ini dinilai penting, karena dapat menggambarkan tingkat pendidikan dari penduduk di daerah tertentu.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Aspek ini hanya terdiri dari satu indikator yaitu Jumlah sekolah menengah pertama (SMA) per 1000 penduduk.

c. Banyakya desa atau kelurahan berdasarkan sarana kesehatan. Aspek ini terdiri dari dua indikator yaitu :

   1. Jumlah Sarana dan Prasarana Kesehatan Per 1000 Penduduk
   2. Jumlah Dokter Per 1000 Penduduk

d. Banyaknya desa atau kelurahan berdasarkan keberadaan pengguna listrik dan sumber penerangan jalan utama

Aspek ini terdiri dari satu indikator, yaitu Persentase rumah tangga pengguna listrik baik PLN maupun non PLN. Aspek selajutnya yaitu komunikasi. Dalam aspek ini indikator yang dibutuhkan yaitu persentase rumah tangga pengguna telepon.

Aspek selanjutnya yaitu lingkungan hidup. Dalam hal ini indikator yang digunakan yaitu persentase rumah tangga pengguna air minum bersih. Air minum yang layak adalah kualitas air minum yang terlindungi. Air jenis ini bersumber dari keran umum, terminal air, air ledeng, hydrant umum, penampungan air hujan, dan sumur. Selain itu, sumber air minum tersebut harus jauh dari pembuangan kotoran (minimal 10 m), pembuangan sampah dan limbah.

Indikator selanjutnya yaitu indikator persentase desa yang mempunyai pasar tanpa bangunan permanen atau semi permanen. Pasar merupakan tempat bertemunya penjual dan pembeli barang dan jasa. Penempatan pasar dapat menggunakan bangunan permanen atau semi

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

permanen, dan dapat juga tanpa bangunan. Pasar dengan bangunan permanen yaitu pasar yang bertempat dalam bangunan tetap yang terdiri dari lantai, dinding dan atap tetap. Selanjutnya pasar dengan bangunan semi permanen adalah pasar dengan bangunan tetap yang memiliki lantai dan atap namun tidak memiliki dinding. Sedangkan pasar tanpa bangunan merupakan pasar yang tidak berada dalam bangunan seperti pasar terapung, pasar subuh dan lain-lain.

4. Kemampuan Keuangan Daerah

Kriteria selanjutnya yaitu kemampuan keuangan daerah. Indikator untuk kriteria ini adalah desentralisasi fiskal. Perhitungan tingkat desentralisasi fiskal didasarkan dari rasio pendapatan asli daerah dengan total penerimaan daerah. Derajat desentralisasi fiskal dapat dihitung menggunakan persamaan sebagai berikut.

$$Desentralisasi\ fiskal = \frac{Pendapatan\ Asli\ Daerah}{Pendapatan\ Daerah} \times 100\%$$

5. Aksessibilitas

Kriteria aksessibiltas merupakan keriteria yang digunakan untuk memberikan penilaian terhadap derajat kemudahan yang dirasakan oleh individu dalam bentuk pelayanan atau objek tertentu. Kriteria aksesibilitas terdiri dari tiga indikator, yaitu :

a. Rata-Rata jarak dari Kabupaten/Kota ke Ibu Kota Provinsi
b. Rata-Rata jarak dari desa Ke Pusat Pelayanan Kesehatan
c. Jarak kemudahan akses ke pusat keamanan (pos polisi) terdekat

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

6. Karakteristik Daerah

Karakteristik daerah merupakan ciri fisik dari daerah yang jarang dimiliki oleh daerah lainnya. Karakteristik dari daerah tertentu ini terdiri dari tujuh indikator, yaitu :

a. Persentase desa terkena bencana gempa bumi

Persentase desa terkena bencana gempa bumi dapat dihitung menggunakan persamaan sebagai berikut:

$$Gempa = \frac{Banyaknya\ desa\ terkena\ gempa}{Jumlah\ seluruh\ desa} x100\%$$

b. Persentase desa terkena bencana tanah longsor

Persentase desa terkena bencana tanah longsor dapat dihitung menggunakan persamaan sebagai berikut:

$$Tanah\ longsor = \frac{Banyaknya\ desa\ terkena\ tanah\ longsor}{Jumlah\ seluruh\ desa} x100\%$$

c. Presentase desa terkena bencana banjir

Persentase desa terkena bencana banjir dapat dihitung menggunakan persamaan sebagai berikut:

$$Banjir = \frac{Banyaknya\ desa\ terkena\ banjir}{Jumlah\ seluruh\ desa} x100\%$$

d. Persentase desa terkena bencana bencana lainnya

Bencana lain dalam hal ini terdiri dari banjir bandang, tsunami dan gelombang pasang laut. Persentase desa terkena bencana bencana lainnya dapat dihitung menggunakan persamaan sebagai berikut :

$$Bencana\ lainnya = \frac{Banyaknya\ desa\ terkena\ bencana\ lain}{Jumlah\ seluruh\ desa} x100\%$$

e. Presentase kawasan hutan lindung

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Presentase kawasan hutan lindung dapat dihitung menggunakan persamaan sebagai berikut :

$$Kawasan\ hutan\ lindung = \frac{luas\ hutan\ lindung}{luas\ seluruh\ hutan} x100\%$$

f. Presentase desa yang melakukan pelestarian lahan kritis

Presentase desa yang melakukan pelestarian lahan kritis dapat dihitung menggunakan persamaan sebagai berikut :

$$Pelestarian\ lahan\ kritis = \frac{luas\ lahan\ kritis}{Jumlah\ desa} x100\%$$

g. Presentase desa yang memiliki konflik

Presentase desa yang mengalami konflik dalam satu tahun terakhir dapat dihitung menggunakan persamaan sebagai berikut:

$$Desa\ konflik = \frac{Banyaknya\ desa\ yang\ mengelami\ perkelahian\ masal}{Jumlah\ desa} x100\%$$

Selain kriteria dan indikator yang disebutkan diatas, harus diperhatikan pula karakteristik daerah tertentu. Diantaranya yaitu :

a. Presentase desa tertinggal

Presentase desa tertinggal merupakan presentase desa yang telah ditetapkan sebagai daerah tertinggal pada tahun penetapan sebelumnya. Presentase daerah tertinggal dapat dihitung menggunakan persamaan sebagai berikut.

$$Desa\ tertinggal\ = \frac{Jumlah\ desa\ tertinggal}{Jumlah\ seluruh\ desa} x100\%$$

b. Daerah perbatasan

Daerah perbatasan merupakan daerah yang berbatasan langsung dengan daerah lain yang bukan termasuk kedalam satu provinsi.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

c. Daerah rawan bencana

Daerah rawan bencana merupakan daerah yang memiliki resiko tinggi terkena bencana karena kondisi geografis, geologis ataupun karena ulah manusia itu sendiri(Badan Perencanaan Pembangunan Daerah Provinsi Jawa Timur, 2017). Bencana yang dimaksud disini terdiri dari tanah longsor, gempa bumi, banjir, tsunami dan gunung meletus.

d. Daerah rawan pangan

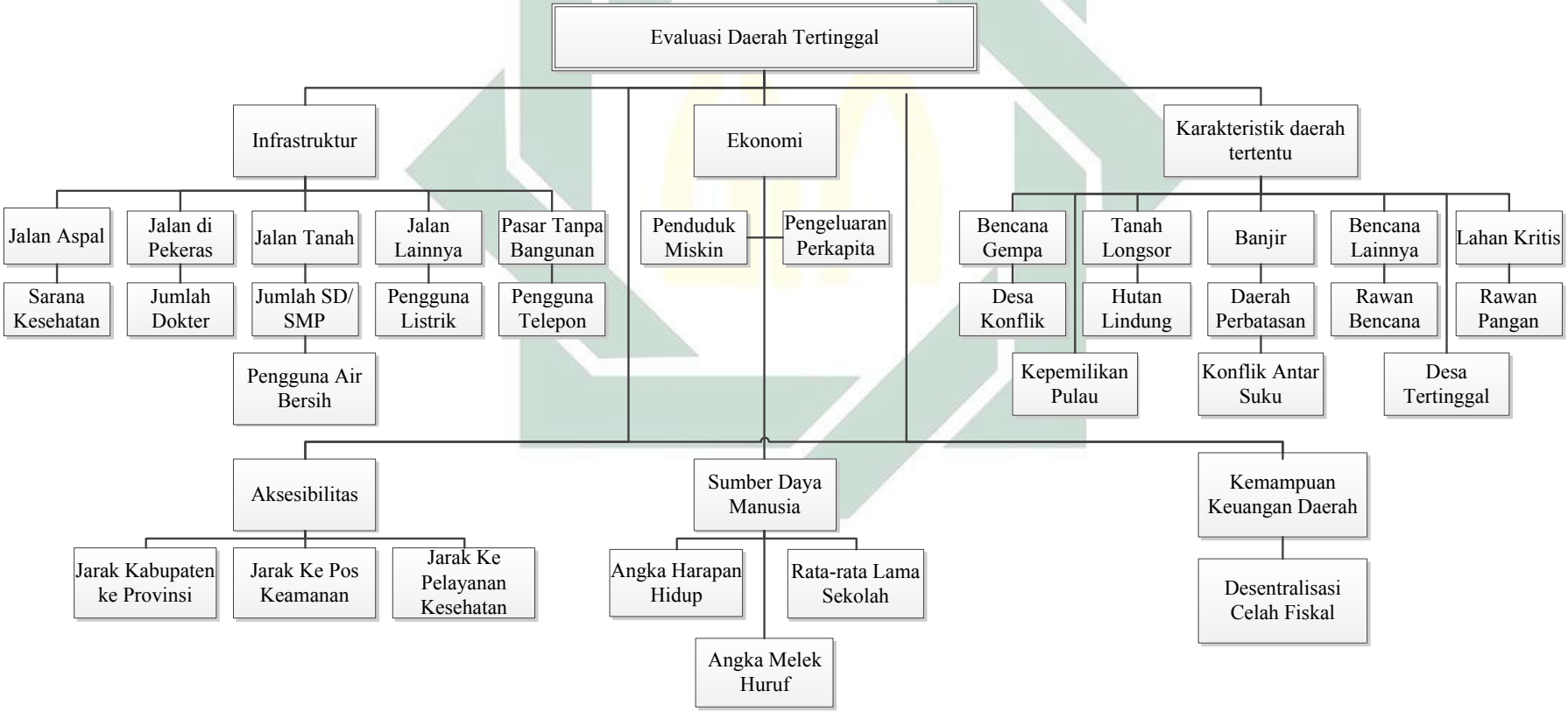
Rawan pangan merupakan kondisi kekurangan atau tidak tercukupinya kebutuhan pangan suatu kondisi ketidakcukupan pangan yang dialami oleh masyarakat dibeberapa daerah yang terjadi pada waktu tertentu sehingga tidak mampu untuk memenuhi kebutuhan pertumbuhan dan kesehatan masyarakat (Badan Penelitian dan Pengembangan Pendidikan dan Pelatiihan dan Informasi Kementrian Desa Pembangunan Daerah Tertinggal dan Transformasi, 2017).

e. Daerah pasca konflik antar suku

Konflik yang terjadi antar suku merupakan perkelahian yang terjadi antara dua pihak yang bersekutu atau tidak dengan alasan utama adalah tentang kesukuan.

Berdasarkan uraian terkait kriteria dan indikator evaluasi daerah tertinggal, yang telah disebutkan diatas, dapat dibuat bagan yang dapat membantu mempermudah proses analisis seperti yang terlihat pada Gambar (2.2).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 2.2 Bagan Kriteria dan Indikator Daerah Tertinggal

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

D. Jenis Data

Data adalah kumpulan dari fakta yang diolah sedemikian rupa sehingga mampu memberikan informasi bagi yang membutuhkannya. Berdasarkan jenis dari variabelnya, data dikelompokkan menjadi dua, yaitu data numerik dan kategorik. Data yang memiliki jenis variabel kuantitatif dapat digolongkan sebagai data numerik. Data jenis ini akan menghasilkan informasi numerik. Menurut Anderson dan Sclove (1974), data numerik dikelompokkan menjadi dua bagian, yaitu data diskrit (pencacahan) dan data kontinu (pengukuran)(Alvionita et al., 2017). Sedangkan data kategorik adalah data yang memiliki jenis variabel kualitatif. Data kategorik dapat dihasilkan melalui penggolongan, pemetaan maupun proses klasifikasi. Data kategorik mempunyai skala pengukuran yang terdapat dalam satu set kategori (Agresti, 2009). Misalnya penyakit kanker yang diukur berdasarkan kategori ganas dan jinak.

E. Analisis Data

Dalam proses analisis data, terdapat tiga jenis analisis data, yaitu analisis *unvariate, bivariate dan multivariate*. Penggunaan analisis data tersebut disesuaikan dengan jenis data yang digunakan. Analisis *unvariate* adalah analisis yang diterapkan dalam satu variabel saja (Setiawan & Muntaha, 2014), seperti analisis deskriptif statistik yang terdiri dari *mean*, median, modus, standard deviasi, nilai minimum, nilai maksimum dan lain sebagainya. Sehingga, analisis *unvariate* merupakan analisis yang dilakukan pada data tunggal yang tidak memiliki keterikatan dengan variabel lainnya. Sedangkan analisis data yang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

berkaitan dengan dua variabel disebut dengan analisis *bivariate* (David & Djamaris, 2018). Dalam analisis *bivariate*, lebih cenderung untuk mengetahui hubungan antar variabel bebas dan variabel terikat. Secara umum, analisis *bivariate* ini biasanya digunakan untuk mengetahui korelasi antar dua variabel. Dimana dua variabel tersebut ada yang berperan sebagai variabel *exogen* dan *endogen*. Sedangkan analisis *multivariate* adalah pengembangan dari analisis *unvariate* dan *bivariate*. Analisis ini digunakan untuk menangani lebih dari dua variabel (Haryono, 2017).

Analisis data yang digunakan dalam penelitian ini adalah analisis *multivariate*. Hal tersebut karena data yang digunakan dalam penelitian ini terdiri dari dua pengamatan lebih dengan beberapa karakteristik yang berbeda-beda. Dalam penelitian ini data pengamatan yang digunakan adalah seluruh Kabupaten/Kota di Jawa Timur (33 daerah) serta variabel yang terdiri dari indikator kriteria daerah tertingal (33 indikator). Gambaran terkait data *multivariate* dalam penelitian yang dilakukan dengan pengamatan sebanyak n serta variabel sebanyak i, disajikan dalam Tabel 2.2.

Tabel 2.2 Gambaran Data *Multivariate*

Pengamatan $ke-i$	Variabel 1 (x_1)	Variabel 2 (x_2)	Variabel 3 (x_3)	...	Variabel (x_i)
1	x_{11}	x_{21}	x_{31}	...	x_{i1}
2	x_{12}	x_{22}	x_{32}	...	x_{i2}
3	x_{13}	x_{23}	x_{33}	...	x_{i3}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$
n	x_{1n}	x_{2n}	x_{3n}	...	x_{in}

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

F. Analisis Faktor

1. Pengertian Analisis Faktor

Analisis faktor adalah jenis teknik analisis *multivariate*. Analisis ini bertujuan untuk mengetahui hubungan yang terjadi antara satu variabel dengan variabel lainnya. Teknik dari analisis ini dilakukan dengan menemukan hubungan diantara variabel-variabel yang awalnya bersifat saling *independent*, menjadi sekumpulan variabel baru yang memiliki korelasi kuat dan berjumlah lebih sedikit dari variabel semula (Simarmata, Arma, & Arnita, 2013). Variabel baru tersebut saling berkorelasi satu sama lain dan memuat banyak informasi yang terdapat dalam variabel asli. Dengan kata lain analisis faktor merupakan teknik untuk melakukan pengujian terhadap hubungan sekumpulan variabel.

Secara umum, analisis faktor bertujuan untuk mengidentifikasi banyaknya faktor umum menggunakan matriks kovarians atau korelasi. Analisis faktor yang dihasilkan mampu memberikan penjelasan terbaik karena memiliki korelasi yang tinggi antara satu faktor dengan faktor yang lain. Analisis faktor ini perlu dilakukan untuk mengurangi dampak dari multikolinearitas. Setiap faktor yang terpilih, akan digunakan sebagai analisis klaster kedepannya . Dimana faktor yang terpilih mempunyai kontribusi yang besar dalam proses pengelompokan. Penghapusan beberapa fitur yang mempunyai tingkat korelasi rendah, akan mnghasilkan suatu matrik jarak serta hasil pengelompokan yang kuat (*robust*) (Ben-Hur & Guyon, 2003).

Analisis faktor memiliki dua asumsi yang harus dipenuhi sebelum melakukan pembentukan faktor. Asumsi pertama yaitu uji kecukupan data

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

korelasi antar variabel. Dalam melakukan pengujian pada asumsi pertama digunakan *Uji Kaiser Meyer Olkin Measure Of Sampling* (KMO). Asumsi kedua yaitu uji kebebasan antar variabel. Dalam melakukan pengujian pada asumsi kedua ini digunakan uji *Barlett*.

2. Model Analisis Faktor

Variabel awal yang digunakan dalam penelitian dinyatakan dengan kombinasi linier antara beberapa faktor yang kemudian ditambahkan dengan nilai residualnya (Johnson & Wichern, 2007). Misal, X adalah peubah acak teramati, dimana u merupakan komponen yang memiliki vektor rata-rata μ serta matriks kovariansi $\sum$. Model analisis faktor menyatakan bahwa X memiliki ketergantungan linier dengan $F_1, F_2, \dots, F_v$ berupa peubah acak tidak teramati yang disebut dengan *common factors* (faktor umum), dan $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_u$ yang disebut dengan *error* (kesalahan) atau *spesific factor* dimana u merupakan variansi tambahan. Model analisis faktor secara khusus dapat dijelaskan dengan Persamaan (2.1).

$$\begin{aligned}
X_1 - \mu_1 &= l_{11}F_1 + l_{12}F_2 + \cdots + l_{1v}F_v + \varepsilon_1 \\
X_2 - \mu_2 &= l_{21}F_1 + l_{22}F_2 + \cdots + l_{2v}F_v + \varepsilon_2 \\
\vdots \quad \vdots \quad & \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
X_u - \mu_u &= l_{u1}F_1 + l_{u2}F_2 + \cdots + l_{uv}F_v + \varepsilon_u
\end{aligned} \tag{2.1}$$

Dalam bentuk matriks dapat dinyatakan dengan persamaan (2.2).

$$X_{(u\times 1)} - \mu_{u\times 1} = L_{(u\times v)}F_{(v\times 1)} + \varepsilon_{u\times 1} \tag{2.2}$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Model analisis faktor memiliki implikasi tertentu dengan hubungan kovariansi sebagai berikut :

a. $F = 0, kov(F) = (FF^T) = 1$

b. $\varepsilon = 0, kov(\varepsilon) = (\varepsilon\varepsilon^T) = \Psi = \begin{bmatrix} \Psi 1 & 0 & \dots & 0 \\ 0 & \Psi 2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Psi p \end{bmatrix}$

c. F dan ε adalah bebas sehingga $kov(\varepsilon, F) = (\varepsilon F^T) = 0$

Persamaan (2.2) merupakan model faktor *orthogonal* yang memiliki hubungan dengan faktor umum. Sehingga, model rotasi matriks dapat ditulis seperti dengan persamaan (2.2).

Keterangan :

X : Variabel yang teramati

μ : Rata-rata peubah

ε : *Spesific factor* (faktor khusus)

F : *Common factor* (faktor umum)

l : *Loading factor* (muatan faktor)

L :Matriks *Loading Factor*

Masing-masing Fdan ε memenuhi syarat sebagai berikut :

Fdan ε bebas sehingga,

a. $F = 0$ dan $kov(F) = 1$

b. $\varepsilon = 0, kov(\varepsilon) = (\varepsilon\varepsilon^T) = \Psi$

Model analisis faktor *orthogonal* berimplikasi dengan struktur kovariansi untuk X, sehingga :

$$(X - \mu) + (X - \mu)^T = (LF + \varepsilon)(LF + \varepsilon)^T$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$= (LF + \varepsilon)[(LF)^T + (\varepsilon)^T]$$

$$= (LF)(LF)^T + (LF)(\varepsilon)^T + \varepsilon(LF)^T + \varepsilon\varepsilon^T$$

$$= LFF^TL^T + LF\varepsilon^T + \varepsilon F^TL^T + \varepsilon\varepsilon^T$$

Sehingga,

$$kov\ (X) = E[(X - \mu)(X - \mu)^T]$$

$$= LE(FF^T)L^T + LE(F\varepsilon^T) + E(\varepsilon F^T)L^T + E(\varepsilon\varepsilon^T)$$

$$= LL^T + \Psi$$

Model faktor *orthogonal*

$$(X - \mu)F^T = (LF + \varepsilon)F^T = LFF^T + \varepsilon F^T$$

Sehingga,

$$kov(X, F) = E[X - \mu]F^T = LE[FF^T] + E[\varepsilon F^T] = L$$

Dengan demikian, struktur kovariansi yang terdapat dalam model analisis faktor *orthogonal* memenuhi sifat sebagai berikut :

a. $kov\ (X) = LL^T + \Psi$

$Var(X_i) = (l_{i1}^2 + l_{i2}^2 + l_{i3}^2 + \cdots + l_{iv}^2) + \Psi_i$

b. $kov(X, F) = L$ atau $kov(X_i, F_j) = L_{ij}$

Keterangan :

X_i : Variabel $ke - i$ yang teramati

F_j : *Common factor* $ke - j$

μ_i : Rata-rata variabel $ke - i$

v : Jumlah *common factor*

Ψ : Keragaman Spesifik

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

3. Proses Analisis Faktor

Secara umum, proses analisis faktor dapat dilakukan dengan beberapa tahapan sebagai berikut :

a. Melakukan uji asumsi kecukupan variabel yang akan diproses dalam analisis faktor menggunakan *Kaiser Meyer Olkin Measure Of Sampling* (KMO)

Uji KMO memiliki tujuan untuk memberikan informasi terkait semua data yang digunkan dalam penelitian cukup untuk difaktorkan. Hipotesis dari pengujian menggunakan KMO yaitu :

Hipotesis :

H_0 : Korelasi parsial yang terbentuk dari data telah cukup untuk difaktorkan

H_1 : Korelasi parsial yang terbentuk dari data tidak cukup untuk difaktorkan

Pengujian Statistik untuk KMO menggunakan Persamaan (2.3).

$$KMO = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}^2}{\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}^2 + \sum_{i=1}^{n}\sum_{j=1}^{n}p_{ij}^2} \tag{2.3}$$

Keterangan :

i : $1, 2, 3, \ldots, n$

j : $1, 2, \ldots, n$

ρ_{ij} : Koefisien korelasi antara variabel i dan j

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

p_{ij} : Koefisien korelasi parsial antara variabel i dan j

Dimana koefisien korelasi dan koefisien korelasi parsial didapatkan dari perhitungan menggunakan Persamaan (2.4) dan (2.5).

$$\rho = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{\{n\sum X^2 - (\sum X)^2\}\{n\sum Y^2 - (\sum Y)^2\}}} \tag{2.4}$$

$$p = \frac{\rho_{yx_1} - \rho_{yx_2}\rho_{x_1x_2}}{\sqrt{1-\rho_{x_1x_2}^2}\sqrt{1-\rho_{yx_2}^2}} \tag{2.5}$$

Keterangan :

n : Jumlah data pengamatan

Y : Variabel terikat

X : Variabel bebas

Jika $H_0 > \alpha(a = 0{,}05)$, maka H_0 diterima. Sehingga dapat disimpulkan bahwa korelasi parsial yang terbentuk dari data telah cukup untuk difaktorkan. Menurut Wibisono (2003), kriteria kesesuaian penggunaan analisis faktor didasarkan pada aturan nilai KMO tercantum dalam Tabel 2.3 (Bagus, 2016).

Tabel 2.3 Kriteria Kesesuaian Penggunaan Analisis Faktor Menggunakan Nilai KMO

No	Nilai KMO	Keterangan
1.	$KMO \geq 0{,}9$	Sangat Baik
2.	$KMO \geq 0{,}8$	Baik
3.	$KMO \geq 0{,}7$	Menengah
4.	$KMO \geq 0{,}6$	Kurang baik
5.	$KMO \geq 0{,}5$	Cukup
6.	$KMO < 0{,}5$	Tidak dapat diterima

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

b. Melakukan pengujian dependensi menggunakan uji *Barlett*

Uji Barlett pertama kali diperkenalkan oleh Maurice Barlett pada tahun 1937(Arsham & Lovric, 2011). Uji *Bartlett* memiliki tujuan untuk mengetahui hubungan antar variabel satu dengan variabel lainnya dalam kasus *multivariate* (Kota, Setiawan, & Linawati, 2012). Tabel 2.3 adalah tabel yang dapat membantu proses perhitungan pada uji *Barlett*.

Hipotesis :

H_0 : $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2$ (seluruh populasi memiliki varian sama)

H_1 : $\sigma_1^2 \neq \sigma_2^2 \neq \cdots \neq \sigma_n^2$ (terdapat populasi memiliki varian berbeda.

Tabel 2.3 Tabel Perhitungan Uji Barlett

Sampel ke-	dk	$1/_{dk}$	s_i^2	$\log s_i^2$	$\mathrm{dk} \log s_i^2$
1	$n_1 - 1$	$1/_{n_1 - 1}$	s_1^2	$\log s_1^2$	$\mathrm{dk} \log s_1^2$
2	$n_2 - 1$	$1/_{n_2 - 1}$	s_2^2	$\log s_2^2$	$\mathrm{dk} \log s_2^2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	$n_k - 1$	$1/_{n_k - 1}$	s_k^2	$\log s_k^2$	$\mathrm{dk} \log s_k^2$

Dimana s adalah varians gabungan dari seluruh kelompok sampel yang digunakan. Nilai s dapat dihitung menggunakan persamaan (2.6).

$$s^2 = \frac{\sum(n_i - 1)s_i^2}{\sum(n_i - 1)} \tag{2.6}$$

Dari hasil perhitungan s^2, akan didapatkan nilai Barlett dengan persamaan (2.7).

$$B = (\log s^2)\sum(n_i - 1) = (\log s^2)\sum dk \tag{2.7}$$

Selanjutnya, uji *Barlett* dilakukan dengan melakukan perhitungan nilai $chi - kuadrat$ (χ^2_{hitung}) yang dapat dicari menggunakan persamaan (2.8).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$\chi^2_{hitung} = (\ln 10)(B - \sum dk \,.\log s_i^2 \tag{2.8}$$

Setelah nilai χ^2_{hitung} didapatkan langkah selanjutnya adalah membandingkan nilai χ^2_{hitung} dengan χ^2_{tabel} dengan taraf nyata $\alpha = 5\%$ serta derajat kebebasan $(dk) = k - 1$, atau dapat ditulis dengan persamaan (2.9).

$$\chi^2_{tabel} = \chi_{(1-a)(k-1)} \tag{2.9}$$

Dengan k adalah banyaknya kelompok dari sampel yang digunakan.

Pengujian Hipotesis :

Apabila $\chi^2_{hitung} > \chi^2_{tabel}$, maka H_0 ditolak yang mengindisikan bahwa terdapat populasi memiliki varian berbeda

c. Proses *Factoring*

Factoring merupakan proses mereduksi data dari beberapa variabel penelitian awal untuk mendapatkan variabel baru yang memiliki faktor dengan jumlah lebih sedikit dari pada variabel awal. Metode yang digunakan adalah *Principal Componen Analysis* (PCA).

d. Menentukan Jumlah Faktor

Penentuan jumlah faktor ini didasarkan pada nilai *eigenvalue* masing-masing variabel faktor yang didapatkan dari proses *factoring*. *Eigenvalue* adalah jumlah varian yang dijelaskan oleh masing-masing faktor. Faktor yang terpilih adalah faktor yang memiliki nilai *eigenvalue* lebih dari 1.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

e. Rotasi faktor

Hasil dari proses *factoring*, tidak akan ada artinya jika tidak dilakukan rotasi faktor. Rotasi faktor ini berfungsi untuk melakukan penafsiran dan pengulasan lebih lanjut terkait hasil dari jumlah faktor. Jenis rotasi yang digunakan adalah rotasi *orthogonal* dengan menggunakan metode *varimax*.

G. Analisis Kelompok (Klaster)

Analisis klaster merupakan adalah salah satu dari jenis teknik analisis *multivariate*. Analisis klaster bertujuan untuk mengelompokkan seluruh objek pengamatan menjadi beberapa klaster yang dikelompokkan berdasarkan kemiripan dari karakteristik objek yang dimilikinya. Teknik dalam analisis klaster yaitu mengelompokkan seluruh objek pengamatan berdasarkan tingkat kemiripan dari karakteristik objek (Johnson & Wichern, 2007) . Secara otomatis, setiap objek yang memiliki kesamaan secara dekat dengan objek lain akan berada dalam sebuah kelompok yang sama dan akan memiliki kemiripan antar anggota dalam satu kelompok tersebut. Dalam proses analisis klaster, akan dilakukan pengelompokkan dengan melihat pola data dari objek pengamatan. Hal tersebut dilakukan untuk mendapatkan kelompok hasil klaster optimum. Kriteria klaster optimum adalah hasil dari masing-masing klaster yang memiliki tingkat homogenitas (kesamaan) yang tinggi terhadap antar anggota dalam satu kelompok yang sama. Selain itu, hasil klaster harus memiliki tingkat heterogenitas (perbedaan) yang tinggi antar kelompok klaster satu dengan yang lainnya.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Berdasarkan penjelasan diatas, dapat diketahui bahwa analisis klaster memiliki hubungan erat dengan pengukuran kemiripan dan ketidakmiripan objek pengamatan. Algoritma pengelompokkan dengan memanfaatkan ukuran kemiripan dan ketidakkemiripan objek pengamatan berguna untuk memisahkan atau menggabungkan data dari suatu objek pengamatan. Pengukuran tingkat kemiripan (d) antara objek $ke - i$ dan objek $ke - j$ (d_{ij}), adalah suatu fungsi yang mempunyai beberapa sifat sebagai berikut $d_{ij} \geq 0, d_{ii} = 0, d_{ij} = d_{ji}$ dan untuk $d_{ik} + d_{jk} \geq d_{ij}$, untuk setiap i, j dan k. Semakin besar nilai dari pengukuran ketidakmiripan dari dua objek tertentu, maka akan semakin besar pula perbedaan antara dua objek tersebut, maka output akhirnya, kedua objek tersebut tidak akan dikelompokkan dalam klaster yang sama(Johnson & Wichern, 2007). Menurut (Subhash. Sharma, 1996) salah satu faktor yang mempengaruhi hasil dari pengelompokan adalah jarak yang terbentuk antar objek pengamatan satu dengan yang lainya. Sehingga diperlukan sebuah alat ukur khusus yang digunakan untuk menentukan jarak diantara objek pengamatan yang digunakan dalam penelitian.

Secara umum, analisis klaster dikelompokkan menjadi dua jenis, diantaranya yaitu analisis klaster dengan metode hirarki dan non hirarki (Johnson & Wichern, 2007). Analisis klaster dengan metode hirarki, dilakukan dengan tanpa menentukan berapa banyak kelompok klaster yang akan dibentuk terlebih dahulu. Pada metode hirarki ini, jumlah kelompok klaster dibentuk dari pendekatan dengan pola tertentu (Johnson & Wichern, 2007). Sedangkan metode non hirarki dilakukan dengan menentukan terlebih dahulu jumlah kelompok yang akan dibentuk dalam analisis klaster, metode ini cocok digunakan untuk

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

menganalisis data yang memiliki ukuran besar (Hair, Black, Babin, & Anderson, 2013).

H. Metode dan Validasi Analisis *Clustering*

Secara umum analisis *clustering* difokuskan pada data numerik saja. Namun, ada beberapa kasus yang memuat data campuran berupa kategorik dan numerik. Analisis *clustering* pada jenis data yang berbeda tidak dapat diberlakukan dengan cara yang sama. Seperti data katagorik yang tidak dapat diberi perlakuan sama seperti data numerik. Hal ini disebabkan karena pengelompokan data kategorik memiliki tingkat kerumitan yang tinggi jika dibandingkan dengan data numerik (Hair et al., 2013). Berikut akan dijelaskan terkait pengelompokan data numerik, kategorik dan campuran.

1. Data Numerik

   Proses klaster pada data numerik didasarkan pada pengukuran jarak. Analisis pengelompokan ini dengan menggunakan ukuran ketidakmiripan dari data yang berskala numerik pada objek pengamatan. Ukuran jarak yang semakin besar mengindikasikan bahwa objek pengamatan memiliki sedikit tingkat kesamaan begitupun sebaliknya. Dalam penelitian ini metode *clustering* yang digunakan adalah metode *K-Means* .

   Metode *K-Means Clustering* merupakan metode yang mempartisi data menjadi K kelompok. Dimana K merupakan jumlah kelompok yang ditentukan oleh peneliti. Algoritm *K-Means* adalah sebagai berikut (Johnson & Wichern, 2007):

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

a. Menentukan banyaknya k klaster yang diinginkan

b. Menentukan *centroid* awal secara random sebanyak k

c. Menentukan jarak terdekat dari setiap objek pengamatan dengan pusat klaster yang telah ditentukan menggunakan jarak *euclidean* dengan persamaan (2.10).

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{n} (x_{ik} - x_{jk})^2} \tag{2.10}$$

Keterangan :

$d(x_i, x_j)$: Jarak antara dua objek $ke - i$ dan $ke - j$, dimana $i = 1,2,\ldots,n$ dan $j = 1,2,\ldots,n\,; i \neq j$

x_{ik} : Nilai objek $ke - i$ pada kelompok $ke - k$

x_{jk} : Nilai objek $ke - j$ pada kelompok $ke - k$

d. Menentukan *centroid* (pusat klaster) baru dengan menghitung nilai rata-rata masing-masing klaster yang terbentuk menggunakan Persamaan (2.11).

$$T_{kj} = \frac{x_{1j} + x_{2j} + \cdots + x_{nj}}{n} \tag{2.11}$$

Keterangan :

T_{kj} : Pusat klaster $ke - k$ pada variabel $ke - j$, dimana $j = 1,2,..,n$

n : Banyaknya data

e. Menentukan jarak terdekat centroid baru menggunakan jarak *euclidean* menggunakan Persamaan (2.10)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

f. Jika belum didapatkan hasil yang tepat, maka kembali ke perhitungan pada step b.

Setelah proses pengelompokan selesai dilakukan, tahapan selanjutnya yaitu validasi pengelompokan optimum. Dalam tahapan ini, akan didapatkan jumlah kelompok paling optimum. Metode yang digunakan dalam penelitian ini adalah *R-Square* dan *Pseudo F-Statistics*. Untuk metode *R-Square* dan *Pseudo F-Statistics,* Jumlah kelompok optimum dapat dilihat dari nilai *R-Square* dan *Pseudo F-Statistics* yang tinggi. Nilai *Pseudo F-Statistics* dapat dihitung menggunakan Persamaan (2.12)(Orpin & Kostylev, 2006).

$$Pseudo\ F - Statistics = \frac{\left(\frac{R^2}{k-1}\right)}{\left(\frac{1-R^2}{n-k}\right)} \tag{2.12}$$

Dimana nilai R^2 didapatkan dengan perhitungan menggunakan Persamaan (2.13).

$$R^2 = \frac{SST - SSE}{SST} = \frac{SSB}{SST} \tag{2.13}$$

Perhitungan *R-Square* melibatkan beberapa perhitungan keragaman data. Diantaranya yaitu keragaman total, keragaman dalam kelompok dan keragaman antar kelompok (Alvionita et al., 2017). Dimana nilai SST, SSW dan SSB dapat dicari dengan menggunakan Persamaan (2.14), (2.15) dan (2.16).

$$SST = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)^2 \tag{2.14}$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$SSW = \sum_{h=1}^{k}\sum_{j=1}^{m}\sum_{i=1}^{n_k}\left(x_{ijh} - \bar{x}_{jh}\right)^2 \tag{2.15}$$

$$SSB = SST - SSW \tag{2.16}$$

Keterangan :

SST : *Sum of Square Total* (Jumlah keragaman total)

SSW : *Sum of Square Within Group* (Jumlah keragaman dalam kelompok)

SSB : *Sum of Square Between Group* (Jumlah keragaman antar kelompok)

m : Banyaknya variabel numerik dalam pengamatan

k : Banyaknya kelompok yang terbentuk dalam pengamatan

n : Total jumlah objek dalam pengamatan

n_k : Jumlah anggota pada kelompok $ke - k$, dimana $k = 1,2,\ldots,k$

x_{ij} : Nilai objek $ke - i$ pada variabel $ke - l$, dimana $l = 1,2,\ldots,m$

x_{ijh} : Nilai objek $ke - i$ pada variabel $ke - j$, dan kelompok $ke - h$ dimana $l = 1,2,\ldots,m$ dan $h = 1,2,\ldots,k$

$\bar{x}_j$: Rata-rata keseluruhan objek variabel $ke - j$, dimana $j = 1,2,\ldots,m$

$\bar{x}_{jh}$: Rata-rata keseluruhan objek variabel $ke - j$, dan kelompok $ke - h$ dimana $j = 1,2,\ldots,m$ dan $h = 1,2,\ldots,k$

Indeks *R-Square* didefinisikan sebagai ukuran perbedaan antar kelompok. Nilai *R-Square* berkisar antara 0 hingga 1. Semakin nilai $R - Square$ mendekati satu, maka semakin terdapat perbedaan yang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

signifikan antar kelompok yang telah terklaster. Sehingga, nilai $R-Square$ yang besar, menunjukkan bahwa tingkat *heterogenitas* antar kelompok tinggi.

2. Data Kategorik

Metode analisis klaster pada data numerik yang melakukan pengelompokan dengan melihat ukuran jarak dan kemiripan data objek penelitian dinilai tidak cocok jika diterapkan untuk melakukan klaster pada data kategorik. Sehingga, dikembangkan metode non hirarki yang digunakan untuk melakukan pengelompokan data kategorik yaitu *K-Modes*. *K-Modes* adalah pengembangan dari metode *K-Means* yang khusus digunakan untuk menangani kasus data berjenis kategorikal. Metode *K-Modes* merupakan metode *clustering* yang menggunakan ukuran ketidaksamaan sederhana. Metode ini memang dikhususkan untuk melakukan *clustering* pada data kategorik. Metode ini memiliki algoritma efisien yang berbasis pada frekuensi untuk menemukan *mode* (Huang, 1998).

Beberapa modifikasi pada metode *K-Modes* yang diakomodasi dari metode *K-Means* adalah sebagai berikut :

a. Jarak dari dua titik data antara X dan Y merupakan jumlah fitur yang terdapat pada X dan Y. Pengukuran kemiripan antara objek X dan Y menggunakan persamaan (2.17).

$$d(x,y) = \sum_{j=1}^{e} \delta \in (X_j, Y_j) \tag{2.17}$$

Keterangan :

e : Jumlah Fitur

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$\delta \in (X_j, Y_j)$: Nilai pencocokan, nilai tersebut didasarkan pada persamaan (2.18).

$$\delta(x, y) = \begin{cases} 0 \ (x_j = y_j) \\ 1 \ (x_j \neq y_j) \end{cases} \tag{2.18}$$

b. Mengubah nilai *means* (rata-rata) menjadi nilai modus (*modes*)

c. Dalam mencari nilai *modes*, digunakan frekuensi data.

d. Titik *centroid* didapatkan dari modus setiap fitur.

Metode validasi untuk mengetahui pengelompokan yang paling optimum pada data kategorik menggunakan tingkat akurasi yang tinggi. Perhitungan tingkat akurasi didapatkan dengan persamaan (2.19).

$$r = \frac{1}{n} \sum_{h=1}^{k} q_h \tag{2.19}$$

Keterangan :

n : Banyaknya pengamatan

q_h : Banyaknya objek terbanyak (dominasi) pada kelompok $ke - h$ dengan $(h = 1,2, \dots, k)$.

Akurasi nilai *r* didefinisikan sebagai ukuran perbedaan antar kelompok. Nilai *r* berkisar antara 0 hingga 1. Semakin nilai *r* mendekati 1, maka semakin terdapat perbedaan yang signifikan antar kelompok yang telah terklaster. Sehingga, nilai *r* yang besar, menunjukkan bahwa tingkat *heterogenitas* antar kelompok tinggi.

3. Data Campuran (Numerik dan Kategorik)

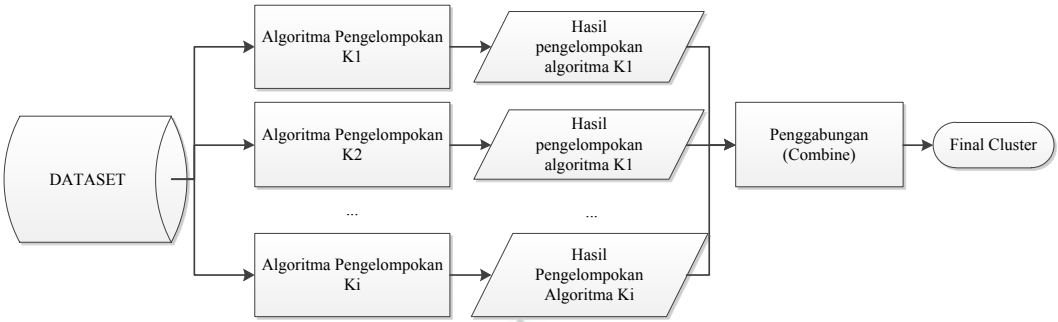
digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Kasus penggunaan data campuran sering kali ditemui dalam beberapa penelitian. Data yang digunakan sering kali memiliki karakteristik data yang berbeda-beda. Sehingga, diperlukan suatu metode khusus yang mampu menangani jenis data tersebut. Metode statistik yang dapat digunakan adalah metode *clustering ensemble*. Apabila terdapat m data yang berskala numerik $(m_{(numerik)})$ dan h data yang berskala kategorik $(h_{(kategorik)})$, maka data l yang berskala campuran yaitu $u = m_{(numerik)} + h_{(kategorik)}$. Selanjutnya, materi terkait klaster *ensemble* akan dibahas pada sub bab selanjutnya.

I. *Ensemble Clustering*

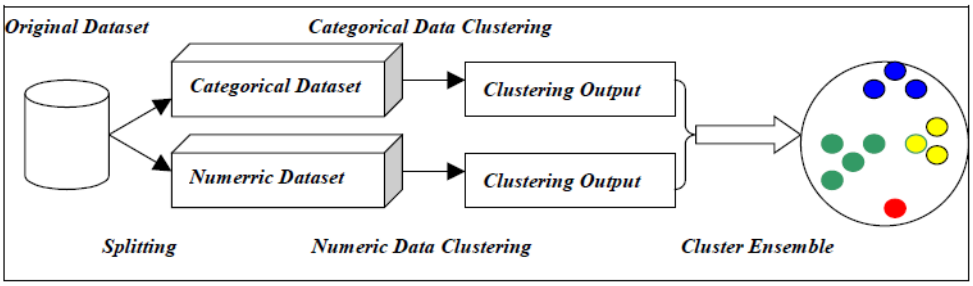
Menurut Suguna & Selvi (2012), *ensemble clustering* adalah suatu metode pengelompokan yang menggabungkan lebih dari satu algoritma agar didapatkan partisi umum dari data. Secara umum, metode *ensemble Clustering* terdiri dari dua tahap algoritma. Tahap pertama yaitu melakukan klaster secara terpisah menggunakan algoritma tertentu yang disesuaikan dengan jenis data (numerik atau kategorik). Hasil dari pengelompokan terpisah tersebut selanjutnya disimpan untuk diolah pada tahap selanjutnya. Tahap kedua yaitu menentukan final klaster berdasarkan hasil pengelompokan pada tahap satu dengan memanfaatkan fungsi konsensus. Secara umum, tahapan untuk penyelesaian kasus klaster *ensambel* seperti yang terlihat pada Gambar (2.3).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 2.3 Tahapan umum pengelompokan *ensemble*

Pengelompokan *ensemble*, dapat dilakukan dengan menggunakan algoritma *Cluster Ensemble Based Mixed Data Clustering* (CEBMDC). Tahapan penyelesaian klaster *ensemble* menggunakan algoritma CEBMDC dapat dilihat pada Gambar (2.4).



Gambar 2.4 Tahapan Algoritma CEBMDC
(Sumber :(Reddy & B.Kavitha, 2012))

Gambar (2.3) merupakan tahapan dalam pengelompokan *ensemble* menggunakan algoritma CEBMDC. Dalam penelitian ini, algoritma CEBMDC akan diterapkan dengan beberapa tahapan sebagai berikut (He, Xu, & Deng, 2005):

1. Melakukan pembagian data menjadi dua sub-data yaitu data numerik dan data kategorik.
2. Melakukan pengelompokan objek untuk sub variabel numerik menggunakan analisis klaster menggunakan metode *K-Means*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

3. Melakukan pengelompokan objek untuk sub variabel kategorik menggunakan analisis klaster menggunakan metode *K-Modes*.
4. Menerapkan proses e*nsemble* dengan menggabung hasil klaster pada tahap 2 dan 3.
5. Melakukan analisis klaster menggunakan *ensemble* ROCK yang akan dijelaskan pada sub bab I.

J. *Robust Clustering Using Links* (ROCK)

Pengelompokan menggunakan metode ROCK ini menggunakan konsep *link* yang digunakan untuk melakukan pengukuran terhadap kesamaan dan kedekatan yang terjadi pada sepasang titik data (Guha, Rastogi, & Shim, 2000). Objek pengamatan yang memiliki *link* (tingkat hubungan) tinggi, maka akan dikelompokkan menjadi satu klaster. Sedangkan objek pengamatan yang memiliki *link* rendah akan dipisahkan dari pengelompokan data tersebut.

Metode ROCK dipilih karena memiliki beberapa kelebihan. Diantaranya yaitu mempunyai tingkat keakurasian yang lebih baik jika dibandingkan dengan metode hirarki lainnya. Selain itu, Metode ROCK memiliki sifat skalabilitas yang baik. Metode ROCK ini mampu menangani data yang mengandung *outlier* secara efektif (Sudipo Guha, Rajeev Rastogi, 1999).

Pengelompokan *ensemble* menggunakan metode ROCK dapat diselesaikan dengan tiga langkah penyelesaian. Tahap pertama yang harus dilakukan adalah menghitung *similaritas*. Ukuran kedekatan atau kemiripan antara objek $ke - i$ dan objek $ke - j$ dapat dihitung menggunakan Persamaan (2.20).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$sim(X_i, X_j) = \frac{X_i \cap X_j}{X_i \cup X_j}, i \neq j \tag{2.20}$$

Dengan $i = 1,2,3,\ldots,n$ dan $j = 1,2,3,\ldots,n$

Keterangan :

X_i : Himpunan kelompok pengamatan $ke - i$

Dimana $X_i = \{x_{1i}, x_{2i}, x_{3i}, \ldots, x_{mi}\}$

X_i : Himpunan kelompok pengamatan $ke - j$

Dimana $X_j = \{x_{1j}, x_{2j}, x_{3j}, \ldots, x_{mj}\}$

Tahapan kedua dalam metode ROCK adalah menentukan tetangga (*Neighbor*), suatu objek pengamatan dikatakan sebagai tetangga apabila nilai $sim(X_i, X_j) \geq \theta$ dimana θ adalah nilai *threshold*. Nilai θ ditentukan sendiri oleh peneliti dengan nilai yang berkisar antara 0 hingga 1(Sudipo Guha, Rajeev Rastogi, 1999).

Langkah terakhir yaitu melakukan perhitungan *link* yang terdapat dalam objek pengamatan. Penentuan nilai *threshold* yang dilakukan pada tahap kedua, sangat mempengaruhi besarnya *link*. Fungsi nilai *threshold* disini sebagai parameter yang ditentukan oleh peneliti dan dapat digunakan sebagai alat kontrol tingkat kedekatan suatu objek. Jika tidak ditemukan lagi *link* antara kelompok objek pengamatan, maka Algortima ROCK akan berhenti (Dutta, Mahanta, & Pujari, 2005). Dengan demikian, dalam mengukur tingkat kemiripan antar objek pengamatan metode ROCK menggunakan informasi l*ink*. Apabila terdapat tiga objek pengamatan yang terdiri dari X_i, X_j dan X_k. Dalam kasus tersebut, diketahui bahawa X_i adalah tetangga dari X_j dan X_j merupakan tetangga dari X_k. Dalam

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kasus ini X_i memiliki *link* dengan X_k maka secara otomatis X_i bertetangga dengan X_k. Dalam menghitung *link* untuk seluruh kemungkinan pasangan dari n objek menggunakan matriks A. Matriks A memiliki ukuran nxn. Apabila X_i dan X_j bertetangga maka akan bernilai 1. Namun jika X_i dan X_j tidak bertetangga, maka akan bernilai 0. Jumlah $link$ antar pasangan objek X_i dan X_j didapatkan dari perkalian matrik antara baris ke X_i dan kolom ke X_j yang terdapat dalam matriks A. Jika terdapat l*ink* yang terbesar antara objek X_i dan X_j semakin besar, maka mereka memiliki kemungkinan yang besar pula untuk berada dalam satu kelompok yang sama (Dutta et al., 2005).

Pengelompokan menggunakan algortima ROCK dapat diukur menggunakan *Goodness measure* (ukuran kebaikan) antar kelompok. Perhitungan *Goodness measure* dapat dihitung menggunakan Persamaan (2.21)

$$G(C_i, C_j) = \frac{link(C_i, C_j)}{(n_i + n_j)^{1+2f(\theta)} - n_i^{1+2f(\theta)} - n_j^{1+2f(\theta)}} \tag{2.21}$$

Keterangan :

$link(C_i, C_j)$: Jumlah *link* dari seluruh kemungkinan pasangan objek yang terdapat dalam C_i dan C_j, dimana $link(C_i, C_j)$ didapatkan dari Persamaan (2.22).

$$link(C_i, C_j) = \sum_{X_{i \in C_i}, X_{j \in C_j}} link(X_i, X_j) \tag{2.22}$$

n_i : Jumlah anggota dalam kelompok $ke - i$

n_j : Jumlah anggota dalam kelompok $ke - j$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$f(\theta)$: Fungsi *threshold* yang didapatkan, dengan $f(\theta) = \frac{1-\theta}{1+\theta}$ dimana θ merupakan nilai *threshold* yang ditentukan oleh peneliti.

K. Validitas Pengelompokan Metode ROCK

Validitas pengelompokan metode ROCK dapat dilakukan dengan melakukan perhitungan nilai $ratio\ sw$ (*sum within*) dan sb (*sum between*) (Bunkers & James, 1996). Jika terdapat pengamatan sebanyak n dimana n_i merupakan jumlah pengamatan pada kategori $ke - i$, dengan $i = 1,2,3, \dots, n$ dan $\sum_{i=1}^{n} n_i = n$. Untuk selanjutnya, n_{ih} adalah jumlah pengamatan pada kategori $ke - i$ serta kelompok $ke - h$, dengan $h = 1,2,3, \dots, k$ dimana k merupakan banyaknya kelompok, dengan demikian $n_{.h} = \sum_{i=1}^{n} n_{ih}$ adalah banyaknya pengamatan pada kelompok $ke - h$ serta $n_{i.} = \sum_{h=1}^{k} n_{ih}$ adalah banyaknya pengamtan pada kategorik $ke - i$.

SST (*Sum Square Total*) untuk data kategorik dapat dirumuskan dengan Persamaan (2.23), SSW (*Sum Square Within*) dirumuskan dengan persamaan (2.24) dan SSB (*Sum Square Between*) dirumuskan dalam persamaan (2.25) sebagai berikut.

$$SST = \frac{n}{2} - \frac{1}{2n}\sum_{i=1}^{n} n_i^2 \tag{2.23}$$

$$SSW = \sum_{h=1}^{k}\left(\frac{n.h}{2} - \frac{1}{2n.h}\sum_{i=1}^{n} n_{ih}^2\right) = \frac{n}{2} - \frac{1}{2}\sum_{h=1}^{k}\frac{1}{n_{.h}}\sum_{i=1}^{n} n_{ih}^2 \tag{2.24}$$

$$SSB = \frac{1}{2}\left(\sum_{h=1}^{k}\frac{1}{n_{.h}}\sum_{i=1}^{n} n_{ih}^2\right) - \frac{1}{2n}\sum_{i=1}^{n} n_{i.}^2 \tag{2.25}$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Selanjutnya untuk MST (*Mean Square Total*) dapat dihitung menggunakan Persamaan (2.26), MSW (*Mean Square Within*) menggunakan Persamaan(2.27) dan MSB (*Mean Square Between*) menggunakan Persamaan (2.28) sebagai berikut.

$$MST = \frac{SST}{(n-1)} \tag{2.26}$$

$$MSW = \frac{SSW}{(n-k)} \tag{2.27}$$

$$MSB = \frac{SSB}{(k-1)} \tag{2.28}$$

Selanjutnya untuk nilai Sw (simpangan baku dalam kelompok) dan Sb (simpangan baku antar kelompok) dirumuskan dengan Persamaan (2.29) dan (2.30).

$$Sw = [MSW]^{\frac{1}{2}} \tag{2.29}$$

$$Sb = [MSB]^{\frac{1}{2}} \tag{2.30}$$

Semakin kecil ratio dari sw dan sb maka akan semakin baik pula kinerja pengelompokan dari hasil klaster yang didapatkan.

L. Pengujian Signifikansi Hasil Klaster

Setelah didapatkan hasil pengelompokan, selanjutnya akan dilakukan pengujian signifikansi atau sering disebut dengan uji perbedaan rata-rata dari kelompok yang didapatkan. Uji signifikansi bertujuan untuk mengetahui apakah terjadi perbedaan yang cukup signifikan antar kelompok dari hasil klaster yang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

terbentuk. Dalam penelitian ini, uji signifikansi akan dilakukan menggunakan uji *Multivariate Analysis of Variance* (MANOVA) dan uji *Kruskal-Wallis.*

1. Uji MANOVA

MANOVA adalah sebuah teknik perluasan dari metode ANOVA (*univariate Analysis of Variance)* yang melibatkan beberapa variabel sekaligus (*Multivariate).* Pengujian hipotesis untuk uji MANOVA adalah sebagai berikut :

- $H_0 : \mu_1 = \mu_2 = \cdots = \mu_K = \mu$ (tidak ada perbedaan yang signifikan antar rata-rata setiap kelompok)
- $H_1 : \mu_K \neq \mu$ (terdapat perbedaan yang signifikan pada rata-rata setiap kelompok)

Salah satu alat uji stastistik yang dapat membantu dalam proses pengambilan keputusan pada rata-rata perbedaan antar kelompok adalah *Wilk's lambda* . Jika nilai stastistik dari uji *Wilk's lambda* semakin rendah, maka perbedaan rata-rata yang terjadi antar kelompok akan semakin signifikan. Nilai *Wilk's lambda* berkisar antara 0 hingga 1. Pengujian statistik *Wilk's lambda* didefinisikan dengan Persamaan (2.28).

$$\Lambda^* = \frac{|SSB|}{|SST|} \qquad (2.28)$$

Setelah nilai *wilk's lambda* selesai didapatkan, tahap selanjutnya adalah mentransformasi nilai *wilk's lambda* kedalam distribusi F. Sehingga, hasil transformasinya akan dibandingkan dengan tabel F. Distribusi *wilk's lambda* disajikan dalam Tabel 2.4.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 2.4 Distribusi *wilk's lambda*

Jumlah Variabel Numerik (m)	Jumlah Kelompok (k)	Distribusi
$m = 1$	$k \geq 2$	$\left(\frac{\sum n_h - k}{k-1}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{h-1,\sum n_h - k}$
$m = 2$	$k \geq 2$	$\left(\frac{\sum n_h - k - 1}{k-1}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2(k-1),2(\sum n_h - k - 1)}$
$m \geq 2$	$k = 2$	$\left(\frac{\sum n_h - m - 1}{m}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{m,(\sum n_h - m - 1)}$
$m \geq 2$	$k = 3$	$\left(\frac{\sum n_h - m - 1}{m}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) F_{2m,2(\sum n_h - m - 2)}$

Pengujian signifikansi menggunakan MANOVA harus memenuhi asumsi ditribusi normal *multivariate*. Pengujian yang dilakukan dapat menggunakan perhitungan nilai *Skewnes* dan *Kurtosis.* Pengujian distribusi normal menggunakan *Skewness* dengan hipotesis sebagai berikut :

- H_0 :$\beta_{1,m} = 0$ (data berdistribusi normal *multivariate*)
- H_1 : $\beta_{1,m} \neq 0$ (data tidak berdistribusi normal *multivariate*)

Statistik uji untuk melakukan pengujian distribusi normal menggunakan *Skewness* dapat melakukan perhitungan dengan Persamaan (2.29).

$$z_{skew} = \frac{(m+1)(N+1)(N+3)}{6\{(N+1)(m+1)-6\}} b_{1,m} \qquad (2.29)$$

Keterangan :

m : Jumlah variabel numerik dalam pengamatan

N : Total objek pengamatan

Apabila $z_{skew} \geq x^2_{\frac{1}{6}m}(m+1)(m+2)$ maka tolak H_0 yang mengindikasikan bahwa data berdistribusi tidak nomal. Pengambilan keputusan juga dapat didasarkan pada nilai $p - value$. Jika $p - value \leq a(a = 5\%)$, maka H_0 ditolak.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Selanjutnya, untuk pengujian distribusi normal menggunakan nilai *Kurtosis* dapat diketahui dengan pengujian hipotesis sebagai berikut :

$H_0 : \beta_{2,m} = 0$ (data berdistribusi normal *multivariate*)

$H_1 : \beta_{2,m} \neq 0$ (data tidak berdistribusi normal *multivariate*)

Statistik uji untuk melakukan pengujian distribusi normal menggunakan *Kurtosis* dapat melakukan perhitungan dengan Persamaan (2.30).

$$z_{kurt} = \frac{b_{1,m} - m(m+2)}{\sqrt{\{8m(m+2)\}/n}} \tag{2.30}$$

Dengan $b_{1,m}$ didapatkan dari $b_{1,m} = \frac{1}{n^2}\sum_{i=1}^{n} g_i^3$ dengan g adalah jarak Mahanlobis yang didapatkan dari $g_i = (x-\mu)'\sum(x-\mu)^{-1}$

Nilai *Kurtosis* dapat dilihat dari kurva, apakah distribusi tersebut terlalu menanjak atau tidak. Jika nilai z_{kurt} lebih besar dari batas atas 2,5%, atau tolak H_0 jika nilai $p - value \leq \alpha$ $(a = 5\%)$.

2. Uji *Kruskal Wallis*

Uji *Kruskal Wallis* adalah pengujian nonparametrik yang bertujuan untuk mengetahui apakah sampel penelitian berasal dari populasi yang sama (identik). Beberapa asumsi yang harus dipenuhi dalam pengujian ini yaitu pengamatan *independent.* Pengujian hipotesis untuk uji *Kruskal Wallis* adalah sebagai berikut :

- H_0 : Kedua populasi identik
- H_1 : Minimal terdapat satu populasi yang identik

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Statistik uji untuk melakukan pengujian perbedaan rata-rata menggunakan *Kruskal Wallis* dapat melakukan perhitungan dengan persamaan (2.31).

$$H = \frac{12}{N(N+1)} \sum_{l=1}^{g} \frac{\underline{R}_l^2}{n_l} - 3(N+1) \tag{2.31}$$

Keterangan :

$\underline{R}_l$: Jumlah seluruh peringkat yang ditetapkan untuk hasil pengamatan dikelompok $ke - i$.

n_l : Banyaknya pengamatan tiap kelompok

N : Banyaknya objek seluruh pengamatan

Pengambilan keputusan didasarkan pada nilai H yang kemudian dibandingkan dengan nilai $X_{(0,05\,;K-1)}$. Jika $H > X_{(0,05\,;K-1)}$, maka H_0 ditolak yang mengindikasikan bahwa tidak semua (minimal satu) populasi identik.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB III
METODE PENELITIAN

Pada bab ini akan dijelaskan metode penelitian yang akan dilakukan dalam proses menyelesaikan permasalahan yang ada dalam penelitian ini. Dalam hal ini, akan dijelaskan secara umum terkait urutan tahapan analisis penyelesaian kasus penelitian data *multivariate* yang memiliki skala campuran berupa data numerik dan kategorik. Metode yang digunakan dalam penelitian ini adalah metode ROCK.

A. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder. Data sekunder adalah data yang didapatkan peneliti secara tidak langsung melalui pihak lain dan media perantara (Susanti, 2016). Data ini didapatkan dari Badan Pusat Statistika (BPS) Jawa Timur dengan alamat website www.jatim.bps.go.id. Data yang digunakan adalah data Potensi Desa (PODES), Statistik Kesejahteraan Rakyat (STAKESRA), Indeks Pembangunan Manusia, Provinsi Jawa Timur dalam angka, Badan Perencanaan dan Pembangunan Daerah dan Data Informasi Perkembangan Daerah tertentu (Daerah Rawan Pangan). Penelitian ini terdiri dari 38 unit pengamatan Kabupaten/Kota di Provinsi Jawa Timur.

B. Variabel Penelitian

Dalam penelitian ini terdapat dua jenis variabel yang digunakan, yaitu variabel alternatif dan variabel kriteria. Variabel alternatif adalah variabel dalam penelitian yang menunjukkan kumpulan objek pengamatan yang secara sengaja ditetapkan oleh peneliti untuk melakukan evaluasi daerah tertinggal di Provinsi

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Jawa Timur. Sedangkan variabel kriteria adalah variabel kriteria penentuan daerah tertinggal di Jawa timur yang terdiri dari beberapa indikator.

Variabel alternatif dalam penelitian ini menyangkut seluruh Kabupaten/Kota di Provinsi Jawa Timur yang terdiri dari 29 Kabupaten dan 9 Kota sesuai dengan rincian yang terdapat pada Tabel (3.1).

Tabel 3.1 Variabel Alternatif Penelitian

Kode	Kabupaten/Kota	Kode	Kabupaten/Kota
01	Kab.Pacitan	20	Kab.Magetan
02	Kab.Ponorogo	21	Kab.Ngawi
03	Kab. Trenggalek	22	Kab.Bojonegoro
04	Kab.Tulungagung	23	Kab. Tuban
05	Kab.Blitar	24	Kab. Lamongan
06	Kab.Kediri	25	Kab. Gresik
07	Kab.Malang	26	Kab.Bangkalan
08	Kab.Lumajang	27	Kab. Sampang
09	Kab. Jember	28	Kab.Pamekasan
10	Kab. Banyuwangi	29	Kab. Sumenep
11	Kab. Bondowoso	30	Kota Kediri
12	Kab.Situbondo	31	Kota Blitar
13	Kab. Probolinggo	32	Kota Malang
14	Kab.Pasuruan	33	Kota Probolinggo
15	Kab. Sidoarjo	34	Kota Pasuruan
16	Kab.Mojokerto	35	Kota Mojokerto
17	Kab.Jombang	36	Kota Madiun
18	Kab.Nganjuk	37	Kota Surabaya
19	Kab.Madiun	38	Kota Batu

Variabel kriteria dalam penelitian ini terdiri dari enam kriteria yang terdiri dari 33 indikator sebagaimana ditunjukkan pada Gambar 2.2 sebelumnya. Variabel-variabel tersebut terdiri dari data yang bertipe numerik dan kategorik. Untuk Variabel kriteria yang bertipe data numerik, penentuan daerah tertinggal memiliki 27 variabel yang ditunjukkan pada Tabel 3.2.

Tabel 3.2 Variabel Kriteria Data Numerik

Kriteria	Indikator	Variabel	Sumber
Perekonomian Masyarakat	Presentase Penduduk Miskin (persen)	x_1	Potensi Desa Jawa Timur
	Pengeluaran penduduk perkapita (ribu rupiah)	x_2	Potensi Desa Jawa Timur

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 3.2 Variabel Kriteria Data Numerik (Lanjutan)

Kriteria	Indikator	Variabel	Sumber
Karakteristik daerah	Presentase jumlah desa terkena bencana gempa	x_3	Potensi Desa Jawa Timur
	Presentase jumlah desa terkena bencana tanah longsor	x_4	Potensi Desa Jawa Timur
	Presentase jumlah desa terkena bencana banjir	x_5	Potensi Desa Jawa Timur
	Presentase desa dengan terkena bencana lainnya	x_6	Potensi Desa Jawa Timur
	Presentase kawasan hutan lindung	x_7	Potensi Desa Jawa Timur
	Presentase desa yang melakukan pelestarian lahan kritis	x_8	Potensi Desa Jawa Timur
	Presentase desa yang mempunyai konflik	x_9	Potensi Desa Jawa Timur
Sumber Daya Manusia	Angka Harapan hidup (tahun)	x_{10}	Indeks Pembangunan Manusia
	Rata-rata lama sekolah (tahun)	x_{11}	Indeks Pembangunan Manusia
	Angka melek huruf (tahun)	x_{12}	Indeks Pembangunan Manusia
Aksesibilitas	Jarak ibu kota Kabupaten ke Ibu Kota provinsi (km)	x_{13}	Jawa Timur Dalam Angka
	Jumlah desa dengan kemudahan akses ke pelayanan keamanan >5 km	x_{14}	Potensi Desa Jawa Timur
Infrastruktur	Jumlah desa dengan permukaan jalan terluas aspal	x_{15}	Potensi Desa Jawa Timur
	Jumlah desa dengan permukaan jalan terluas diperkeras	x_{16}	Potensi Desa Jawa Timur
	Jumlah desa dengan permukaan jalan terluas tanah	x_{17}	Potensi Desa Jawa Timur
	Jumlah desa dengan permukaan jalan terluas lainnya	x_{18}	Potensi Desa Jawa Timur
	Jumlah desa mempunyai pasar tanpa bangunan permanen	x_{19}	Potensi Desa Jawa Timur
	Jumlah prasarana kesehatan per 1000 penduduk	x_{20}	Potensi Desa Jawa Timur
	Jumlah dokter per 1000 penduduk	x_{21}	Potensi Desa Jawa Timur
	Jumlah SMA per 1000 penduduk	x_{22}	Potensi Desa Jawa Timur
	Presentase rumah tangga pengguna listrik (persen)	x_{23}	Potensi Desa Jawa Timur
	Presentase rumah tangga pengguna telepon (persen)	x_{24}	Potensi Desa Jawa Timur
	Presentase rumah tangga pengguna air bersih	x_{25}	Potensi Desa Jawa Timur
Kemampuan Keuangan Daerah	Derajat Desentralisasi Fiskal	x_{26}	Jawa Timur Dalam Angka
Karakteristik daerah tertentu	Presentase Desa Tertinggal	x_{27}	Indeks Pembangunan Desa

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Selanjutnya untuk Varibel kriteria yang bertipe data ketegorik untuk penentuan daerah tertinggal terdapat enam variabel ditunjukkan pada Tabel (3.3).

Tabel 3.3 Variabel Kriteria Data Kategorik

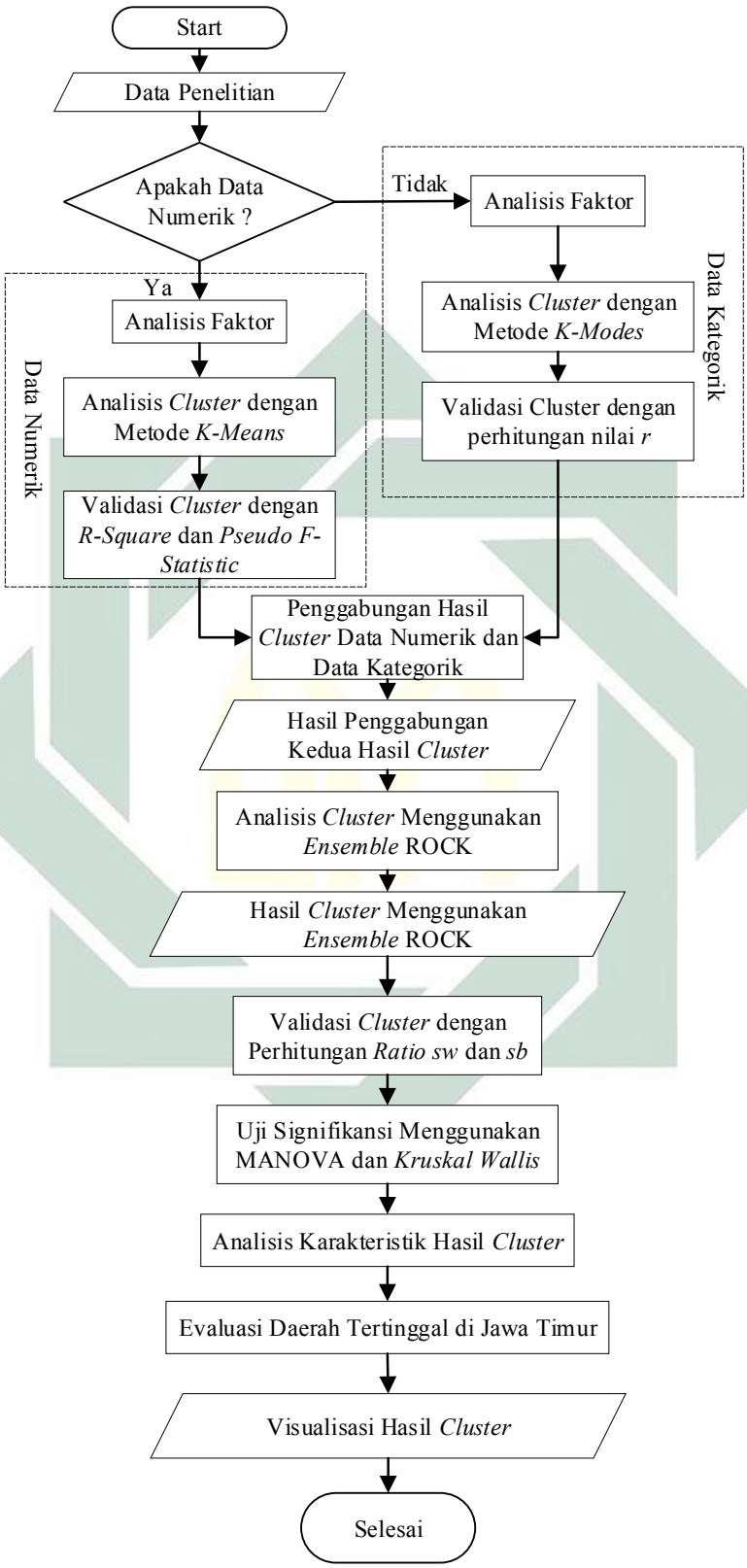
Kriteria	Indikator	Kategori	Variabel	Sumber
	Daerah Perbatasan	0:Bukan Perbatasan 1 :Perbatasan	x_{28}	Jawa Timur Dalam Angka
	Keberadaan Kepulauan	0:Memiliki pulau 1 : Tidak pulau	x_{29}	Jawa Timur Dalam Angka
	Daerah pasca konflik antar suku	0:Tidak ada konflik 1 : ada konflik	x_{30}	Jawa Timur Dalam Angka
	Daerah rawan pangan	0: Bukan daerah rawan pangan 1: Daerah rawan pangan	x_{31}	Pusat Data dan Informasi Perkembangan Daerah Tertentu Rawan Pangan
Karakteristik Daerah Tertentu	Tanah Longsor	0 : Tidak ada 1 : menengah 2 : tinggi	x_{32}	Badan Perencanaan dan Pembangunan Daerah
	Banjir	0 : Tidak ada 1 : Menengah 2 : Tinggi	x_{33}	Badan Perencanaan dan Pembangunan Daerah
	Gempa Bumi	0 : Tidak ada 1 : Ada	x_{34}	Badan Perencanaan dan Pembangunan Daerah
	Tsunami	0 : Tidak ada 1 : Sedang 2 : Besar	x_{35}	Badan Perencanaan dan Pembangunan Daerah
	Letusan Gunung	0 : Tidak ada 1 : Ada	x_{36}	Badan Perencanaan dan Pembangunan Daerah

Berdasarkan data yang termuat dalam Tabel (3.1), (3.2) dan (3.3), diketahui bahwa penelitian ini terdiri dari 38 variabel alternatif dan 36 variabel kriteria.

C. Tahap Analisis Data Penelitian

Tahapan analisis penyelesaian untuk penelitian ini digambarkan pada diagram alir yang ditunjukkan pada gambar (3.1).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

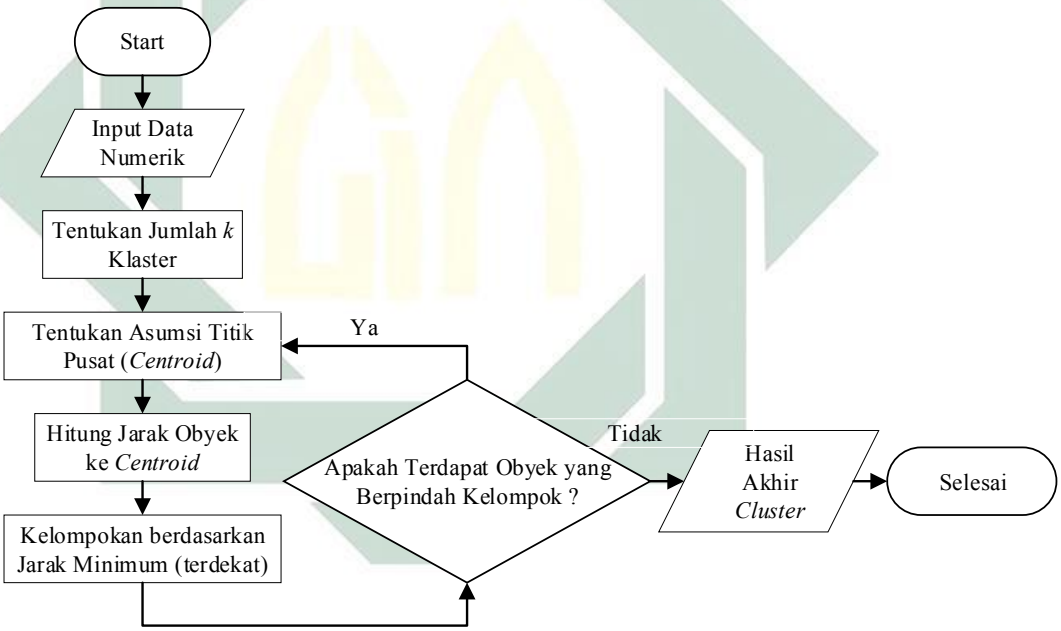


Gambar 3.1 Diagram Alir Tahapan Analisis Data

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Berdasarkan diagram alir penelitian pada Gambar 3.1, tahapan analisis data dalam penelitian ini untuk mencapai tujuan penelitian adalah sebagai berikut.

1. Melakukan pemisahan data campuran menjadi data kategorik dan data numerik.
2. Melakukan reduksi variabel penelitian dengan analisis faktor baik data numerik maupun kategorik.
3. Melakukan analisis klaster data numerik menggunakan metode *K-Means* dengan tahapan analisis yang ditunjukkan pada Gambar 3.2.



Gambar 3.2 Tahapan Analisis *Clustering* Menggunakan *K-Means*

4. Setelah didapatkan hasil *clustering* data numerik, selanjutnya dilakukan validasi pengelompokan optimum menggunakan *Pseudo F-Statistic* dan R^2 berdasarkan Persamaan (2.13) dan (2.14)
5. Melakukan analisis klaster data kategorik menggunakan metode *K-Modes* dengan algoritma yang hampir sama dengan metode *K-Means* pada

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Gambar 3.2, namun perhitungan jarak kemiripan dihitung berdasarkan frekuensi dari *modes* menggunakan Persamaan (2.17) dan (2.18)

6. Melakukan validasi pengelompokan optimum menggunakan perhitungan nilai *r* (akurasi tertinggi) berdasarkan Persamaan (2.19).
7. Menggabungkan hasil pengelompokan dari tahap 4 dan 5 sebagai variabel baru yang bertipe data kategorik.
8. Melakukan analisis klaster hasil pengelompokan pada tahap 6 menggunakan metode ROCK dengan tahap analisis sebagai berikut :
   a. Menghitung $sim(X_i, X_j)$ sebagai pengukuran *similaritas* menggunakan Persamaan (2.20).
   b. Menentukan *Neighbors* (tetangga) dengan menghitung nilai *link* menggunakan Persamaan (2.22).
   c. Mengitung nilai *Goodness measure* menggunakan Persamaan (2.21).
   d. Membandingkan hasil pengelompokan dari masing-masing *threshold* (θ) yang telah ditentukan peneliti.
   a. Validasi *clustering* menggunakan perhitungan nilai $ratio\ sw\ dan\ sb$ berdasarkan Persamaan (2.29) dan (2.30).
9. Melakukan pengujian MANOVA pada data numerik dengan menggunakan persamaan (2.28) dan uji Kruskal Wallis untuk data kategorik dengan menggunakan persamaan (2.31).
10. Melakukan visualisasi hasil *cluster* dengan perwanaan pada peta.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

11. Melakukan analisis karakteristik dari hasil *cluster* yang terbentuk berdasarkan karakteristik dari data seluruh kriteria penetapan daerah tertinggal.
12. Melakukan evaluasi hasil klaster yang tergolong sebagai daerah tertinggal di Jawa Timur.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB IV

HASIL DAN PEMBAHASAN

Dalam penelitian ini akan dilakukan analisis klaster untuk penentuan daerah tertinggal menggunakan survey penduduk tahun 2014. Hasil dari analisis klaster daerah tertinggal tahun 2014 akan dibandingkan dengan hasil penetapan daerah tertinggal oleh Pemerintah. Selanjutnya, akan dilakukan analisis klaster daerah tertinggal menggunakan hasil survey penduduk tahun 2018 dengan tujuan untuk memberikan gambaran terkait penetapan daerah tertinggal tahun 2020 mendatang.

A. Analisis Faktor Data Numerik

Sebelum melakukan analisis klaster untuk data numerik, akan dilakukan analisis faktor data penelitian. Analisis faktor ini dilakukan dengan cara memilih beberapa variabel yang mempunyai kontribusi besar dari faktor yang terbentuk. Data yang akan dianalisis terdiri dari 27 variabel kriteria penentuan daerah tertinggal masing-masing Kabupaten/kota mengacu pada Tabel 3.2. Dari analisis faktor ini akan dihasilkan variabel baru yang berjumlah lebih sedikit dari variabel semula. Variabel baru yang terbentuk akan mewakili variabel dalam penelitian tanpa mengikutsertakan secara keseluruhan.

Langkah pertama yang harus dilakukan dalam analisis faktor adalah melakukan beberapa uji asumsi, diantaranya yaitu uji kecukupan data korelasi antar variabel menggunakan uji KMO dan uji dependensi antar variabel menggunakan uji *Barlett.* Hasil dari uji KMO dan uji *Barlett* dapat dilihat pada Tabel 4.1.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.1 Hasil uji KMO dan Uji Barlett

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,524
Bartlett's Test of Sphericity	Approx. Chi-Square	744,995
	Df	300
	Sig.	,000

Berdasarkan Tabel 4.1 diketahui bahwa nilai KMO (0,524) > 0,500 dan nilai signifikansi uji *Bartlett's Test of Sphericity* diketahui bahwa nilai sig (0,000) < ($\alpha = 0{,}05$). Hal tersebut mengindikasikan bahwa data penelitian telah memenuhi korelasi dan cukup untuk difaktorkan. Hal ini disesuaikan dengan kriteria kesesuaian penggunaan analisis faktor menggunakan nilai KMO yang terdapat pada Tabel 2.3.

Langkah selanjutnya adalah proses *factoring*. Proses *factoring* merupakan proses pemisahan beberapa variabel yang memiliki korelasi kuat. Metode yang dipakai adalah PCA (*Principal Components Analysis*). Pada Tabel 4.2, akan ditunjukkan hasil dari proses *factoring* untuk menentukan hasil dari komponen yang terbentuk dengan melihat nilai *eigenvalue* lebih dari 1.

Tabel 4.2 Hasil *factoring* menggunakan PCA

Component	*Initial Eigenvalues*		
	Total	*% of Variance*	*Cumulative %*
1	**7,092**	28,366	**28,366**
2	**4,273**	17,091	**45,457**
3	**2,157**	8,626	**54,084**
4	**1,898**	7,594	**61,678**
5	**1,587**	6,348	**68,026**
6	**1,430**	5,718	**73,744**
7	**1,073**	4,293	**78,037**
8	**1,030**	4,121	**82,158**
9	,781	3,123	85,282
10	,658	2,633	87,914
11	,642	2,569	90,484
12	,544	2,178	92,661
13	,362	1,449	94,110
14	,299	1,198	95,308
15	,240	,961	96,269

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.2 Hasil *factoring* menggunakan PCA (Lanjutan)

Component	*Initial Eigenvalues*		
	Total	*% of Variance*	*Cumulative %*
16	,234	,937	97,206
17	,197	,789	97,995
18	,128	,510	98,505
19	,114	,457	98,963
20	,076	,304	99,267
21	,059	,236	99,503
22	,053	,214	99,717
23	,035	,138	99,855
24	,024	,097	99,952
25	,012	,048	100,000

Berdasarkan Tabel 4.2, terlihat bahwa jumlah komponen yang memiliki nilai *eigenvalue* lebih dari 1 sebanyak 8 komponen dengan besar nilai *varians comulatif* sebesar 82,16 % yang menunjukkan bahwa 8 komponen yang terpilih dapat mewakili seluruh variabel penelitian sebesar 82,16 %. Dengan demikian, dalam proses analisis klaster, akan digunakan 8 komponen itu untuk mewakili 27 variabel numerik data penelitian. Selanjutnya, untuk menentukan variabel dari 8 komponen yang terbentuk, dapat ditunjukkan dengan nilai *loading factor* terbesar setiap variabelnya yang terlihat pada Tabel 4.3.

Tabel 4.3 Nilai *Loading factor* masing-masing variabel

Indikator	Faktor							
	1	2	3	4	5	6	7	8
x_{11}	**,937**	,047	-,207	-,054	-,078	-,089	,139	,048
x_1	-,889	-,034	,088	-,031	, 148	-,017	,104	-,010
x_2	,849	,247	-,216	-,167	,105	,024	,073	,063
x_{12}	,760	-,056	-,037	-,256	-,058	,098	,099	-,137
x_{25}	,741	,188	-,241	-,078	,374	,064	,062	-,022
x_{27}	-,712	-,075	-,058	-,232	,369	-,026	,088	-,184
x_{10}	,661	,178	-,048	,164	-,247	-,334	,145	,100
x_{24}	,646	,108	-,413	,019	,097	-,230	,395	-,058
x_{15}	-,543	,461	,007	-,086	-,504	-,077	,133	,337
x_{22}	-,034	**,894**	,021	-,152	,093	-,022	,198	-,086
x_{20}	,335	,857	-,100	-,059	-,126	,024	,118	,013
x_{21}	,198	,839	,022	-,062	-,308	-,069	,124	,153
x_{19}	-,013	,755	,138	,022	,173	-,174	-,344	-,090
x_{26}	,336	,502	,198	-,033	-,021	,171	,389	,270

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Indikator	Factor						
	1	2	3	4	5	6	7
x_{17}	-,105	,125	**,900**	,092	,146	-,024	-,150
x_{16}	-,368	,008	,769	-,156	-,102	,134	-,051
x_3	,007	-,073	-,013	**,909**	,098	-,062	-,097
x_4	-,119	-,139	-,040	,872	,053	,113	-,074
x_8	-,226	-,004	,087	,176	**,810**	,077	-,114
x_{14}	-,345	,498	,046	-,150	-,573	-,070	,069
x_{23}	,335	,016	-,441	,100	-,441	,246	-,236
x_5	,112	-,117	,134	,115	-,012	**,854**	,216
x_6	-,429	,019	-,180	-,144	,157	,650	-,331
x_{13}	-,109	-,272	,286	,310	,124	-,092	**,793**

Berdasarkan masing-masing nilai *loading factor* dari setiap variabel, akan dipilih nilai *loading factor* terbesar dari setiap faktornya untuk mengetahui variabel mana yang termasuk kedalam 8 komponen yang terbentuk. Variabel terpilih dari masing-masing faktor dapat dilihat pada Tabel 4.4.

Tabel 4.4 Delapan Variabel Terpilih

Faktor	Indikator	Nilai *Loading*
1	Rata-rata lama sekolah (x_{11})	0,937
2	Jumlah SMA per 1000 penduduk (x_{22})	0,894
3	Jumlah desa dengan permukaan jalan terluas tanah (x_{17})	0,900
4	Presentase desa terkena bencana gempa (x_3)	0,909
5	Presentase desa yang melakukan pelestarian lahan kritis (x_8)	0,810
6	Presentase jumlah desa terkena banjir (x_5)	0,854
7	Jarak ibu kota Kabupaten ke ibu kota Provinsi (x_{13})	0,793
8	Presentase desa yang memiliki konflik (x_9)	0,926

Berdasarkan Tabel 4.4, delapan variabel terpilih tersebut yang akan dianalisis lebih lanjut menggunakan metode *K-Means*.

B. Analisis Klaster Data Numerik Menggunakan *K-Means*

Berdasarkan analisis faktor yang telah dibahas pada bagian A, langkah selanjutnya adalah melakukan analisis klaster data 8 komponen hasil analisis faktor menggunakan metode *K-Means*. Data tersebut akan dilakukan pengklasteran menjadi 2 kelompok, 3 kelompok dan 4 kelompok.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Data yang diklaster terdiri dari 38 data Kabupaten/Kota dan akan dilakukan pengklasteran menurut kriteria dan indikator penentuan daerah tertinggal berdasarkan 8 variabel terpilih dari tahap analisis faktor. Dalam hal ini, akan diberikan contoh perhitungan manual untuk 10 data dari 38 data Kabupaten/Kota sebagai sampel perhitungan yang diambil secara random. Sepuluh data terpilih adalah data ke 1, 4, 9, 12, 15, 18, 21, 27, 31 dan 36. Tabel 4.5 adalah data sampel yang digunakan untuk perhitungan manual.

Tabel 4.5 Data sampel perhitungan manual *K-Means*

Data ke-	Kabupaten	x_3	x_5	x_8	x_9	x_{11}	x_{13}	x_{17}	x_{22}
1	Pacitan	79,532	12,865	89,474	0,585	6,430	276	2	0,009
4	Tulungagung	2,952	5,904	73,801	3,321	7,450	154	3	0,024
9	Jember	1,210	21,371	88,710	1,613	5,633	198	7	0,069
12	Situbondo	1,471	27,941	77,941	0,735	5,540	194	0	0,015
15	Sidoarjo	0,000	12,286	63,714	1,714	10,090	23	0	0,06
18	Nganjuk	0,000	16,549	84,507	5,634	7,310	119	0	0,025
21	Ngawi	0,000	22,581	84,793	0	6,520	181	9	0,019
27	Sampang	0,000	10,753	95,699	1,075	3,490	90	0	0,028
31	Kota Blitar	0,000	0,000	85,714	0	9,810	167	0	0,008
36	Kota Madiun	0,000	14,815	81,481	7,407	10,900	169	0	0,013

Tahap awal yang akan dilakukan adalah menentukan jumlah k. Dalam hal ini akan diberikan contoh $k = 4$. Selanjutnya, akan dilakukan pengolahan untuk iterasi ke-1 dengan menentukan *centroid* secara random untuk masing-masing k. Tabel 4.6 adalah *centroid* iterasi ke-1 yang dipilih secara random dari Tabel 4.5

Tabel 4.6 *Centroid* iterasi ke-1

Data ke-	Kabupaten	x_3	x_5	x_8	x_9	x_{11}	x_{13}	x_{17}	x_{22}	*Centroid*
9	Jember	1,210	21,371	88,710	1,613	5,630	198	7	0,069	C1
18	Nganjuk	0,000	16,549	84,507	5,633	7,310	119	0	0,025	C2
27	Sampang	0,000	10,753	95,699	1,075	3,490	90	0	0,028	C3
36	Kota Madiun	0,000	14,815	81,481	7,407	10,900	169	0	0,013	C4

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Setelah ditetapkan 4 *centroid* secara random, langkah selanjutnya adalah menghitung jarak masing-masing data dari keempat *centroid* menggunakan jarak *euclidean* berdasarkan Persamaan 2.13.

Jarak *euclidean* data ke-1 dengan *centroid* 1

$d(x_i, x_j) = \sqrt{\sum_{k=1}^{p}(x_{ik} - x_{jk})^2} = (79{,}532 - 1{,}210)^2 + (12{,}865 - 21{,}371)^2 + (89{,}474 - 88{,}710)^2 + (0{,}585 - 1{,}613)^2 + (6{,}430 - 5{,}630)^2 + (276 - 198)^2 + (2 - 7)^2 + (0{,}009 - 0{,}069)^2 = 12318{,}039$

Jarak *euclidean* data ke-1 dengan *centroid* 2

$d(x_i, x_j) = \sqrt{\sum_{k=1}^{p}(x_{ik} - x_{jk})^2} = (79{,}532 - 0{,}000)^2 + (12{,}865 - 16{,}549)^2 + (89{,}474 - 84{,}507)^2 + (0{,}585 - 5{,}634)^2 + (6{,}430 - 7{,}310)^2 + (276 - 119)^2 + (2 - 0)^2 + (0{,}009 - 0{,}025)^2 = 31042{,}870$

Jarak *euclidean* data ke-1 dengan *centroid* 3

$d(x_i, x_j) = \sqrt{\sum_{k=1}^{p}(x_{ik} - x_{jk})^2} = (79{,}532 - 0{,}000)^2 + (12{,}865 - 10{,}753)^2 + (89{,}474 - 95{,}699)^2 + (0{,}585 - 1{,}075)^2 + (6{,}430 - 3{,}490)^2 + (276 - 90)^2 + (2 - 0)^2 + (0{,}009 - 0{,}013)^2 = 40977{,}467$

Jarak *euclidean* data ke-1 dengan *centroid* 4

$d(x_i, x_j) = \sqrt{\sum_{k=1}^{p}(x_{ik} - x_{jk})^2} = (79{,}532 - 0{,}000)^2 + (12{,}865 - 14{,}815)^2 + (89{,}474 - 7{,}407)^2 + (0{,}585 - 10{,}900)^2 + (6{,}430 - 7{,}310)^2 + (276 - 169)^2 + (2 - 0)^2 + (0{,}009 - 0{,}013)^2 = 17912{,}569$

Jarak *euclidean* data ke-4 dengan *centroid* 1

$d(x_i, x_j) = \sqrt{\sum_{k=1}^{p}(x_{ik} - x_{jk})^2} = (2{,}952 - 1{,}210)^2 + (5{,}904 - 21{,}371)^2 + (73{,}801 - 88{,}710)^2 + (3{,}321 - 1{,}613)^2 + (7{,}450 - 5{,}630)^2 + (154 - 198)^2 + (3 - 7)^2 + (0{,}024 - 0{,}069)^2 = 12318{,}039$

Dan seterusnya hingga perhitungan jarak *euclidean* data ke-36 dengan *centroid* 4

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Setelah didapatkan hasil perhitungan jarak *euclidean* setiap data dengan keempat *centroid*. Selanjutnya dari hasil perhitungan tersebut, dipilih jarak terdekat antara data dengan salah satu dari keempat *centroid* pada Tabel 4.6. Kemudian, data tersebut akan dialokasikan ke *centroid* terdekat sebagai hasil klaster. Hasil perhitungan jarak *euclidean* dan hasil klaster iterasi ke-1 dapat dilihat pada Tabel 4.7.

Tabel 4.7 Hasil perhitungan jarak *euclidean* dan hasil klaster iterasi ke-1

Data ke-	C1	C2	C3	C4	Hasil
1	**12318,03921**	31042,87	40977,46718	17912,57	1
4	2422,769663	1476,029	4637,479338	**409,7109**	4
9	**0**	6351,366	11880,93109	1048,046	1
12	**224,9775436**	5827,177	11433,26263	885,2445	1
15	31402,67672	9689,609	**5558,336669**	21671,14	3
18	6351,365952	**0**	1035,231534	2528,196	2
21	**314,6659819**	3993,823	8631,185087	370,3259	1
27	11880,93109	1035,232	**0**	6554,641	3
31	1497,231531	2617,327	6185,412322	**297,4532**	4
36	1048,045592	2528,196	6554,64087	**0**	4

Berdasarkan Tabel 4.7, diketahui bahwa data ke-1 masuk kedalam klaster 1, karena data ke-1 memiliki jarak terdekat dengan *centroid* 1. Data ke-4 masuk kedalam klaster 4, karena data ke-4 memiliki jarak terdekat dengan *centroid* 4. Langkah selanjutnya, adalah melakukan iterasi ke-2 dengan menentukan *centroid* baru yang didapatkan dengan cara mengelompokkan data sesuai dengan hasil klaster yang sama kemudian menghitung rata-rata masing-masing data pada klaster yang sama. Tabel 4.8 adalah pengelompokan data yang memiliki klaster 1.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.8 Pengelompokan data berdasarkan klaster 1

Data ke-	x_3	x_5	x_8	x_9	x_{11}	x_{13}	x_{17}	x_{22}
1	79,532	12,866	89,474	0,585	6,43	276	2	0,009
9	1,209	21,371	88,709	1,613	5,63	198	7	0,069
12	1,471	27,941	77,941	0,735	5,54	194	0	0,015
21	0	22,581	84,793	0	6,52	181	9	0,019

Berdasarkan Tabel 4.8, akan dihasilkan *centroid* 1 (C1) baru untuk melakukan iterasi ke-2 dengan menghitung rata-rata ($\bar{x}$) setiap data.

$\bar{x}_3 = \frac{79,532+1,209+1,471+0}{4} = \frac{82,212}{4} = 20,553$

$\bar{x}_5 = \frac{12,866+21,371+27,941+22,581}{4} = \frac{84,758}{4} = 21,190$

Dan seterusnya hingga perhitungan $\bar{x}_{22}$

Selanjutnya untuk pengelompokan data berdasarkan klaster 2, dapat dilihat pada Tabel 4.9.

Tabel 4.9 Pengelompokan data berdasarkan klaster 2

Data ke-	x_3	x_5	x_8	x_9	x_{11}	x_{13}	x_{17}	x_{22}
18	0	16,549	84,507	5,634	7,31	119	0	0,025

Berdasarkan Tabel 4.9, akan dihasilkan *centroid* 2 (C2) baru untuk melakukan iterasi ke-2 dengan menghitung rata-rata ($\bar{x}$) setiap data.

$\bar{x}_3 = \frac{0}{1} = 0$

$\bar{x}_5 = \frac{16,549}{1} = 16,549$

Dan seterusnya hingga perhitungan $\bar{x}_{22}$

Selanjutnya untuk pengelompokan data berdasarkan klaster 3, dapat dilihat pada Tabel 4.10.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.10 Pengelompokan data berdasarkan klaster 3

Data ke-	x_3	x_5	x_8	x_9	x_{11}	x_{13}	x_{17}	x_{22}
15	0	12,286	63,714	1,714	10,09	23	0	0,06
27	0	10,753	95,69892	1,075	3,49	90	0	0,028

Berdasarkan Tabel 4.9, akan dihasilkan *centroid* 3 (C3) baru untuk melakukan iterasi ke-2 dengan menghitung rata-rata ($\bar{x}$) setiap data.

$\bar{x}_3 = \frac{0+0}{2} = 0$

$\bar{x}_5 = \frac{12,286+10,753}{2} = \frac{23,038}{2} = 11,519$

Dan seterusnya hingga perhitungan $\bar{x}_{22}$

Selanjutnya untuk pengelompokan data berdasarkan klaster 4, dapat dilihat pada Tabel 4.11.

Tabel 4.11 Pengelompokan data berdasarkan klaster 4

Data ke-	x_3	x_5	x_8	x_9	x_{11}	x_{13}	x_{17}	x_{22}
4	2,952	5,904	73,801	3,321	7,45	154	3	0,024
31	0	0	85,714	0	9,81	167	0	0,008
36	0	14,815	81,481	7,407	10,9	169	0	0,013

Berdasarkan Tabel 4.9, akan dihasilkan *centroid* 4 (C4) baru untuk melakukan iterasi ke-2 dengan menghitung rata-rata ($\bar{x}$) setiap data.

$\bar{x}_3 = \frac{2,952+0+0}{3} = \frac{2,952}{3} = 0,984$

$\bar{x}_5 = \frac{5,904+0+14,815}{3} = \frac{20,719}{3} = 6,906$

Dan seterusnya hingga perhitungan $\bar{x}_{22}$

Berdasarkan perhitungan rata-rata dari Tabel 4.8, 4.9, 4.10 dan 4.11, didapatkan *centroid* baru untuk melakukan iterasi ke-2 sesuai dengan Tabel 4.12.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.12 *Centroid* untul iterasi ke-2

Centroid	x_3	x_5	x_8	x_9	x_{11}	x_{13}	x_{17}	x_{22}
C1	20,553	21,19	85,229	0,733	6,03	212,25	4,5	0,028
C2	0	16,549	84,507	5,634	7,31	119	0	0,025
C3	0	11,519	79,707	1,395	6,79	56,5	0	0,044
C4	0,984	6,906	80,332	3,576	9,387	163,333	1	0,015

Setelah ditetapkan 4 *centroid* baru, langkah selanjutnya adalah menghitung jarak masing-masing data dari keempat *centroid* menggunakan jarak *euclidean* berdasarkan persamaan 2.13. Dengan cara yang sama pada iterasi ke-1 didapatkan hasil perhitungan jarak *euclidean* setiap data dengan keempat *centroid*. Selanjutnya dari hasil perhitungan tersebut, dipilih jarak terdekat antara data dengan salah satu dari keempat *centroid* pada Tabel 4.12. Kemudian, data tersebut akan dialokasikan ke *centroid* terdekat sebagai hasil klaster. Tabel 4.13 adalah hasil perhitungan jarak *euclidean* dan hasil klaster untuk iterasi ke-2.

Tabel 4.13 Hasil perhitungan jarak *euclidean* dan hasil klaster iterasi ke-2

Data ke-	C1	C2	C3	C4	Hasil
1	**7558,715**	31042,870	54607,610	19001,360	1
4	3829,150	1476,029	9594,520	**142,464**	4
9	**596,711**	6351,366	20252,220	1535,208	1
12	**777,564**	5827,177	19183,209	1412,732	1
15	36729,681	9689,609	**1389,584**	20004,468	3
18	9159,778	0,000	3972,836	2086,376	2
21	1421,602	3993,823	15731,492	**663,667**	4
27	15493,488	**1035,232**	1389,584	5671,703	2
31	2484,115	2617,327	12390,101	**105,045**	4
36	2389,176	2528,196	12723,306	**114,914**	4

Berdasarkan Tabel 4.13, diketahui bahwa data ke-1 masuk kedalam klaster 1, karena data ke-1 memiliki jarak terdekat dengan *centroid* 1. Data ke-4 masuk

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kedalam klaster 4, karena data ke-4 memiliki jarak terdekat dengan *centroid* 4, dan seterusnya.

Setelah didapatkan hasil klaster pada iterasi ke-1 dan ke-2, selanjutnya akan dilakukan perbandingan apakah terdapat data yang menempati kelompok klaster yang berbeda antara iterasi ke-1 dan iterasi ke-2. Jika memang ada, maka akan dilanjut dengan iterasi selanjutnya sampai didapatkan hasil klaster yang sama dengan iterasi sebelumnya. Perbandingan hasil klaster iterasi ke-1 dan iterasi ke-2 dapat dilihat pada Tabel 4.14.

Tabel 4.14 Perbandingan hasil klaster iterasi ke-1 dan iterasi ke-2

Data ke-	Iterasi ke-1	Iterasi ke-2	Keterangan
1	1	1	Sama
4	4	4	Sama
9	1	1	Sama
12	1	1	Sama
15	3	3	Sama
18	2	2	Sama
21	1	4	Beda
27	3	2	Beda
31	4	4	Sama
36	4	4	Sama

Berdasarkan Tabel 4.14, pada data ke-21 dan data ke-27, didapatkan perbedaan antara iterasi ke-1 dan ke-2. Dengan demikian, harus dilanjutkan iterasi selanjutnya, yaitu iterasi ke-3.

Sebagaimana yang dilakukan pada iterasi ke-2, untuk melakukan iterasi selanjutnya, harus menentukan *centroid* baru yang didapatkan dengan cara mengelompokkan data sesuai dengan hasil klaster yang sama kemudian menghitung rata-rata masing-masing data dalam klaster yang sama. Tabel 4.14

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

adalah hasil perhitungan rata-rata masing-masing data dalam kelompok yang sama. Dimana hasil perhitungan tersebut akan dijadikan sebagai *centroid* baru pada iterasi ke-3.

Tabel 4.15 *Centroid* iterasi ke-3

Centroid	x_3	x_5	x_8	x_9	x_{11}	x_{13}	x_{17}	x_{22}
C1	27,404	20,726	85,375	0,978	5,867	222,667	3	0,031
C2	0	13,651	90,103	3,355	5,4	104,5	0	0,027
C3	0	12,286	63,714	1,714	10,09	23	0	0,06
C4	0,738	10,825	81,447	2,682	8,67	167,75	3	0,016

Dengan cara yang sama pada iterasi ke-1 dan iterasi ke-2, didapatkan hasil perhitungan jarak *euclidean* setiap data dengan keempat *centroid*. Selanjutnya dari hasil perhitungan tersebut, dipilih jarak terdekat antara data dengan salah satu dari keempat *centroid* pada Tabel 4.15. Kemudian, data tersebut akan dialokasikan ke *centroid* terdekat sebagai hasil klaster. Tabel 4.16 adalah hasil perhitungan jarak *euclidean* dan hasil klaster untuk iterasi ke-3.

Tabel 4.16 Hasil perhitungan jarak *euclidean* dan hasil klaster iterasi ke-3

Data ke-	C1	C2	C3	C4	Hasil
1	**5641,833**	35751,361	71016,922	18005,585	1
4	5674,662	2797,946	17330,729	**278,545**	4
9	1322,593	8857,341	31402,677	**1105,635**	4
12	1610,813	8371,412	29712,323	**1017,447**	4
15	41185,561	7365,165	**0,000**	21281,655	3
18	11548,725	**258,808**	9689,609	2438,800	2
21	2528,260	6053,696	25610,967	**373,313**	4
27	18572,141	**258,808**	5558,337	6287,136	2
31	4304,948	4142,562	21373,959	**153,986**	4
36	3756,874	4282,611	21671,140	**54,329**	4

Berdasarkan Tabel 4.16, diketahui bahwa data ke-1 masuk kedalam klaster 1, karena data ke-1 memiliki jarak terdekat dengan *centroid* 1. Data ke-4 masuk

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kedalam klaster 4, karena data ke-4 memiliki jarak terdekat dengan *centroid* 4, dan seterusnya.

Setelah didapatkan hasil klaster pada iterasi ke-1 dan ke-2, dan ke-3, selanjutnya akan dilakukan perbandingan apakah terdapat data yang menempati kelompok klaster yang berbeda antara iterasi ke-3 dengan iterasi sebelumnya, yaitu iterasi ke-2. Jika memang ada, maka akan dilanjut dengan iterasi selanjutnya sampai didapatkan hasil klaster yang sama dengan iterasi sebelumnya. Perbandingan hasil klaster iterasi ke-2 dan iterasi ke-3 dapat dilihat pada Tabel 4.17.

Tabel 4.17 Perbandingan hasil klaster iterasi ke-2 dan iterasi ke-3

Data ke-	Iterasi ke-2	Iterasi ke-3	Keterangan
1	1	1	Sama
4	4	4	Sama
9	1	4	Beda
12	1	4	Beda
15	3	3	Sama
18	2	2	Sama
21	4	4	Sama
27	2	2	Sama
31	4	4	Sama
36	4	4	Sama

Berdasarkan Tabel 4.17, pada data ke-9 dan data ke-12, didapatkan perbedaan antara iterasi ke-2 dan ke-3. Dengan demikian, harus dilanjutkan iterasi selanjutnya, yaitu iterasi ke-4.

Sebagaimana yang dilakukan pada iterasi sebelumnya, untuk melakukan iterasi selanjutnya, harus menentukan *centroid* baru yang didapatkan dengan cara menghitung rata-rata masing-masing data dalam klaster yang sama. Tabel 4.18 adalah hasil perhitungan rata-rata masing-masing data dalam kelompok yang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

sama. Dimana hasil perhitungan tersebut akan dijadikan sebagai *centroid* baru pada iterasi ke-4.

Tabel 4.18 *Centroid* iterasi ke-4

Centroid	x_3	x_5	x_8	x_9	x_{11}	x_{13}	x_{17}	x_{22}
C1	79,532	12,865	89,474	0,585	6,430	276,000	2,000	0,009
C2	0,000	13,651	90,103	3,355	5,400	104,500	0,000	0,027
C3	0,000	12,286	63,714	1,714	10,090	23,000	0,000	0,060
C4	0,939	15,435	82,073	2,179	7,642	177,167	3,167	0,025

Dengan cara yang sama pada iterasi sebelumnya, didapatkan hasil perhitungan jarak *euclidean* setiap data dengan keempat *centroid*. Selanjutnya dari hasil perhitungan tersebut, dipilih jarak terdekat antara data dengan salah satu dari keempat *centroid* pada Tabel 4.18. Kemudian, data tersebut akan dialokasikan ke *centroid* terdekat sebagai hasil klaster. Tabel 4.19 adalah hasil perhitungan jarak *euclidean* dan hasil klaster untuk iterasi ke-4.

Tabel 4.19 Hasil perhitungan jarak *euclidean* dan hasil klaster iterasi ke-4

Data Ke-	C1	C2	C3	C4	Hasil
1	**0,000**	35751,361	71016,922	16011,699	1
4	21052,147	2797,946	17330,729	**701,396**	4
9	12318,039	8857,341	31402,677	**532,439**	4
12	13182,699	8371,412	29712,323	**473,647**	4
15	71016,922	7365,165	**0,000**	24131,456	3
18	31042,870	**258,808**	9689,609	3413,477	2
21	15516,012	6053,696	25610,967	**114,062**	4
27	40977,467	**258,808**	5558,337	7834,976	2
31	18401,786	4142,562	21373,959	**375,226**	4
36	17912,569	4282,611	21671,140	**116,287**	4

Berdasarkan Tabel 4.19, diketahui bahwa data ke-1 masuk kedalam klaster 1, karena data ke-1 memiliki jarak terdekat dengan *centroid* 1. Data ke-4 masuk

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kedalam klaster 4, karena data ke-4 memiliki jarak terdekat dengan *centroid* 4, dan seterusnya.

Setelah didapatkan hasil klaster pada iterasi ke-1 dan ke-2, dan ke-3, dan ke-4, selanjutnya akan dilakukan perbandingan apakah terdapat data yang menempati kelompok klaster yang berbeda antara iterasi ke-4 dengan iterasi sebelumnya, yaitu iterasi ke-3. Jika memang ada, maka akan dilanjut dengan iterasi selanjutnya sampai didapatkan hasil klaster yang sama dengan iterasi sebelumnya. Perbandingan hasil klaster iterasi ke-3 dan iterasi ke-4 dapat dilihat pada Tabel 4.20.

Tabel 4.20 Perbandingan hasil klaster iterasi ke-3 dan iterasi ke-4

Data ke-	Iterasi ke-3	Iterasi ke-4	Keterangan
1	1	1	Sama
4	4	4	Sama
9	4	4	Sama
12	4	4	Sama
15	3	3	Sama
18	2	2	Sama
21	4	4	Sama
27	2	2	Sama
31	4	4	Sama
36	4	4	Sama

Berdasarkan Tabel 4.20, diketahui bahwa seluruh data tidak ada yang berpindah klaster antara iterasi ke-3 dan ke-4. Dengan demikian, hasil klaster optimal terdapat pada iterasi ke-4. Setelah didapatkan hasil klaster, langkah selanjutnya adalah melakukan validasi hasil pengelompokan menggunakan perhitungan *R-Square* dan *Pseudo F-Statistic*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Langkah pertama yang harus dilakukan adalah menghitung nilai SST (Sum Square Total) data Tabel 4.5 berdasarkan persamaan 2.16. Penentuan nilai SST dimulai dengan menghitung nilai $\bar{x}_l$.

$$\bar{x}_3 = \frac{79{,}532+2{,}952+\cdots+0{,}000}{10} = \frac{85{,}164}{10} = 8{,}516$$

$$\bar{x}_5 = \frac{12{,}865+5{,}904+\cdots+14{,}815}{10} = \frac{145{,}065}{10} = 14{,}506$$

Dan seterusnya hingga didapatkan $\bar{x}_{22}$

Setelah didapatkan nilai $\bar{x}_l$, langkah selanjutnya adalah menghitung $\sum_{i=1}^{n}(x_{il} - \bar{x}_l)^2$.

$$\sum_{i=1}^{n}(x_{i3} - \bar{x}_3)^2 = (\,79{,}532 - 8{,}516)^2 + (\,2{,}952 - 8{,}516)^2 + \cdots + (\,0{,}000 - 8{,}516)^2$$

$$= 5612{,}407$$

$$\sum_{i=1}^{n}(x_{i5} - \bar{x}_5)^2 = (\,12{,}865 - 14{,}506\,)^2 + (\,5{,}904 - 14{,}506)^2 + \cdots + (\,14{,}815 - 14{,}506)^2$$

$$= 603{,}228$$

Dan seterusnya hingga didapatkan $\sum_{i=1}^{n}(x_{i22} - \bar{x}_{22})^2$

Setelah didapatkan nilai $\bar{x}_l$ dan $\sum_{i=1}^{n}(x_{il} - \bar{x}_l)^2$, langkah selanjutnya akan dihitung nilai SST sebagai berikut.

$$SST = \sum_{l=1}^{m}\sum_{i=1}^{n}(x_{il} - \bar{x}_l)^2 = 5612{,}407 + 603{,}228 + \cdots + 0{,}004 = 49128{,}825$$

Setelah didapatkan nilai SST, langkah selanjutnya adalah menghitung nilai SSW berdasarkan Persamaan 2.17. Perhitungan nilai SSW dimulai dengan menghitung nilai rata-rata keseluruhan objek variabel $ke - l$, dan kelompok $ke - c$ $(\bar{x}_{lc})$. Berdasarkan hasil pengelompokan pada Tabel 4.19, didapatkan nilai $\bar{x}_{l1}, \bar{x}_{l2}, \bar{x}_{l3}$ dan $\bar{x}_{l4}$ dapat diperoleh sebagai berikut.

$$\bar{x}_{31} = \frac{79{,}532}{1} = 79{,}532\,,\ \bar{x}_{51} = \frac{2{,}952}{1} = 2{,}952\,,\ldots,\bar{x}_{221} = \frac{0{,}013}{1} = 0{,}013$$

$$\bar{x}_{32} = \frac{0+0}{2} = 0\,,\ \bar{x}_{52} = \frac{16{,}549+10{,}753}{2} = 13{,}651\,,\ldots,\bar{x}_{22\ 2} = \frac{0{,}025+0{,}028}{2} = 0{,}0265$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$\bar{x}_{33} = \frac{0}{1} = 0$, $\bar{x}_{53} = \frac{12,286}{1} = 12,286 , \ldots , \bar{x}_{22\ 3} = \frac{0,06}{1} = 0,06$

$\bar{x}_{34} = \frac{2,952+1,210+1,471+0+0}{6} = 0,938 , \bar{x}_{54} = \frac{5,904+21,371+27,941+22,581+0+14,815}{6} =$

$15,435 , \ldots , \bar{x}_{224} = \frac{0,024+0,069+0,015+0,019+0,008+0,013}{6} = 0,025$

Setelah didapatkan nilai nilai $\bar{x}_{l1}, \bar{x}_{l2}, \bar{x}_{l3}$ dan $\bar{x}_{l4}$ langkah selanjutnya adalah menghitung $\sum_{l=1}^{m} \sum_{i=1}^{n_c} (x_{ilc} - \bar{x}_{lc})^2$.

$\sum_{l=1}^{m} \sum_{i=1}^{n_c} (x_{il1} - \bar{x}_{l1})^2 = (79,532 - 79,532)^2 + (2,952 - 2,952)^2 + \cdots + (0,013 - 0,013)^2 = 0$

$\sum_{l=1}^{m} \sum_{i=1}^{n_c} (x_{il2} - \bar{x}_{l2})^2 = ((0 - 0)^2 + (0 - 0)^2) + ((16,549 - 13,651)^2 + (10,753 - 13,651)^2) + \cdots + ((0,025 - 0,0265)^2 + (0,028 - 0,0265)^2)$
$= 517,616$

$\sum_{l=1}^{m} \sum_{i=1}^{n_c} (x_{il3} - \bar{x}_{l2})^2 = (0 - 0)^2 + (12,286 - 12,286\ )^2 + \cdots + (0,06 - 0,06\ )^2 = 0$

$\sum_{l=1}^{m} \sum_{i=1}^{n_c} (x_{il4} - \bar{x}_{l4})^2 = ((2,952 - 0,938)^2 + (1,210 - 0,938)^2 + \cdots + (0 - 0,938)^2) + ((5,904 - 15,435)^2 + (21,371 - 15,435)^2 + \cdots + (14,815 - 15,435)^2) + \cdots + ((0,024 - 0,025)^2 + (0,069 - 0,025)^2 + \cdots + (0,013 - 0,025)^2) = 2313,057$

Setelah didapatkan nilai nilai $\bar{x}_{l1}, \bar{x}_{l2}, \bar{x}_{l3}$ dan $\bar{x}_{l4}$ dan $\sum_{l=1}^{m} \sum_{i=1}^{n_c} (x_{ilc} - \bar{x}_{lc})^2$, selanjutnya akan dihitung nilai SSW sebagai berikut.

$SSW = \sum_{c=1}^{C} \sum_{l=1}^{m} \sum_{i=1}^{n_c} (x_{ilc} - \bar{x}_{lc})^2 = 0 + 517,616 + 0 + 2313,057 = 2830,673$

Berdasarkan nilai SST dan SSW didapatkan nilai SSB berdasarkan Persamaan 2.18 sebagai berikut.

$SSB = 49128,825 - 2830,673 = 46298,152$

Setelah nilai SST, SSW, dan SSB didapatkan, akan dihitung nilai *R-Square* dan *Pseudo F-Statistic* menggunakan Persamaan 2.15 dan 2.14.

$R - Square = \frac{SSW}{SST} = \frac{46298,152}{49128,825} = 0,942383$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$Pseudo\ F - Statistic = \frac{\frac{0{,}942383^{\ 2}}{4-1}}{\frac{(1-0{,}942383^{\ 2})}{10-4}} = 15{,}87071$$

Berdasarkan nilai dari *R-Square* dan *Pseudo F-Statistic*, diketahui bahwa pengelompokan hasil klaster sangat baik dengan nilai *R-Square* sebesar 94% dan nilai *Pseudo F-Statistic* sebesar 15,87.

Analisis klaster data secara keseluruhan, dihitung menggunakan sistem. Dimana hasil output yang didapatkan relevan dengan perhitungan sampel data yang telah dijelaskan diatas. Selanjutnya hasil pengklasteran 2 kelompok, 3 kelompok dan 4 kelopok yang dihitung menggunakan sistem dapat dilihat pada Tabel 4.21, Tabel 4.22 dan Tabel 4.23.

Tabel 4.21 Hasil klaster 2 kelompok menggunakan *K-Means*

Kabupaten	Hasil Klaster	Kabupaten	Hasil Klaster
Kab.Pacitan	2	Kab.Magetan	2
Kab.Ponorogo	2	Kab.Ngawi	2
Kab. Trenggalek	2	Kab.Bojonegoro	1
Kab.Tulungagung	2	Kab. Tuban	1
Kab.Blitar	2	Kab. Lamongan	1
Kab.Kediri	1	Kab. Gresik	1
Kab.Malang	1	Kab.Bangkalan	1
Kab.Lumajang	2	Kab. Sampang	1
Kab. Jember	2	Kab.Pamekasan	1
Kab. Banyuwangi	2	Kab. Sumenep	2
Kab. Bondowoso	2	Kota Kediri	1
Kab.Situbondo	2	Kota Blitar	2
Kab. Probolinggo	1	Kota Malang	1
Kab.Pasuruan	1	Kota Probolinggo	1
Kab. Sidoarjo	1	Kota Pasuruan	1
Kab.Mojokerto	1	Kota Mojokerto	1
Kab.Jombang	1	Kota Madiun	2
Kab.Nganjuk	1	Kota Surabaya	1
Kab.Madiun	2	Kota Batu	1

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.22 Hasil klaster 3 kelompok menggunakan *K-Means*

Kabupaten	Hasil Klaster	Kabupaten	Hasil Klaster
Kab.Pacitan	1	Kab.Magetan	1
Kab.Ponorogo	1	Kab.Ngawi	1
Kab. Trenggalek	1	Kab.Bojonegoro	3
Kab.Tulungagung	1	Kab. Tuban	3
Kab.Blitar	1	Kab. Lamongan	2
Kab.Kediri	3	Kab. Gresik	2
Kab.Malang	3	Kab.Bangkalan	2
Kab.Lumajang	3	Kab. Sampang	3
Kab. Jember	1	Kab.Pamekasan	3
Kab. Banyuwangi	1	Kab. Sumenep	1
Kab. Bondowoso	1	Kota Kediri	3
Kab.Situbondo	1	Kota Blitar	1
Kab. Probolinggo	3	Kota Malang	3
Kab.Pasuruan	2	Kota Probolinggo	3
Kab. Sidoarjo	2	Kota Pasuruan	2
Kab.Mojokerto	2	Kota Mojokerto	2
Kab.Jombang	3	Kota Madiun	1
Kab.Nganjuk	3	Kota Surabaya	2
Kab.Madiun	1	Kota Batu	3

Tabel 4.23 Hasil klaster 4 kelompok menggunakan *K-Means*

Kabupaten	Hasil Klaster	Kabupaten	Hasil Klaster
Kab.Pacitan	4	Kab.Magetan	2
Kab.Ponorogo	2	Kab.Ngawi	2
Kab. Trenggalek	2	Kab.Bojonegoro	3
Kab.Tulungagung	2	Kab. Tuban	3
Kab.Blitar	2	Kab. Lamongan	1
Kab.Kediri	3	Kab. Gresik	1
Kab.Malang	3	Kab.Bangkalan	1
Kab.Lumajang	3	Kab. Sampang	3
Kab. Jember	3	Kab.Pamekasan	3
Kab. Banyuwangi	4	Kab. Sumenep	2
Kab. Bondowoso	2	Kota Kediri	3
Kab.Situbondo	2	Kota Blitar	2
Kab. Probolinggo	3	Kota Malang	3
Kab.Pasuruan	1	Kota Probolinggo	3
Kab. Sidoarjo	1	Kota Pasuruan	1
Kab.Mojokerto	1	Kota Mojokerto	1
Kab.Jombang	3	Kota Madiun	2
Kab.Nganjuk	3	Kota Surabaya	1
Kab.Madiun	2	Kota Batu	3

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Berdasarkan hasil klaster dengan pembentukan 2 kelompok, 3 kelompok dan 4 kelompok, selanjutnya akan dibandingkan mana diantara 3 pengelompokan tersebut yang paling optimum menggunakan nilai *R-Square* dan nilai *Pseudo F-Statistic* terbesar. Perbandingan nilai *R-Square* dan *Pseudo F-Statistic* masing-masing pengelompokan dapat dilihat pada Tabel 4.24.

Tabel 4.24 Perbandingan nilai *R-Square* dan *Pseudo F-Statistic* masing-masing pengelompokan

Kelompok Klaster	*R-Square*	*Pseudo F-Statistic*
Kelompok 2	0,6367291	63,0996
Kelompok 3	0,7625969	56,21429
Kelompok 4	0,879009	82,33755

Berdasarkan Tabel 4.24, diketahui bahwa pengelompokan yang paling optimum menggunakan metode *K-Means* adalah kelompok 4 dengan nilai *R-Square* sebesar 87,90% dan nilai *Pseudo F-Statistic* sebesar 82,338%.

C. Analisis Faktor Data Kategorik

Sama halnya dengan data numerik, Sebelum melakukan analisis klaster untuk data kategorik, terlebih dahulu akan dilakukan analisis faktor data penelitian. Analisis faktor ini dilakukan dengan cara memilih beberapa variabel yang mempunyai kontribusi besar dari faktor yang terbentuk. Data yang akan dianalisis terdiri dari 9 variabel kriteria penentuan daerah tertinggal masing-masing Kabupaten/Kota. Langkah pertama yang harus dilakukan dalam analisis faktor adalah melakukan beberapa uji asumsi, diantaranya yaitu uji kecukupan data korelasi antar variabel menggunakan uji KMO dan uji dependensi antar

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

variabel menggunakan uji *Barlett*. Hasil dari uji KMO dan uji *Barlett* dapat dilihat pada Tabel 4.25.

Tabel 4.25 Hasil Uji Asumsi KMO dan *Barlett Test*

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,683
Bartlett's Test of Sphericity	Approx. Chi-Square	123,616
	Df	36
	Sig.	,000

Berdasarkan Tabel 4.25 diketahui bahwa nilai KMO (0,683) > 0,500 dan nilai signifikansi uji *Bartlett's Test of Sphericity* diketahui bahwa nilai sig (0,000) < ($\alpha = 0{,}05$). Hal tersebut mengindikasikan bahwa data penelitian telah memenuhi korelasi dan cukup untuk difaktorkan. Hal ini disesuaikan dengan kriteria kesesuaian penggunaan analisis faktor menggunakan nilai KMO yang terdapat pada Tabel 2.3.

Langkah selanjutnya adalah proses *factoring*. Proses *factoring* merupakan proses pemisahan beberapa variabel yang memiliki korelasi kuat. Metode yang dipakai adalah PCA (*Principal Components Analysis*). Pada Tabel 4.26, akan ditunjukkan hasil dari proses *factoring* untuk menentukan hasil dari komponen yang terbentuk dengan melihat nilai *eigenvalue* lebih dari 1.

Tabel 2.26 Hasil *Factoring* Menggunakan PCA

Component	*Initial Eigenvalues*		
	Total	*% of Variance*	*Cumulative %*
1	**3,373**	**37,481**	**37,481**
2	**1,572**	**17,466**	**54,947**
3	**1,235**	**13,724**	**68,670**
4	,983	10,927	79,597
5	,660	7,335	86,933
6	,435	4,838	91,770
7	,348	3,870	95,640
8	,230	2,560	98,201
9	,162	1,799	100,000

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Berdasarkan Tabel 4.26, terlihat bahwa jumlah komponen yang memiliki nilai *eigenvalue* lebih dari 1 sebanyak 3 komponen dengan besar nilai varians komulatif sebesar 68,670 % yang menunjukkan bahwa 3 komponen yang terpilih dapat mewakili seluruh variabel penelitian sebesar 68,670 %. Dengan demikian, dalam proses analisis klaster, akan digunakan 3 komponen itu untuk mewakili 9 variabel kategorik data penelitian. Selanjutnya, untuk menentukan variabel dari 3 komponen yang terbentuk, dapat ditunjukkan dengan nilai *loading factor* terbesar setiap variabelnya yang terlihat pada Tabel 4.27.

Tabel 4.27 Nilai *Loading factor* masing-masing variabel

Indikator	*Component*		
	1	2	3
x_{34}	**,818**	,341	-,152
x_{32}	,800	,054	,170
x_{36}	,777	,130	-,134
x_{35}	,655	-,050	,310
x_{30}	-,121	**,874**	-,146
x_{31}	,385	,786	,334
x_{33}	,319	,710	,385

Berdasarkan masing-masing nilai *loading factor* dari setiap variabel, akan dipilih nilai *loading factor* terbesar dari setiap faktornya untuk mengetahui variabel mana yang termasuk kedalam 3 komponen yang terbentuk. Variabel terpilih dari masing-masing faktor dapat dilihat pada Tabel 4.28.

Tabel 4.28 Tiga Variabel Terpilih

Faktor	Indikator	Nilai *Loading*
1	Daerah rawan bencana gempa bumi (x_{34})	0,818
2	Daerah pasca konflik (x_{30})	0,874
3	Daerah perbatasan (x_{28})	0,867

Berdasarkan Tabel 4.28, 3 variabel terpilih tersebut yang akan dianalisis lebih lanjut menggunakan metode *K-Modes*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

D. Analisis Klaster Data Kategorik Menggunakan *K-Modes*

Berdasarkan analisis faktor yang telah dibahas pada bagian C, langkah selanjutnya adalah melakukan analisis klaster data 3 komponen hasil analisis faktor menggunakan metode *K-Modes*. Data tersebut akan dilakukan pengklasteran menjadi 2 kelompok, 3 kelompok dan 4 kelompok.

Data yang diklaster terdiri dari 38 data Kabupaten/Kota dan akan dilakukan pengklasteran menurut kriteria dan indikator penentuan daerah tertinggal berdasarkan 3 variabel terpilih dari tahap analisis faktor. Dalam hal ini, akan diberikan contoh perhitungan manual untuk 10 data dari 38 data Kabupaten/Kota sebagai sampel yang dipilih secara random. Sepuluh data terpilih adalah data ke 1, 4, 8, 12, 15, 18, 25, 27, 31 dan 37. Tabel 4.29 adalah data sampel yang digunakan untuk perhitungan manual.

Tabel 4.29 Data sampel perhitungan manual *K-Modes*

Data Ke-	Kabupaten	x_{28}	x_{28}	x_{28}
1	Pacitan	1	0	0
4	Tulungagung	0	1	1
8	Lumajang	1	0	1
12	Situbondo	1	1	1
15	Sidoarjo	0	1	0
18	Nganjuk	0	1	1
25	Gresik	1	1	0
27	Sampang	1	1	0
31	Kota Blitar	0	0	0
37	Kota Surabaya	1	0	0

Tahap awal yang akan dilakukan adalah menentukan jumlah k. Dalam hal ini akan diberikan nilai $k = 4$. Selanjutnya, akan dilakukan pengolahan untuk iterasi ke-1 dengan menentukan *centroid* secara random untuk masing-masing k.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.30 adalah *centroid* iterasi ke-1 yang dipilih secara random dari Tabel 4.30.

Tabel 4.30 *Centroid* iterasi ke-1

Centroid	Data ke-	Kabupaten	x_{28}	x_{30}	x_{34}
C1	8	Lumajang	1	0	1
C2	15	Sidoarjo	0	1	0
C3	25	Gresik	1	1	0
C4	31	Kota Blitar	0	0	0

Setelah ditetapkan 4 *centroid* secara random, langkah selanjutnya adalah menghitung jarak masing-masing data dari keempat *centroid* menggunakan jarak berdasarkan persamaan 2.19 dan 2.20.

Jarak data ke-1 dan *centroid* 1 (C1)

$d(x,y) = \sum_{j=1}^{r} \delta \in (X_j, Y_j) = \delta \in (1,1) + \delta \in (0,0) + \delta \in (0,1) = 0 + 0 + 1 = 1$

Jarak data ke-1 dan *centroid* 2 (C2)

$d(x,y) = \sum_{j=1}^{r} \delta \in (X_j, Y_j) = \delta \in (1,0) + \delta \in (0,1) + \delta \in (0,0) = 1 + 1 + 0 = 2$

Jarak data ke-1 dan *centroid* 3 (C3)

$d(x,y) = \sum_{j=1}^{r} \delta \in (X_j, Y_j) = \delta \in (1,1) + \delta \in (0,1) + \delta \in (0,0) = 0 + 1 + 0 = 1$

Jarak data ke-1 dan *centroid* 4 (C4)

$d(x,y) = \sum_{j=1}^{r} \delta \in (X_j, Y_j) = \delta \in (1,0) + \delta \in (0,0) + \delta \in (0,0) = 1 + 0 + 0 = 1$

Jarak data ke-4 dan *centroid* 1 (C1)

$d(x,y) = \sum_{j=1}^{r} \delta \in (X_j, Y_j) = \delta \in (0,1) + \delta \in (1,0) + \delta \in (1,1) = 0 + 0 + 1 = 0$

Dan seterusnya hingga perhitungan Jarak data ke-37 dan *centroid* 4 (C4)

Setelah didapatkan hasil perhitungan jarak setiap data dengan keempat *centroid*. Selanjutnya dari hasil perhitungan tersebut, dipilih jarak terdekat antara data dengan salah satu dari keempat *centroid* pada Tabel 4.29. Kemudian, data

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

tersebut akan dialokasikan ke *centroid* terdekat sebagai hasil klaster. Hasil perhitungan jarak dan hasil klaster iterasi ke-1 dapat dilihat pada Tabel 4.31.

Tabel 4.31 Hasil perhitungan jarak dan hasil klaster iterasi ke-1

Data Ke-	C1	C2	C3	C4	Hasil
1	1	2	1	1	1
4	2	1	2	2	2
8	0	3	2	2	1
12	1	2	1	3	1
15	3	0	1	1	2
18	2	1	2	2	2
25	2	1	0	2	3
27	2	1	0	2	3
31	2	1	2	0	4
37	1	2	1	1	1

Berdasarkan Tabel 4.31, diketahui bahwa data ke-1 masuk kedalam klaster 1, karena data ke-1 memiliki jarak terdekat dengan *centroid* 1. Data ke-4 masuk kedalam klaster 4, karena data ke-4 memiliki jarak terdekat dengan *centroid* 4, dan seterusnya.

Langkah selanjutnya, adalah melakukan iterasi ke-2 dengan menentukan *centroid* baru yang didapatkan dengan cara mengelompokkan data sesuai dengan hasil klaster yang sama kemudian menentukan modus dari masing-masing data pada klaster yang sama. Tabel 4.32 adalah pengelompokan data yang memiliki klaster 1.

Tabel 4.32 Pengelompokan data dan modus klaster berdasarkan klaster 1

Data ke-	x_{28}	x_{30}	x_{34}
1	1	0	0
9	1	0	1
12	1	1	1
36	1	0	0
Modus	1	0	1

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Selanjutnya untuk pengelompokan data berdasarkan klaster 2, dapat dilihat pada Tabel 4.32.

Tabel 4.32 Pengelompokan data dan modus klaster berdasarkan klaster 2

Data ke-	x_{28}	x_{30}	x_{34}
4	0	1	1
15	0	1	0
18	0	1	1
Modus	0	1	1

Selanjutnya untuk pengelompokan data berdasarkan klaster 3, dapat dilihat pada Tabel 4.33.

Tabel 4.33 Pengelompokan data dan modus klaster berdasarkan klaster 3

Data ke-	x_{28}	x_{30}	x_{34}
21	1	1	0
27	1	1	0
Modus	1	1	0

Selanjutnya untuk pengelompokan data berdasarkan klaster 4, dapat dilihat pada Tabel 4.34.

Tabel 4.34 Pengelompokan data dan modus klaster berdasarkan klaster 4

Data ke-	x_{28}	x_{30}	x_{34}
31	0	0	0
Modus	0	0	0

Berdasarkan perhitungan rata-rata dari Tabel 4.31, 4.32, 4.33 dan 4.34, didapatkan *centroid* baru untuk melakukan iterasi ke-2 sesuai dengan Tabel 4.35.

Tabel 4.35 *Centroid* iterasi ke-2

Centroid	x_{28}	x_{30}	x_{34}
C1	1	0	1
C2	0	1	1
C3	1	1	0
C4	0	0	0

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Setelah ditetapkan 4 *centroid* baru, langkah selanjutnya adalah menghitung jarak masing-masing data dari keempat *centroid* menggunakan persamaan 2.19 dan 2.20. Dengan cara yang sama pada iterasi ke-1 didapatkan hasil perhitungan jarak setiap data dengan keempat *centroid*. Selanjutnya dari hasil perhitungan tersebut, dipilih jarak terdekat antara data dengan salah satu dari keempat *centroid* pada Tabel 4.33. Kemudian, data tersebut akan dialokasikan ke *centroid* terdekat sebagai hasil klaster. Tabel 4.36 adalah hasil perhitungan jarak dan hasil klaster untuk iterasi ke-2.

Tabel 4.36 Hasil perhitungan jarak dan hasil klaster iterasi ke-2

Data Ke-	C1	C2	C3	C4	Hasil
1	1	3	1	1	1
4	2	0	2	2	2
8	0	2	2	2	1
12	1	1	1	3	2
15	3	1	1	1	2
18	2	0	2	2	2
25	2	2	0	2	3
27	2	2	0	2	3
31	2	2	2	0	4
37	1	3	1	1	1

Berdasarkan Tabel 4.36, diketahui bahwa data ke-1 masuk kedalam klaster 1, karena data ke-1 memiliki jarak terdekat dengan *centroid* 1. Data ke-4 masuk kedalam klaster 2, karena data ke-4 memiliki jarak terdekat dengan *centroid* 2, dan seterusnya hingga semua data teralokasikan secara keseluruhan. Setelah didapatkan hasil klaster pada iterasi ke-1 dan ke-2, selanjutnya akan dilakukan perbandingan apakah terdapat data yang menempati kelompok klaster yang berbeda antara iterasi ke-1 dengan iterasi sebelumnya, yaitu iterasi ke-2. Jika

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

memang ada, maka akan dilanjut dengan iterasi selanjutnya sampai didapatkan hasil klaster yang sama dengan iterasi sebelumnya. Perbandingan hasil klaster iterasi ke-1 dan iterasi ke-2 dapat dilihat pada Tabel 4.37.

Tabel 4.37 Perbandingan hasil klaster iterasi ke-1 dan iterasi ke-2

Data ke-	Iterasi ke-1	Iterasi ke-2	Keterangan
1	1	1	Sama
4	2	2	Sama
9	1	1	Sama
12	1	2	Beda
15	2	2	Sama
18	2	2	Sama
21	3	3	Sama
27	3	3	Sama
31	4	4	Sama
36	1	1	Sama

Berdasarkan Tabel 4.35, pada data ke-12, didapatkan perbedaan klaster antara iterasi ke-1 dan ke-2. Dengan demikian, harus dilanjutkan iterasi selanjutnya, yaitu iterasi ke-3.

Sebagaimana yang dilakukan pada iterasi ke-2, untuk melakukan iterasi selanjutnya, harus menentukan *centroid* baru yang didapatkan dengan cara mengelompokkan data sesuai dengan hasil klaster yang sama kemudian menghitung rata-rata masing-masing data dalam klaster yang sama. Tabel 4.38 adalah modus masing-masing data dalam kelompok yang sama. Dimana hasil perhitungan tersebut akan dijadikan sebagai *centroid* baru pada iterasi ke-3.

Tabel 4.38 *Centroid* iterasi ke-3

Centroid	x_{28}	x_{30}	x_{34}
C1	1	0	0
C2	0	1	1
C3	1	1	0
C4	0	0	0

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Setelah ditetapkan 4 *centroid* baru, langkah selanjutnya adalah menghitung jarak masing-masing data dari keempat *centroid* menggunakan persamaan 2.19 dan 2.20. Dengan cara yang sama pada iterasi ke-1 dan ke-2 didapatkan hasil perhitungan jarak setiap data dengan keempat *centroid*. Selanjutnya dari hasil perhitungan tersebut, dipilih jarak terdekat antara data dengan salah satu dari keempat *centroid* pada Tabel 4.38. Kemudian, data tersebut akan dialokasikan ke *centroid* terdekat sebagai hasil klaster. Tabel 4.39 adalah hasil perhitungan jarak dan hasil klaster untuk iterasi ke-3.

Tabel 4.39 Hasil perhitungan jarak dan hasil klaster iterasi ke-3

Data Ke-	C1	C2	C3	C4	Hasil
1	0	3	1	1	1
4	3	0	2	2	2
8	1	2	2	2	1
12	2	1	1	3	2
15	2	1	1	1	2
18	3	0	2	2	2
25	1	2	0	2	3
27	1	2	0	2	3
31	1	2	2	0	4
37	0	3	1	1	1

Berdasarkan Tabel 4.39, diketahui bahwa data ke-1 masuk kedalam klaster 1, karena data ke-1 memiliki jarak terdekat dengan *centroid* 1. Data ke-4 masuk kedalam klaster 2, karena data ke-4 memiliki jarak terdekat dengan *centroid* 2, dan seterusnya hingga semua data teralokasikan secara keseluruhan.

Setelah didapatkan hasil klaster pada iterasi ke-1 dan ke-2, selanjutnya akan dilakukan perbandingan apakah terdapat data yang menempati kelompok klaster yang berbeda antara iterasi ke-3 dengan iterasi sebelumnya, yaitu iterasi ke-2. Jika

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

memang ada, maka akan dilanjut dengan iterasi selanjutnya sampai didapatkan hasil klaster yang sama dengan iterasi sebelumnya. Perbandingan hasil klaster iterasi ke-2 dan iterasi ke-3 dapat dilihat pada Tabel 4.37.

Tabel 4.40 Perbandingan hasil klaster iterasi ke-2 dan iterasi ke-3

Data ke-	Iterasi ke-2	Iterasi ke-3	Keterangan
1	1	1	Sama
4	2	2	Sama
9	1	1	Sama
12	2	2	Sama
15	2	2	Sama
18	2	2	Sama
21	3	3	Sama
27	3	3	Sama
31	4	4	Sama
36	1	1	Sama

Berdasarkan Tabel 4.37, diketahui bahwa seluruh data tidak ada yang berpindah klaster antara iterasi ke-2 dan ke-3. Dengan demikian, hasil klaster optimal terdapat pada iterasi ke-3. Setelah didapatkan hasil klaster, langkah selanjutnya adalah melakukan validasi hasil pengelompokan menggunakan perhitungan nilai akurasi (r).

Berdasarkan tabel 4.32, 4.33 dan 4.34, diketahui bahwa pada klaster 1 terdiri dari 4 objek, klaster 2 terdiri dari 3 objek, klaster 3 terdiri dari 2 objek dan klaster 4 terdiri dari 1 objek. Berdasarkan pengelompokan objek tiap klasternya, dilihat modus data dari masing-masing variabel dan dihitung jumlahnya. Seperti pada klaster 1, diketahui bahwa pada variabel x_{28} memiliki nilai modus 1, dengan kemunculan sebanyak 4 kali, pada klaster 2 variabel x_{28} memiliki mnilai modus 0, dengan kemunculan sebanyak 3 kali, dan seterusnya. Kemudian dihitung berdasarkan Persamaan 2.21. Untuk lebih jelasnya, Tabel 4.41 adalah tabel perhitungan nilai r.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.41 Perhitungan nilai r

Klaster	x_{28}	x_{30}	x_{34}
C1	4	3	2
C2	3	3	2
C3	2	2	2
C4	1	1	1
Jumlah	10	9	7
Rata-rata	1	0,9	0,7

Berdasarkan tabel 4.41, didapatkan perhitungan r sebagai berikut.

$$r = \frac{1 + 0{,}9 + 0{,}7}{3} = 0{,}86$$

Selanjutnya, untuk analisis klaster data secara keseluruhan, dihitung menggunakan sistem. Dimana hasil output yang didapatkan relevan dengan perhitungan sampel data yang telah dijelaskan diatas. Selanjutnya hasil pengklasteran 2 kelompok, 3 kelompok dan 4 kelompok yang dihitung pada Tabel 4.42, Tabel 4.43 dan Tabel 4.44.

Tabel 4.42 Hasil klaster 2 kelompok menggunakan *K-Modes*

Kabupaten	Hasil Klaster	Kabupaten	Hasil Klaster
Kab.Pacitan	1	Kab.Magetan	1
Kab.Ponorogo	1	Kab.Ngawi	1
Kab. Trenggalek	1	Kab.Bojonegoro	1
Kab.Tulungagung	2	Kab. Tuban	1
Kab.Blitar	2	Kab. Lamongan	1
Kab.Kediri	2	Kab. Gresik	1
Kab.Malang	2	Kab.Bangkalan	1
Kab.Lumajang	1	Kab. Sampang	1
Kab. Jember	2	Kab.Pamekasan	1
Kab. Banyuwangi	2	Kab. Sumenep	1
Kab. Bondowoso	2	Kota Kediri	1
Kab.Situbondo	2	Kota Blitar	1
Kab. Probolinggo	1	Kota Malang	1
Kab.Pasuruan	1	Kota Probolinggo	1
Kab. Sidoarjo	2	Kota Pasuruan	1
Kab.Mojokerto	2	Kota Mojokerto	1
Kab.Jombang	2	Kota Madiun	1
Kab.Nganjuk	2	Kota Surabaya	1
Kab.Madiun	2	Kota Batu	1

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.43 Hasil klaster 3 kelompok menggunakan *K-Modes*

Kabupaten	Hasil Klaster	Kabupaten	Hasil Klaster
Kab.Pacitan	1	Kab.Magetan	1
Kab.Ponorogo	1	Kab.Ngawi	1
Kab. Trenggalek	1	Kab.Bojonegoro	3
Kab.Tulungagung	2	Kab. Tuban	3
Kab.Blitar	2	Kab. Lamongan	3
Kab.Kediri	2	Kab. Gresik	3
Kab.Malang	2	Kab.Bangkalan	1
Kab.Lumajang	1	Kab. Sampang	3
Kab. Jember	2	Kab.Pamekasan	3
Kab. Banyuwangi	2	Kab. Sumenep	3
Kab. Bondowoso	2	Kota Kediri	1
Kab.Situbondo	2	Kota Blitar	1
Kab. Probolinggo	1	Kota Malang	1
Kab.Pasuruan	1	Kota Probolinggo	1
Kab. Sidoarjo	2	Kota Pasuruan	1
Kab.Mojokerto	2	Kota Mojokerto	1
Kab.Jombang	2	Kota Madiun	1
Kab.Nganjuk	2	Kota Surabaya	1
Kab.Madiun	2	Kota Batu	1

Tabel 4.44 Hasil klaster 4 kelompok menggunakan *K-Modes*

Kabupaten	Hasil Klaster	Kabupaten	Hasil Klaster
Kab.Pacitan	1	Kab.Magetan	1
Kab.Ponorogo	1	Kab.Ngawi	1
Kab. Trenggalek	1	Kab.Bojonegoro	3
Kab.Tulungagung	2	Kab. Tuban	3
Kab.Blitar	2	Kab. Lamongan	3
Kab.Kediri	2	Kab. Gresik	3
Kab.Malang	2	Kab.Bangkalan	1
Kab.Lumajang	1	Kab. Sampang	3
Kab. Jember	2	Kab.Pamekasan	3
Kab. Banyuwangi	2	Kab. Sumenep	3
Kab. Bondowoso	2	Kota Kediri	4
Kab.Situbondo	2	Kota Blitar	4
Kab. Probolinggo	1	Kota Malang	4
Kab.Pasuruan	1	Kota Probolinggo	1
Kab. Sidoarjo	2	Kota Pasuruan	1
Kab.Mojokerto	2	Kota Mojokerto	4
Kab.Jombang	2	Kota Madiun	4
Kab.Nganjuk	2	Kota Surabaya	1
Kab.Madiun	2	Kota Batu	4

Berdasarkan hasil klaster dengan pembentukan 2 kelompok, 3 kelompok dan 4 kelompok, selanjutnya akan dibandingkan mana diantara 3 pengelompokan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

tersebut yang paling optimum menggunakan perhitungan akurasi nilai r. Perbandingan nilai r masing-masing pengelompokan dapat dilihat pada Tabel 4.45.

Tabel 4.45 Perbandingan nilai r masing-masing pengelompokan

Kelompok Klaster	r
Kelompok 2	0,7894737
Kelompok 3	0,850877193
Kelompok 4	0,903508772

Berdasarkan Tabel 4.44, diketahui bahwa pengelompokan yang paling optimum menggunakan metode *K-Modes* adalah kelompok 4 dengan nilai r sebesar 0,903508772 atau 90,4%.

E. Analisis Klaster Data Campuran Menggunakan ROCK

Berdasarkan hasil analisis klaster yang telah dibahas pada bagian B dan D, langkah selanjutnya adalah melakukan analisis klaster data campuran hasil klaster dari data numerik yang didapatkan dari metode *K-Means* dan data kategorik yang didapatkan dari data *K-Modes* menggunakan ROCK. Hasil pengelompokan metode *K-Means* dan *K-Modes* masing-masing terdiri dari 4 kelompok. Kedua hasil klaster tersebut digabung dan dianggap sebagai data baru yang memiliki skala kategorik, hasil gabungan data tersebut selanjutnya akan dianalisis menggunakan metode *ensemble* ROCK. Struktur data campuran yang akan dianalisis menggunakan metode *ensemble* ROCK dapat dilihat pada Tabel 4.45 sebagai berikut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.46 Struktur data campuran

Kabupaten	Hasil Klaster Numerik	Hasil Klaster Kategorik
Pacitan	4	1
Ponorogo	2	1
Trenggalek	2	1
.	.	.
.	.	.
.	.	.
Kota Mojokerto	1	4
Kota Madiun	2	4
Kota Surabaya	1	1
Kota Batu	3	4

Dalam penelitian ini, data campuran tersebut akan dikelompokkan menjadi 3 dan 4 klaster. Nilai *threshold* (θ) yang akan diuji cobakan sebanyak 10 *threshold* (θ) yaitu 0,1; 0,2; 0,3; 0,01; 0,02; 0,03; 0,04; dan 0,05. Dari 10 *threshold* yang digunakan, akan dipilih nilai *threshold* yang menghasilkan kelompok optimum dengan melihat perbandingan $ratio\ sw$ dan sb terkecil. Tabel 4.47 adalah tabel hasil nilai $ratio\ sw$ dan sb dari masing-masing *threshold* untuk pengelompokan 3 dan 4 klaster.

Tabel 4.47 Nilai $ratio\ sw$ dan sb dari masing-masing *threshold* untuk pengelompokan 3 dan 4 klaster

Threshold	3 Klaster	4 Klaster
0,1	0,7237014	0,855867
0,15	0,5414150	0,9331397
0,2	0,4737199	0,5209103
0,25	1,3563698	0,8428222
0,3	0,4157780	0,4938904
0,01	0,5260732	0,8196005
0,02	0,5140327	**0,3917473**
0,03	**0,4041851**	0,8833502
0,04	0,5883901	0,8390868
0,05	0,5940107	0,5515427

Berdasarkan Tabel 4.47, diketahui bahwa nilai $ratio\ sw$ dan sb terkecil dimiliki oleh *threshold* 0,02 dengan pengelompokan sebanyak 4 klaster. Seperti pada analisis klaster data numerik dan kategorik sebelumnya, dalam hal ini juga akan diberikan contoh perhitungan manual untuk 10 data dari 38 data

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Kabupaten/Kota yang diambil secara acak. Sepuluh data terpilih adalah data ke 1, 4, 8, 12, 15, 18, 25, 27, 31 dan 37 dengan menggunakan nilai *threshold* 0,02 dengan pengelompokan sebanyak 4 klaster. Tabel 4.29 adalah data sampel yang digunakan untuk perhitungan manual.

Tabel 4.48 Data sampel perhitungan manual ROCK

Kode	Data Ke-	Kabupaten	Numerik	Kategorik
P1	1	Pacitan	1	1
P2	4	Tulungagung	4	2
P3	8	Lumajang	4	1
P4	12	Situbondo	4	1
P5	15	Sidoarjo	3	2
P6	18	Nganjuk	2	2
P7	25	Gresik	3	3
P8	27	Sampang	2	3
P9	31	Kota Blitar	4	4
P10	37	Kota Surabaya	3	1

Langkah pertama dalam analisis klaster menggunakan metode ROCK adalah menghitung *similaritas* menggunakan Persamaan 2.25.

$$sim(P1, P2) = \frac{P1 \cap P2}{P1 \cup P2} = \frac{n(P1 \cap P2)}{n(P1) + n(P2) - n(P1 \cap P2)} = \frac{0}{2 + 2 - 0} = \frac{0}{4} = 0$$

$$sim(P1, P3) = \frac{P1 \cap P3}{P1 \cup P3} = \frac{n(P1 \cap P3)}{n(P1) + n(P3) - n(P1 \cap P3)} = \frac{1}{2 + 2 - 1} = \frac{1}{3} = 0{,}3333$$

$$sim(P1, P4) = \frac{P1 \cap P4}{P1 \cup P4} = \frac{n(P1 \cap P4)}{n(P1) + n(P4) - n(P1 \cap P4)} = \frac{1}{2 + 2 - 1} = \frac{1}{3} = 0{,}3333$$

Dan seterusnya hingga $sim(P9, P10)$

Untuk nilai *similaritas* dengan dirinya sendiri pasti bernilai 1. Hasil perhitungan tersebut akan dibentuk matrik *similaritas* sebagai berikut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$sim = \begin{array}{c} P1 \\ P2 \\ P3 \\ P4 \\ P5 \\ P6 \\ P7 \\ P8 \\ P9 \\ P10 \end{array} \begin{bmatrix} 1 & 0 & 0{,}333 & 0{,}333 & 0 & 0 & 0 & 0 & 0 & 0{,}333 \\ 0 & 1 & 0{,}333 & 0{,}333 & 0{,}333 & 0{,}333 & 0 & 0 & 0{,}333 & 0 \\ 0{,}333 & 0{,}333 & 1 & 1 & 0 & 0 & 0 & 0 & 0{,}333 & 0{,}333 \\ 0{,}333 & 0{,}333 & 1 & 1 & 0 & 0 & 0 & 0 & 0{,}333 & 0{,}333 \\ 0 & 0{,}333 & 0 & 0 & 1 & 0{,}333 & 0{,}333 & 0 & 0 & 0{,}333 \\ 0 & 0{,}333 & 0 & 0 & 0{,}333 & 1 & 0 & 0{,}333 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0{,}333 & 0 & 1 & 0{,}333 & 0 & 0{,}333 \\ 0 & 0 & 0 & 0 & 0 & 0{,}333 & 0{,}333 & 1 & 0 & 0 \\ 0 & 0{,}333 & 0{,}333 & 0{,}333 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0{,}333 & 0 & 0{,}333 & 0{,}333 & 0{,}333 & 0 & 0{,}333 & 0 & 0 & 1 \end{bmatrix}$$

Setelah didapatkan matrik *similaritas* diatas, langkah selanjutnya adalah menentukan *Neighbor*. Dalam menentukan *Neighbor*, akan digunakan nilai *threshold* sebagai ambang batas apakah objek tersebut bertetangga atau tidak. Dalam hal ini akan digunakan nilai *threshold* $(\theta) = 0{,}02$. Suatu objek dikatakan bertetangga apabila nilai *similaritas*nya lebih dari nilai *threshold* yang ditentukan. Jika objek tersebut bertetangga, maka akan bernilai 1. Jika tidak, maka akan bernilai 0. Dengan demikian, didapatkan matriks tetangga sebagai berikut.

$$A = \begin{array}{c} P1 \\ P2 \\ P3 \\ P4 \\ P5 \\ P6 \\ P7 \\ P8 \\ P9 \\ P10 \end{array} \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Berdasarkan matriks A diatas, telah diketahui hubungan tetangga antar objek pengamatan dengan keterangan sebagai berikut.

$Neighbor\ [P1] = \{P1, P3, P4, P10\}$

$Neighbor\ [P2] = \{P2, P3, P4, P5, P6, P9\}$

$Neighbor\ [P3] = \{P1, P2, P3, P4, P9, P10\}$

Dan seterusnya hingga $Neighbor\ [P10]$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Dari matriks *Neighbor* A diatas, didapatkan banyaknya *link* antar objek pengamatan satu dengan yang lainnya dengan melakukan perkalian matriks $A \times A$. Matriks *link* dengan ukuran 10×10 dapat dilihat sebagai berikut.

$$A \times A = \begin{array}{c} P1 \\ P2 \\ P3 \\ P4 \\ P5 \\ P6 \\ P7 \\ P8 \\ P9 \\ P10 \end{array} \begin{bmatrix} 4 & 2 & 4 & 4 & 1 & 0 & 1 & 0 & 2 & 4 \\ 2 & 6 & 4 & 4 & 3 & 3 & 1 & 1 & 4 & 3 \\ 4 & 4 & 6 & 6 & 2 & 1 & 1 & 0 & 4 & 4 \\ 4 & 4 & 6 & 6 & 2 & 1 & 1 & 0 & 4 & 4 \\ 1 & 3 & 2 & 2 & 5 & 3 & 3 & 2 & 1 & 3 \\ 0 & 3 & 1 & 1 & 3 & 4 & 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 3 & 2 & 4 & 2 & 0 & 3 \\ 0 & 1 & 0 & 0 & 2 & 2 & 2 & 3 & 0 & 1 \\ 2 & 4 & 4 & 4 & 1 & 1 & 0 & 0 & 4 & 2 \\ 4 & 3 & 4 & 4 & 3 & 1 & 3 & 1 & 2 & 6 \end{bmatrix}$$

Karena *link* dengan dirinya sendiri didefinisikan dengan 0, maka matriks *link* B didapatkan sebagai berikut.

$$B = \begin{array}{c} P1 \\ P2 \\ P3 \\ P4 \\ P5 \\ P6 \\ P7 \\ P8 \\ P9 \\ P10 \end{array} \begin{bmatrix} 0 & 2 & 4 & 4 & 1 & 0 & 1 & 0 & 2 & 4 \\ 2 & 0 & 4 & 4 & 3 & 3 & 1 & 1 & 4 & 3 \\ 4 & 4 & 0 & 6 & 2 & 1 & 1 & 0 & 4 & 4 \\ 4 & 4 & 6 & 0 & 2 & 1 & 1 & 0 & 4 & 4 \\ 1 & 3 & 2 & 2 & 0 & 3 & 3 & 2 & 1 & 3 \\ 0 & 3 & 1 & 1 & 3 & 0 & 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 3 & 2 & 0 & 2 & 0 & 3 \\ 0 & 1 & 0 & 0 & 2 & 2 & 2 & 0 & 0 & 1 \\ 2 & 4 & 4 & 4 & 1 & 1 & 0 & 0 & 0 & 2 \\ 4 & 3 & 4 & 4 & 3 & 1 & 3 & 1 & 2 & 0 \end{bmatrix}$$

link tersebut didapatkan dari irisan anggota *Neighbor*, misal diketahui $Neighbor\ [P1] = \{P3, P4, P10\}$ dan $Neighbor\ [P2] = \{P3, P4, P5, P6, P9\}$ maka $Neighbor\ [P1] \cap Neighbor\ [P2] = \{P3, P4\}, n = 2$. Dengan demikian, didapatkan $Link\ (P1, P2)$ adalah 2, dan seterusnya. Setelah didapatkan matriks $Link\ B$, langkah selanjutnya adalah melakukan penggabungan kelompok dengan menghitung nilai *Goodness measure* menggunakan Persamaan 2.26. Sebelum menghitung nilai *Goodness Measure*, akan dihitung nilai $f(\theta)$ sebagai berikut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$f(\theta) = \frac{1 - 0{,}07}{1 + 0{,}07} = 0{,}960784$$

$$G(P1, P2) = \frac{2}{(1+1)^{1+2(0{,}961)} - (1)^{1+2(0{,}961)} - (1)^{1+(0{,}961)}} = 0{,}358635$$

$$G(P1, P3) = \frac{4}{(1+1)^{1+2(0{,}961)} - (1)^{1+2(0{,}961)} - (1)^{1+2(0{,}961)}} = 0{,}717270744$$

Dan seterusnya hingga $G(P9, P10)$ dan didapatkan matriks *Goodnees Measure* atau *Local Heap* sebagai berikut.

$$GM = \begin{array}{c} P1 \\ P2 \\ P3 \\ P4 \\ P5 \\ P6 \\ P7 \\ P8 \\ P9 \\ P10 \end{array} \begin{bmatrix} 0 & 0{,}359 & 0{,}717 & 0{,}717 & 0{,}179 & 0 & 0{,}179 & 0 & 0{,}359 & 0{,}717 \\ 0{,}359 & 0 & 0{,}717 & 0{,}717 & 0{,}538 & 0{,}538 & 0{,}179 & 0{,}179 & 0{,}717 & 0{,}538 \\ 0{,}717 & 0{,}717 & 0 & 1{,}076 & 0{,}359 & 0{,}179 & 0{,}179 & 0 & 0{,}717 & 0{,}717 \\ 0{,}717 & 0{,}717 & 1{,}076 & 0 & 0{,}359 & 0{,}179 & 0{,}179 & 0 & 0{,}717 & 0{,}717 \\ 0{,}179 & 0{,}538 & 0{,}358 & 0{,}358 & 0 & 0{,}538 & 0{,}538 & 0{,}259 & 0{,}179 & 0{,}538 \\ 0 & 0{,}538 & 0{,}179 & 0{,}179 & 0{,}538 & 0 & 0{,}359 & 0{,}359 & 0{,}179 & 0{,}179 \\ 0{,}179 & 0{,}179 & 0{,}179 & 0{,}179 & 0{,}538 & 0{,}359 & 0 & 0{,}359 & 0 & 0{,}538 \\ 0 & 0{,}179 & 0 & 0 & 0{,}359 & 0{,}359 & 0{,}359 & 0 & 0 & 0{,}179 \\ 0{,}359 & 0{,}717 & 0{,}717 & 0{,}717 & 0{,}179 & 0{,}179 & 0 & 0 & 0 & 0{,}359 \\ 0{,}717 & 0{,}538 & 0{,}717 & 0{,}717 & 0{,}538 & 0{,}179 & 0{,}538 & 0{,}179 & 0{,}359 & 0 \end{bmatrix}$$

Setelah didapatkan nilai *Goodness measure* atau *Local Heap*, langkah selanjutnya adalah menentukan *Global Heap* yang didapatkan dari nilai maksimum setiap baris dari *Local Heap*. Sehingga, didapatkan nilai *Global Heap* yang ditunjukkan dengan matriks berukuran 10×1 sebagai berikut.

$$GH = \begin{array}{c} P1 \\ P2 \\ P3 \\ P4 \\ P5 \\ P6 \\ P7 \\ P8 \\ P9 \\ P10 \end{array} \begin{bmatrix} 0{,}717 \\ 0{,}717 \\ \mathbf{1{,}076} \\ \mathbf{1{,}076} \\ 0{,}538 \\ 0{,}538 \\ 0{,}538 \\ 0{,}359 \\ 0{,}717 \\ 0{,}717 \end{bmatrix}$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Setelah didapatkan nilai *Global Heap*, langkah selanjutnya adalah menggabungkan objek yang memiliki *Global Heap* terbesar. Berdasarkan matriks GH diatas, diketahui bahwa nilai *Global Heap* terbesar dimiliki oleh $P3$ dan $P4$, dengan demikian, $P3$ dan $P4$ dikelompokkan menjadi 1 kelompok. Berikut adalah nilai *Neighbor* baru dengan penggabungan $P3$ dan $P4$.

$Neighbor\ [P1] = \{P1, P3, P4, P10\}$

$Neighbor\ [P2] = \{P2, P3, P4, P5, P6, P9\}$

$Neighbor\ [P3, P4] = \{P1, P2, P3, P4, P9, P10\}$

$Neighbor\ [P5] = \{P2, P5, P6, P7, P10\}$

$Neighbor\ [P6] = \{P2, P5, P6, P8\}$

$Neighbor\ [P7] = \{P5, P7, P8, P10\}$

$Neighbor\ [P8] = \{P6, P7, P8\}$

$Neighbor\ [P9] = \{P2, P3, P4, P9\}$

$Neighbor\ [P10] = \{P1, P3, P4, P5, P7, P10\}$

Sama halnya dengan langkah sebelumnya, dari nilai *Neighbor* didapatkan jumlah *link* antar objek satu dengan yang lainnya sebagai berikut.

$$C = \begin{array}{r} P1 \\ P2 \\ P3, P4 \\ P5 \\ P6 \\ P7 \\ P8 \\ P9 \\ P10 \end{array} \begin{bmatrix} 0 & 2 & 4 & 1 & 0 & 1 & 0 & 2 & 4 \\ 2 & 0 & 4 & 3 & 3 & 1 & 1 & 4 & 3 \\ 4 & 4 & 0 & 2 & 1 & 1 & 0 & 4 & 4 \\ 1 & 3 & 2 & 0 & 3 & 3 & 2 & 1 & 3 \\ 0 & 3 & 1 & 3 & 0 & 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 3 & 2 & 0 & 2 & 0 & 3 \\ 0 & 1 & 0 & 2 & 2 & 2 & 0 & 0 & 1 \\ 2 & 4 & 4 & 1 & 1 & 0 & 0 & 0 & 2 \\ 4 & 3 & 4 & 3 & 1 & 3 & 1 & 2 & 0 \end{bmatrix}$$

Setelah didapatkan matriks *link* C, langkah selanjutnya adalah menentukan *Goodness measure* atau *Local Heap* dengan perhitungan sebagai berikut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$G(P1,P2) = \frac{2}{(1+1)^{1+2(0,961)} - (1)^{1+2(0,961)} - (1)^{1+(0,961)}} = 0,358635$$

$$G(P1,(P3,P4)) = \frac{4}{(1+2)^{1+2(0,961)} - (1)^{1+2(0,961)} - (2)^{1+2(0,961)}} = 0,247001321$$

Dan seterusnya hingga didapatkan $G(P9,P10)$ dan didapatkan matriks *Goodnees Measure* atau *Local Heap* sebagai berikut.

$$GM = \begin{array}{c} P1 \\ P2 \\ P3,P4 \\ P5 \\ P6 \\ P7 \\ P8 \\ P9 \\ P10 \end{array} \begin{bmatrix} 0 & 0,359 & 0,247 & 0,179 & 0 & 0,179 & 0 & 0,359 & 0,717 \\ 0,359 & 0 & 0,247 & 0,538 & 0,538 & 0,179 & 0,179 & 0,717 & 0,538 \\ 0,247 & 0,247 & 0 & 0,126 & 0,062 & 0,062 & 0 & 0,247 & 0,247 \\ 0,179 & 0,538 & 0,126 & 0 & 0,538 & 0,538 & 0,359 & 0,179 & 0,538 \\ 0 & 0,538 & 0,062 & 0,538 & 0 & 0,359 & 0,359 & 0,179 & 0,179 \\ 0,179 & 0,179 & 0,062 & 0,538 & 0,359 & 0 & 0,359 & 0 & 0,538 \\ 0,359 & 0,179 & 0 & 0,359 & 0359 & 0,359 & 0 & 0 & 0,179 \\ 0 & 0,717 & 0,247 & 0,179 & 0,179 & 0 & 0 & 0 & 0,359 \\ 0,717 & 0,539 & 0,247 & 0,538 & 0,179 & 0,538 & 0,179 & 0,359 & 0 \end{bmatrix}$$

Setelah didapatkan nilai *Goodness measure* atau *Local Heap*, langkah selanjutnya adalah menentukan *Global Heap* yang didapatkan dari nilai maksimum setiap baris dari *Local Heap*. Sehingga, didapatkan nilai *Global Heap* yang ditunjukkan dengan matriks berukuran 10×1 sebagai berikut.

$$GH = \begin{array}{c} P1 \\ P2 \\ P3,P4 \\ P5 \\ P6 \\ P7 \\ P8 \\ P9 \\ P10 \end{array} \begin{bmatrix} \mathbf{0,717} \\ \mathbf{0,717} \\ 0,247 \\ 0,538 \\ 0,538 \\ 0,538 \\ 0,359 \\ \mathbf{0,717} \\ \mathbf{0,717} \end{bmatrix}$$

Setelah didapatkan nilai *Global Heap*, langkah selanjutnya adalah menggabungkan objek yang memiliki *Global Heap* terbesar. Berdasarkan matriks GH diatas, diketahui bahwa nilai *Global Heap* terbesar dimiliki oleh $P1, P2, P9,$ dan $P10$. Dengan demikian, oleh $P1, P2, P9,$ dan $P10$ dikelompokkan menjadi 1

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kelompok. Berikut adalah nilai *Neighbor* baru dengan penggabungan $P1, P2, P9$, dan $P10$.

$Neighbor\ [P1, P2, P9, P10] = \{P1, P2, P3, P4, P5, P6, P7, P9, P10\}$

$Neighbor\ [P3, P4] = \{P1, P2, P3, P4, P9, P10\}$

$Neighbor\ [P5] = \{P2, P5, P6, P7, P10\}$

$Neighbor\ [P6] = \{P2, P5, P6, P8\}$

$Neighbor\ [P7] = \{P5, P7, P8, P10\}$

$Neighbor\ [P8] = \{P6, P7, P8\}$

Sama halnya dengan langkah sebelumnya, dari nilai *Neighbor* didapatkan jumlah *link* antar objek satu dengan yang lainnya sebagai berikut.

$$D = \begin{array}{c} P1, P2, P9, P10 \\ P3, P4 \\ P5 \\ P6 \\ P7 \\ P8 \end{array} \begin{bmatrix} 0 & 6 & 5 & 3 & 3 & 2 \\ 6 & 0 & 2 & 1 & 1 & 0 \\ 5 & 2 & 0 & 3 & 3 & 2 \\ 3 & 1 & 3 & 0 & 2 & 2 \\ 3 & 1 & 3 & 2 & 0 & 2 \\ 2 & 0 & 2 & 2 & 2 & 0 \end{bmatrix}$$

Setelah didapatkan matriks *link* D, langkah selanjutnya adalah menentukan *Goodness measure* atau *Local Heap* dengan perhitungan sebagai berikut.

$$G\big((P1, P2, P9, P10), (P3, P4)\big) = \frac{6}{(4+2)^{1+2(0,961)} - (4)^{1+2(0,961)} - (2)^{1+(0,961)}} = 0{,}048900212$$

$$G\big((P1, P2, P9, P10), P5\big) = \frac{5}{(4+1)^{1+2(0,961)} - (4)^{1+2(0,961)} - (1)^{1+2(0,961)}} = 0{,}096580729$$

Dan seterusnya hingga didapatkan $G(P7, P8)$ dan didapatkan matriks *Goodnees Measure* atau *Local Heap* sebagai berikut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$D = \begin{array}{c} P1, P2, P9, P10 \\ P3, P4 \\ P5 \\ P6 \\ P7 \\ P8 \end{array} \begin{bmatrix} 0 & 0{,}0489 & 0{,}0966 & 0{,}0579 & 0{,}0579 & 0{,}0386 \\ 0{,}0489 & 0 & 0{,}1235 & 0{,}0618 & 0.0618 & 0 \\ 0{,}0966 & 0{,}1235 & 0 & 0{,}538 & 0{,}538 & 0{,}359 \\ 0{,}0579 & 0{,}0618 & 0{,}538 & 0 & 0{,}359 & 0{,}359 \\ 0{,}0579 & 0.0618 & 0{,}538 & 0{,}359 & 0 & 0{,}359 \\ 0{,}0386 & 0 & 0{,}359 & 0{,}359 & 0{,}359 & 0 \end{bmatrix}$$

Setelah didapatkan nilai *Goodness measure* atau *Local Heap*, langkah selanjutnya adalah menentukan *Global Heap* yang didapatkan dari nilai maksimum setiap baris dari *Local Heap*. Sehingga, didapatkan nilai *Global Heap* yang ditunjukkan dengan matriks berukuran 10×1 sebagai berikut.

$$GH = \begin{array}{c} P1, P2, P9, P10 \\ P3, P4 \\ P5 \\ P6 \\ P7 \\ P8 \end{array} \begin{bmatrix} 0{,}137 \\ 0{,}159 \\ \mathbf{0,641} \\ \mathbf{0,641} \\ \mathbf{0,641} \\ 0{,}428 \end{bmatrix}$$

Setelah didapatkan nilai *Global Heap*, langkah selanjutnya adalah menggabungkan objek yang memiliki *Global Heap* terbesar. Berdasarkan matriks GH diatas, diketahui bahwa nilai *Global Heap* terbesar dimiliki oleh $P5, P6$, dan $P7$. Dengan demikian $P5, P6,$ dan $P7$ dikelompokkan menjadi 1 kelompok. Berikut adalah nilai *Neighbor* baru dengan penggabungan P5, P6, dan P7.

$Neighbor\ [P1, P2, P9, P10] = \{P1, P2, P3, P4, P5, P6, P7, P9, P10\}$

$Neighbor\ [P3, P4] = \{P1, P2, P3, P4, P9, P10\}$

$Neighbor\ [P5, P6, P7] = \{P2, P5, P6, P7, P8, P10\}$

$Neighbor\ [P8] = \{P6, P7, P8\}$

Karena peneliti membatasi sebanyak 4 klaster, maka iterasi berhenti dan didapatkan pengelompokan yaitu, kelompok 1 terdiri dari 4 anggota yaitu Pacitan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

($P1$), Tulungagung ($P2$), Kota Blitar ($P9$) dan Kota Surabaya ($P10$). Kelompok 2 terdiri 2 anggota dari Lumajang ($P3$) dan Situbondo ($P4$). Kelompok 3 terdiri dari 3 anggota, yaitu Sidoarjo ($P5$), Nganjuk ($P6$), Gresik ($P7$) dan kelompok 4 terdiri dari 1 anggota yaitu Tuban. Pengklasteran 38 Kabupaten/Kota menggunakan *ensemble* ROCK dihitung menggunakan sistem dengan hasil pengklasteran seperti pada Tabel 4.49.

Tabel 4.49 Hasil klaster metode ROCK ($\theta = 0{,}02$)

Kelompok	Hasil Klaster
Kelompok 1	Pacitan, Blitar, Kediri, Pasuruan, Ngawi, Kota Kediri, Tuban, Kota Blitar, Kota Mojokerto, Kota Madiun
Kelompok 2	Ponorogo, Tulungagung, Malang, Sidoarjo, Nganjuk, Madiun, Magetan, Kota Malang, Kota Surabaya, Kota Batu
Kelompok 3	Trenggalek, Jember, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Bangkalan, Sampang, Pamekasan, Sumenep, Kota Probolinggo
Kelompok 4	Lumajang, Mojokerto, Jombang, Bojonegoro, Lamongan, Gresik, Kota Pasuruan.

F. Karakteristik Hasil Klaster Metode ROCK

Berdasarkan hasil klaster yang terbentuk, selanjutnya akan dilakukan uji signifikansi menggunakan uji *Wilks Lambda* untuk data numerik dan uji *Kruskal Wallis* untuk data kategorik. Hasil uji *Wilks Lambda* berdasarkan Tabel yang pada lampiran 8, menunjukan bahwa nilai signifikansi sebesar 0,048 < 0,05. Sehingga H_0 ditolak yang menunjukkan bahwa hasil klaster penetapan daerah tertinggal di Jawa Timur berdasarkan indikator KPDT bertipe numerik secara *multivariate* memiliki perbedaan yang signifikan. Sedangkan untuk uji signifikansi pada data kategorik menggunakan uji *Kruskal Wallis* dapat dilihat pada Tabel 4.50.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.50 Hasil Uji *Kruskal Wallis*

Keterangan	Variabel							
	x_{28}	x_{29}	x_{30}	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}
Chi-Square	7,206	7,944	3,194	3,244	2,753	4,838	2,078	0,866
Df	3	3	3	3	3	3	3	3
Asymp. Sig.	0,056	**0,047**	**0,036**	**0,036**	**0,043**	**0,018**	0,056	0,834

Berdasarkan Tabel 4.50, apabila dilihat dari variabel $x_{29}, x_{30,}x_{31}, x_{32}$, dan x_{33}, menunjukkan bahwa terdapat perbedaan antar klaster yang terbentuk, karena 5 variabel tersebut memiliki nilai signifikansi kurang dari (0,05). Namun, ketika dilihat berdasarkan variabel x_{28}, x_{34}, , x_{35} dan x_{36}, menunjukkan bahwa tidak ada perbedaan antar klaster yang terbentuk, karena 3 variabel tersebut memiliki nilai signifikansi lebih dari (0,05).

Berdasarkan Peraturan Menteri Desa, Pembangunan Daerah Tertinggal Dan Transmigrasi Nomor 3 Tahun 2016 Pasal 17 dalam Menentukan Arah Ketertinggalan, diketahui bahwa masing-masing indikator yang telah ditetapkan oleh KPDT memiliki dua arah ketertinggalan. Dua arah tersebut terdiri dari arah positif dan negatif (Republik Indonesia, 2016). Arah positif menandakan bahwa semakin tinggi nilai indikator tersebut, maka semakin buruk juga keadaan suatu daerah tersebut. Sedangkan arah negatif menandakan bahwa semakin rendah nilai indikator tersebut, maka akan semakin buruk pula keadaan suatu daerah tersebut. Berikut akan disajikan statistik deskripstif (rata-rata) indikator dengan skala numerik tiap-tiap klaster dan arah ketertinggalannya pada Tabel 4.51.

Tabel 4.51 Statistik Deskripstif (Rata-rata) Indikator Skala Numerik Anggota Tiap Klaster

Indikator	Arah Ketertinggalan	Klaster 1	Klaster 2	Klaster 3	Klaster 4
Presentase Penduduk Miskin (persen)	Positif	10,14	8,991	**16,12**	12,14571
Pengeluaran penduduk perkapita (ribu rupiah)	Negatif	10 408,43	11304,12	**8618,577**	10051,7

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.51 Statistik Deskripstif (Rata-rata) Indikator Skala Numerik Anggota Tiap Klaster (Lanjutan)

Indikator	Arah Ketertinggalan	Klaster 1	Klaster 2	Klaster 3	Klaster 4
Presentase jumlah desa terkena bencana gempa	Positif	**9,02**	4,400129	1,661975	0
Presentase jumlah desa terkena bencana tanah longsor	Positif	8,69	**12,02233**	9,332308	4,149003
Presentase jumlah desa terkena bencana banjir	Positif	9,57	11,02762	**16,22358**	24,12029
Presentase desa dengan terkena bencana lainnya	Positif	0,61	0,73766	**4,018237**	1,237932
Presentase kawasan hutan lindung	Negatif	8,95	9,87059	14,30402	**0,82165**
Presentase desa yang melakukan pelestarian lahan kritis	Negatif	77,04	73,20741	84,98853	**73,20521**
Presentase desa yang mempunyai konflik	Positif	**1,34**	1,263825	0,997666	1,992424
Angka Harapan hidup (tahun)	Negatif	71,94	72,072	**68,57917**	70,93143
Rata-rata lama sekolah (tahun)	Negatif	8,21	8,131	**5,811667**	7,454286
Angka melek huruf (tahun)	Negatif	12,57	12,958	**12,07083**	12,56143
Jarak ibu kota Kabupaten ke Ibu Kota provinsi (km)	Positif	146,56	113,3	**147,0833**	71,71429
Jumlah desa dengan kemudahan akses ke pelayanan keamanan	Negatif	**44,22**	75,8	56,83333	90,57143
Jumlah desa dengan permukaan jalan terluas aspal	Negatif	**142,00**	216,2	194,5	278,4286
Jumlah desa dengan permukaan jalan terluas diperkeras	Positif	18,56	9,2	**24,75**	22
Jumlah desa dengan permukaan jalan terluas tanah	Positif	1,33	0,5	**1,666667**	0,857143
Jumlah desa dengan permukaan jalan terluas lainnya	Negatif	**0**	0,1	**0**	**0**
Jumlah desa mempunyai pasar tanpa bangunan permanen	Positif	11,78	27,8	**28,66667**	16
Jumlah prasarana kesehatan per 1000 penduduk	Negatif	**0,00**	0,009	0,00325	0,007143

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.51 Statistik Deskripstif (Rata-rata) Indikator Skala Numerik Anggota Tiap Klaster (Lanjutan)

Indikator	Arah Ketertinggalan	Klaster 1	Klaster 2	Klaster 3	Klaster 4
Jumlah dokter per 1000 penduduk	Negatif	0,08	0,1549	**0,073583**	0,120857
Jumlah SMA per 1000 penduduk	Negatif	0,02	0,0448	**0,038417**	0,041
Presentase rumah tangga pengguna listrik (persen)	Negatif	99,49	99,706	**97,775**	98,83571
Presentase rumah tangga pengguna telepon (persen)	Negatif	88,13	88,987	**83,135**	87,12
Presentase rumah tangga pengguna air bersih	Negatif	18,26	26,57229	**9,069813**	14,19978
Derajat Desentralisasi Fiskal	Negatif	12,14	18,03763	**10,26706**	27,92596
Presentase Desa Tertinggal	Positif	0,61	0,881	**5,2975**	1,061429

Keterangan : = Berpotensi tertinggal

Selanjutnya, untuk modus dari data bertipe kategorik untuk tiap-tiap klaster dapat dilihat pada tabel 4.52.

Tabel 4.52 Statistik Deskriptif (Modus) Indikator Skala Kategorik Anggota Tiap Klaster

Indikator	Klaster 1	Klaster 2	Klaster 3	Klaster 4
Daerah perbatasan	Tidak	Tidak	**Ya**	**Ya**
Kepemilikan pulau	Tidak	Tidak	Tidak	Tidak
Daerah pasca konflik	Tidak	Tidak	**Ya**	**Ya**
Daerah rawan pangan	**Ya**	**Ya**	**Ya**	**Ya**
Daerah rawan bencana tanah longsor	Tinggi	Tinggi	Tinggi	Tidak
Daerah rawan bencana banjir	Tinggi	Tinggi	Tinggi	Tinggi
Daerah rawan bencana gempa bumi	Tidak	**Ya**	**Ya**	Tidak
Daerah rawan bencana tsunami	Tidak	Tidak	Tidak	Tidak
Daerah rawan bencana letusan gunung	Tidak	Tidak	Tidak	Tidak

Keterangan : = Berpotensi tertinggal

Berdasarkan Tabel 4.50, diketahui bahwa rata-rata indikator presentase penduduk miskin memiliki arah ketertingalan positif. Oleh karena itu, semakin tinggi presentase penduduk miskin suatu daerah, maka akan semakin buruk juga kondisi Kabupaten/Kota dari anggota klaster. Sehingga Kabupaten/Kota yang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

tergolong kedalam klaster tersebut memiliki potensi yang tinggi menjadi daerah tertinggal. Klaster 3 memiliki rata-rata presentase penduduk miskin sebesar 16,117% yang menunjukkan angka paling tinggi jika dibandingkan klaster lainnya. Sehingga, klaster 3 memiliki potensi sebagai daerah tertinggal yang tinggi. Selanjutnya untuk indikator pengeluaran penduduk perkapita memiliki arah ketertinggalan negatif. Oleh karena itu, semakin rendah pengeluaran penduduk perkapita suatu daerah, maka akan semakin buruk juga kondisi Kabupaten/Kota dari anggota klaster. Klaster 3 memiliki rata-rata pengeluaran penduduk perkapita sebesar 9212,364 ribu yang menunjukkan angka paling redah jika dibandingkan klaster lainnya. Sehingga, klaster 3 memiliki potensi sebagai daerah tertinggal yang tinggi, dan seterusnya hingga indikator presentase desa tertinggal.

Berdasarkan Tabel 4.50, dapat diketahui bahwa klaster 1 memiliki potensi ketertinggalan sebanyak 6 indikator. Klaster 2 memiliki potensi ketertinggalan sebanyak 1 indikator. Klaster 3 memiliki potensi ketertinggalan sebanyak 19 indikator. Sedangkan klaster 4 memiliki potensi ketertinggalan sebanyak 3 indikator. Selanjutnya untuk potensi ketertinggalan ditinjau dari variabel yang bertipe kategorik berdasarkan Tabel 4.51, terlihat bahwa untuk klaster 1 memiliki potensi ketertinggalan sebanyak 1 indikator. Klaster 2 memiliki potensi ketertinggalan sebanyak 2 indikator. Klaster 3 memiliki potensi ketertinggalan sebanyak 4 indikator. Sedangkan klaster 4 memiliki potensi ketertinggalan sebanyak 3 indikator. Oleh karena itu, urutan klaster untuk Kabupaten/Kota tertinggal, berkembang, mandiri dan maju berturut-turut yaitu klaster 3, klaster 1

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

klaster 4 dan klaster 2. Tabel 4.53 adalah anggota dari masing-masing klaster mulai dari tertinggal hingga maju.

Tabel 4.53 Anggota klaster daerah tertinggal, berkembang, mandiri dan maju

Klaster	Anggota
Tertinggal	Trenggalek, Jember, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Bangkalan, Sampang, Pamekasan, Sumenep, Kota Probolinggo
Berkembang	Pacitan, Blitar, Kediri, Pasuruan, Ngawi, Kota Kediri, Tuban, Kota Blitar, Kota Mojokerto, Kota Madiun
Mandiri	Lumajang, Mojokerto, Jombang, Bojonegoro, Lamongan, Gresik, Kota Pasuruan.
Maju	Ponorogo, Tulungagung, Malang, Sidoarjo, Nganjuk, Madiun, Magetan, Kota Malang, Kota Surabaya, Kota Batu

Berdasarkan Tabel 4.53, diketahui bahwa Kabupaten/Kota yang termasuk kedalam daerah tertinggal adalah Trenggalek, Jember, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Bangkalan, Sampang, Pamekasan, Sumenep, dan Kota Probolinggo, hal ini sejalan dengan penetapan daerah tertinggal yang dilakukan oleh Pemerintah. Berdasarkan Peraturan Presiden Nomor 131 Tahun 2015 diketahui bahwa Kabupaten/Kota di Provinsi Jawa Timur yang termasuk kedalam daerah tertinggal adalah Bangkalan, Sampang, Bondowoso dan Situbondo. Dimana, 4 Kabupaten tersebut masuk kedalam hasil cluster yang tercantum dalam Tabel 4.53. Selanjutnya, untuk visualisasi dan pemetaan penetapan daerah tertinggal pada peta dapat dilihat pada Gambar 4.2.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Gambar 4.1 Pemetaan Daerah Tertinggal Provinsi Jawa Timur Tahun 2015

G. Evaluasi Penetapan Daerah Tertinggal Tahun 2015

Berdasarkan Tabel 4.52 yang dibahas sebelumnya diketahui bahwa terjadi perubahan daerah yang ditetapkan sebagai daerah tertinggal. Terlihat bahwa pada Tabel 4.52 wilayah yang termasuk kedalam penetapan daerah tertinggal tahun 2015 yaitu Trenggalek, Jember, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Bangkalan, Sampang, Pamekasan, Sumenep, Kota Probolinggo.

Pada penetapan tahun 2015, Kabupaten Trenggalek tergolong sebagai daerah tertinggal dengan tingkat kemiskinan yang cukup tinggi yaitu 13,10%. Kemiskinan di Trenggalek terus meningkat dari tahun sebelumnya, yaitu dari 267.274 jiwa menjadi 272.792 jiwa. Sama halnya dengan berita yang termuat dalam www.bangsaonline.com, yang menyebutkan bahwa penduduk miskin di Trenggalek terus mengalami peningkatan (Rossa, 2016).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Kabupaten selanjutnya yang ditetapkan sebagai daerah tertinggal pada tahun 2015 yaitu Kabupaten Jember. Hal ini dapat dilihat dari data pengeluaran penduduk perkapita yang rendah, yaitu Rp. 8.226.660 hal inilah yang menjadikan Kabupaten Jember tergolong sebagai daerah tertinggal. Berdasarkan berita yang termuat dalam Prosalina.com menyebutkan bahwa terdapat 20 kecamatan di Kabupaten Jember yang masih tergolong tertinggal (Prosalina, 2013). Selain itu, Kabupaten Jember memiliki presentase penduduk miskin yang cukup tinggi yaitu 11,28%. Hal ini termuat dalam FaktualNews.com yang menyebutkan bahwa buruknya pengelolaan dana desa, mengakibatkan tingkat kemiskinan di Jember tinggi (Hatta, 2014).

Kabupaten selanjutnya yang ditetapkan sebagai daerah tertinggal pada tahun 2015 yaitu Kabupaten Banyuwangi. Pada dasarnya, Kabupaten Banyuwangi memiliki karaketistik daerah yang cukup baik. Berbeda dengan Kabupaten yang telah ditetapkan tertinggal lainnya. Indikator yang perlu diperbaiki adalah infrastruktur daerah khususnya pengguna listrik. Berdasarkan berita yang termuat dalam NasionalTempo.com, menyebutkan bahwa masih ada sekitar 100 desa terpencil di Banyuwangi yang belum di aliri listrik (Ningstyas, 2011). Selain listrik, terdapat beberapa desa di Banyuwangi yang belum memiliki sarana pendidikan yang memadai, terbukti dari berita yang termuat dalam Pojoksatu.id, menyebutkan bahwa untuk pergi kesekolah saja, mereka harus mempertaruhkan nyawa untuk pergi ke sekolah, pasalnya daerah yang sangat terpencil menjadikan tidak ada angkutan umum yang dapat dijangkau (Subekti, 2015).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Kabupaten selanjutnya yang ditetapkan sebagai daerah tertinggal pada tahun 2015 yaitu Kabupaten. Salah satu penyebab Probolinggo masuk kedalam daerah tertinggal adalah tingkat kemiskinan yang tinggi. Bahkan lebih tinggi daripada Kabupaten Jember dan Trenggalek, yaitu sebesar 20,44%. Hal ini termuat dalam www.SurabayaPagi.com yang menyebutkan bahwa Kabupaten Probolinggo menduduki peringkat ketiga sebagai Kabupaten termiskin se-Jawa Timur (Surabayapagi.com, 2012). Sama halnya dengan Kabupaten Probolinggo yang mengalami masalah kemiskinan, Kota Probolinggo juga mengalaminya meskipun tidak separah di Kabupaten Probolinggo. Hal ini termuat dalam RadarBromo.co.id yang menyebutkan bahwa terdapat 2000 rumah warga miskin di Kota Probolinggo tidak *tercover* Bantuan Pangan Non Tunai (BPNT) (Arifin, 2014)

Bondowoso, Situbondo, Bangkalan dan Sampang merupakan Kabupaten yang tergolong sebagai daerah tertinggal. Hal ini sejalan dengan keputusan Presiden tentang penetapan daerah tertinggal yang menyebutkan bahwa keempat Kabupaten tersebut digolongkan sebagai daerah tertinggal (Republik Indonesia, 2014) (Yunita, 2015)(Humas, 2015)(Lumanauw, 2015). Jika dilihat dari indikator yang telah ditetapkan KPDT, memang keempat Kabupaten tersebut memiliki arah ketertinggalan yang besar. Seperti Kabupaten Sampang yang memiliki tingkat presentase penduduk miskin tertinggi di Jawa Timur 25,80%, selain itu rata-rata sekolah, angka melek huruf, jumlah rumah sakit terendah di Jawa Timur menjadikan Kabupaten ini tergolong sebagai daerah tertinggal. Sedangkan Kabupaten Bangkalan memiliki tingkat presentase desa tertinggal tertinggi di Jawa Timur. Selanjutnya Kabupaten Situbondo menduduki Kabupaten yang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

sering terkena bencana dengan presentase sebesar 16,912%. Selanjutnya Kabupaten Bondowoso memiliki tingkat angka harapan hidup dan pengguna telepon terendah di Jawa Timur. Beberapa hal itulah yang menjadikan Bondowoso termasuk kedalam daerah tertinggal.

Salain Bangkalan dan Sampang, Kabupaten Madura lainnya yang tergolong tertinggal adalah Sumenep dan Pamekasan. Hal yang menjadi pokok perhatian Kabupaten Sumenep adalah perekonomian yang buruk dan infrastruktur yang tidak layak (Syaiful Anwar & Arif, 2017). Berdasarkan berita yang termuat dalam infopublik.id menyebutkan bahwa Kabupaten Sumenep masih berada dibawa garis kemiskinan selama 5 tahun terakhir, yaitu dari tahun 2013-2015 (Yanto, 2018). Selain presentase penduduk miskin yang tinggi, permasalahan ekonomi Kabupaten Sumenep lainnya yaitu tingkat pengeluran perkapita Kabupaten Sumenep terendah di Jawa Timur yaitu sebesar Rp.7.143.080. Selain itu, permasalahan infrastruktur Kabupaten Sumenep yaitu pasokan listrik yang kurang memadai. Terlihat dari presentase penduduk pengguna listrik terendah di Jawa Timur adalah Kabupaten Sumenep. Untuk Kabupaten Pamekasan memiliki tingkat desentralisasi fiskal terendah ketiga setelah Sampang dan Bangkalan yaitu sebesar 7,082%. Hal ini juga termuat dalam mediamadura.com yang menyebutkan bahwa serapan APBD Pamekasan tergolong rendah (Rifqi, 2016). Selain itu, Kabupaten Pamekasan menduduki presentase penduduk miskin nomor 4 se-Jawa Timur setelah Sampang, Bangkalan dan Sumenep.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

H. Prediksi Penetapan Daerah Tertinggal Tahun 2020

Prediksi penetapan daerah tertinggal tahun 2020 mendatang merupakan suatu bentuk upaya untuk memberikan gambaran dan masukan bagi Pemerintah dalam mengambil kebijakan yang disesuaikan dengan karakteristik dari daerahnya masing-masing. Data yang digunakan adalah data survey tahun 2018 yang akan dianalisis menggunakan metode *ensemble* ROCK seperti yang dilakukan pada sub bab sebelumnya. Tahapan awal yang harus dilakukan adalah analisis faktor baik untuk data numerik maupun data kategorik. Tabel 4.54 dan 4.55 merupakan hasil analisis faktor untuk data numerik dan kategorik dengan pengujian asumsi dan hasil *Loading Factor* seperti yang tercantum dalam Lampiran 7.

Tabel 4.54 Hasil analisis faktor data numerik

Faktor	Indikator	Nilai *Loading*
1	Rata-rata lama sekolah (x_{11})	0,947
2	Jumlah SMA per 1000 penduduk (x_{22})	0,784
3	Jumlah desa dengan permukaan jalan terluas tanah (x_{17})	0,251
4	Presentase desa terkena bencana gempa (x_3)	0,745
5	Jumlah desa dengan kemudahan akses ke pelayanan keamanan (x_{14})	0,745
6	Presentase jumlah desa terkena banjir (x_5)	0,420
7	Jumlah desa dengan permukaan jalan terluas lainnya(x_{18})	0,322

Tabel 4.55 Hasil analisis faktor data kategorik

Faktor	Indikator	Nilai *Loading*
1	Daerah rawan bencana tanah longsor (x_{32})	0,818
2	Daerah pasca konflik (x_{30})	0,874
3	Daerah perbatasan (x_{28})	0,867

Setelah didapatkan hasil analisis faktor untuk data numerik dan data kategorik, langkah selanjutnya adalah melakukan analisis klaster untuk data numerik menggunakan metode *K-Means*. Dalam hal ini akan dilakukan pengelompokan data menjadi 2,3 dan 4 klaster . Pengelompokan terbaik adalah dengan melihat nilai *Pseudo F-Statistic* dan *R-Square* yang paling besar. Tabel

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4.56 adalah perbandingan nilai *Pseudo F-Statistic* dan *R-Square* dari 2,3 dan 4 klaster.

Tabel 4.56 Perbandingan nilai *R-Square* dan *Pseudo F-Statistic* masing-masing pengelompokan

Kelompok Klaster	*R-Square*	*Pseudo F-Statistic*
Kelompok 2	0,5680428	47,34160
Kelompok 3	0,7443780	50,96045
Kelompok 4	0,8258974	53,76237

Berdasarkan Tabel 4.55, diketahui bahwa pengelompokan yang paling optimum menggunakan metode *K-Means* adalah kelompok 4 dengan nilai *R-Square* sebesar 82,59% dan nilai *Pseudo F-Statistic* sebesar 53,76237. Hasil pengelompokan 4 klaster dapat dilihat pada Tabel 4.57.

Tabel 4.57 Hasil klaster 4 kelompok menggunakan *K-Means*

Kabupaten	Hasil Klaster	Kabupaten	Hasil Klaster
Kab.Pacitan	2	Kab.Magetan	1
Kab.Ponorogo	1	Kab.Ngawi	4
Kab. Trenggalek	1	Kab.Bojonegoro	1
Kab.Tulungagung	1	Kab. Tuban	3
Kab.Blitar	1	Kab. Lamongan	3
Kab.Kediri	3	Kab. Gresik	1
Kab.Malang	1	Kab.Bangkalan	4
Kab.Lumajang	1	Kab. Sampang	4
Kab. Jember	1	Kab.Pamekasan	4
Kab. Banyuwangi	1	Kab. Sumenep	1
Kab. Bondowoso	1	Kota Kediri	4
Kab.Situbondo	4	Kota Blitar	4
Kab. Probolinggo	1	Kota Malang	4
Kab.Pasuruan	1	Kota Probolinggo	4
Kab. Sidoarjo	3	Kota Pasuruan	4
Kab.Mojokerto	1	Kota Mojokerto	4
Kab.Jombang	3	Kota Madiun	4
Kab.Nganjuk	1	Kota Surabaya	4
Kab.Madiun	1	Kota Batu	4

Selanjutnya untuk analisis klaster data kategorik akan digunakan metode *K-Modes*. Sama halnya dengan analisis klaster menggunakan metode *K-Means*, pengelompokan dilakukan menjadi 2,3 dan 4 klaster. Pemilihan klaster terbaik dilihat dari hasil perhitungan akurasi nilai r yang terbesar. Tabel 4.58 adalah

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

perbandingan hasil perhitungan akurasi nilai r dari pengelompokan 2, 3 dan 4 klaster.

Tabel 4.58 Perbandingan nilai r masing-masing pengelompokan

Kelompok Klaster	r
Kelompok 2	0,6666667
Kelompok 3	0,7894736
Kelompok 4	0,8245614

Berdasarkan Tabel 4.57, diketahui bahwa pengelompokan yang paling optimum menggunakan metode *K-Modes* adalah kelompok 4 dengan nilai r sebesar 0,8245614 atau 82,46%. Hasil pengelompokan 4 klaster dapat dilihat pada Tabel 4.59.

Tabel 4.59 Hasil klaster 4 kelompok menggunakan *K-Modes*

Kabupaten	Hasil Klaster	Kabupaten	Hasil Klaster
Kab.Pacitan	1	Kab.Magetan	1
Kab.Ponorogo	1	Kab.Ngawi	1
Kab. Trenggalek	2	Kab.Bojonegoro	2
Kab.Tulungagung	3	Kab. Tuban	2
Kab.Blitar	2	Kab. Lamongan	4
Kab.Kediri	2	Kab. Gresik	4
Kab.Malang	2	Kab.Bangkalan	4
Kab.Lumajang	1	Kab. Sampang	4
Kab. Jember	2	Kab.Pamekasan	4
Kab. Banyuwangi	2	Kab. Sumenep	4
Kab. Bondowoso	1	Kota Kediri	3
Kab.Situbondo	1	Kota Blitar	3
Kab. Probolinggo	2	Kota Malang	3
Kab.Pasuruan	2	Kota Probolinggo	4
Kab. Sidoarjo	3	Kota Pasuruan	4
Kab.Mojokerto	3	Kota Mojokerto	3
Kab.Jombang	3	Kota Madiun	3
Kab.Nganjuk	2	Kota Surabaya	4
Kab.Madiun	1	Kota Batu	1

Berdasarkan Tabel 4.57 dan 4.59 didapatkan hasil klaster optimum dari 2 metode tersebut masing-masing pengelompokan sebanyak 4 klaster. Kedua hasil klaster tersebut digabung dan dianggap sebagai data baru yang memiliki skala kategorik, hasil gabungan data tersebut selanjutnya akan dianalisis menggunakan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

metode *ensemble* ROCK. Sama halnya dengan penetapan daerah tertinggal sebelumnya, dalam hal ini akan digunakan 10 *threshold* dengan percobaan pengelompokan data menjadi 3 dan 4 klaster. Pemilihan klaster terbaik didasarkan dengan nilai $ratio\ sw$ dan sb terkecil. Tabel 4.61 Nilai $ratio\ sw$ dan sb dari masing-masing *threshold* untuk pengelompokan 3 dan 4 klaster

Tabel 4.61 Nilai $ratio\ sw$ dan sb dari masing-masing *threshold* untuk pengelompokan 3 dan 4 klaster

Threshold	3 Klaster	4 Klaster
0,1	0,7582026	0,5840034
0,15	0,617584	0,5804819
0,2	**0,4320012**	0,6196247
0,25	0,5308498	0,5935667
0,3	15.272.234	0,7116942
0,01	0,7144611	10.654.771
0,02	0,7845632	0,6013504
0,03	0,5305423	0,8132498
0,04	0,762686	**0,3873984**
0,05	0,5059842	0,6727432

Berdasarkan Tabel 4.59, diketahui bahwa nilai *threshold* terkecil yaitu 0,3873984. Dengan demikian pengelompokan yang paling optimum adalah 4 klaster dengan menggunakan nilai *threshold* sebesar 0,04. Hasil pengelompokan 4 klaster dengan *threshold* 0,04 dapat dilihat pada tabel 4.62.

Tabel 4.62 Hasil klaster metode ROCK ($\theta = 0{,}04$)

Kelompok	Hasil Klaster
Kelompok 1	Pacitan, Tulungagung, Blitar, Jember, Mojokerto, Jombang, Magetan, Kota Blitar, Kota Madiun, Kota Probolinggo.
Kelompok 2	Kediri, Banyuwangi, Nganjuk, Madiun, Bojonegoro, Gresik, Kota Kediri, Kota Pasuruan, Kota Mojokerto.
Kelompok 3	Trenggalek, Bondowoso, Situbondo, Probolinggo, Tuban, Bangkalan, Pamekasan, Sumenep, Sampang.
Kelompok 4	Ponorogo, Malang, Lumajang, Pasuruan, Sidoarjo, Ngawi, Lamongan, Kota Malang, Kota Surabaya, Kota Batu.

Berdasarkan hasil klaster yang terbentuk, selanjutnya akan dilakukan uji signifikansi menggunakan uji *Wilks Lambda* untuk data numerik dan uji *Kruskal*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Wallis untuk data kategorik. Hasil uji *Wilks Lambda* berdasarkan Tabel yang pada lampiaran 8, menunjukan bahwa nilai signifikansi sebesar 0,037 < 0,05. Sehingga H_0 ditolak yang menunjukkan bahwa hasil klaster penetapan daerah tertinggal di Jawa Timur berdasarkan indikator KPDT bertipe numerik secara *multivariate* memiliki perbedaan yang signifikan. Sedangkan untuk uji signifikansi pada data kategorik menggunakan uji *Kruskal Wallis* dapat dilihat pada Tabel 4.63.

Tabel 4.63 Hasil Uji Kruskal Wallis

Keterangan	Variabel							
	x_{28}	x_{29}	x_{30}	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}
Chi-Square	7,206	7,944	3,194	3,244	2,753	4,838	2,078	0,866
Df	3	3	3	3	3	3	3	3
Asymp. Sig.	**0,044**	**0,016**	**0,078**	0,751	**0,045**	**0,048**	0,623	**0,275**

Berdasarkan Tabel 4.63, apabila dilihat dari variabel $x_{28}, x_{29}, x_{30,} x_{32}, x_{33}$, dan x_{36}, menunjukkan bahwa terdapat perbedaan antar klaster yang terbentuk, karena 6 variabel tersebut memiliki nilai signifikansi kurang dari (0,05). Namun, ketika dilihat berdasarkan variabel x_{31}, x_{34}, dan x_{35}, menunjukkan bahwa tidak ada perbedaan antar klaster yang terbentuk, karena 3 variabel tersebut memiliki nilai signifikansi lebih dari (0,05).

Selanjutnya, dari hasil klaster yang terbentuk, akan dilakukan analisis terkait karakteristik masing-masing anggota klaster yang terbentuk. Berikut akan disajikan statistik deskriptif (rata-rata) indikator dengan skala numerik tiap-tiap klaster dan arah ketertinggalannya pada Tabel 4.64.

Tabel 4.64 Statistik Deskriptif (Rata-rata) Indikator Skala Numerik Anggota Tiap Klaster

Indikator	Arah Ketertinggalan	Klaster 1	Klaster 2	Klaster 3	Klaster 4
Presentase Penduduk Miskin (persen)	Positif	10,061071	11,317737	**16,10839**	9,3064009
Pengeluaran penduduk perkapita (ribu rupiah)	Negatif	11230,333	10963,5	**9382,889**	11651,7

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.64 Statistik Deskriptif (Rata-rata) Indikator Skala Numerik Anggota Tiap Klaster (Lanjutan)

Indikator	Arah Ketertinggalan	Klaster 1	Klaster 2	Klaster 3	Klaster 4
Presentase jumlah desa terkena bencana gempa	Positif	**7,696983**	2,0924952	1,0812273	4,7986011
Presentase jumlah desa terkena bencana tanah longsor	Positif	**14,92308**	4,7375212	12,268514	14,260464
Presentase jumlah desa terkena bencana banjir	Positif	14,894436	**18,80293**	14,800775	17,425333
Presentase desa dengan terkena bencana lainnya	Positif	4,3425851	2,713742	**5,573127**	1,2831262
Presentase kawasan hutan lindung	Negatif	17,519057	**13,85731**	23,400075	15,255145
Presentase desa yang melakukan pelestarian lahan kritis	Negatif	37,89342	38,076521	**35,846539**	42,870441
Presentase desa yang mempunyai konflik	Positif	1,921463	1,4446526	1,012992	**1,9605049**
Angka Harapan hidup (tahun)	Negatif	71,807778	71,678	**68,92**	72,001
Rata-rata lama sekolah (tahun)	Negatif	8,1577778	7,909	**6,112222**	8,069
Angka melek huruf (tahun)	Negatif	13,135556	13,104	**12,61778**	13,438
Jarak ibu kota Kabupaten ke Ibu Kota provinsi (km)	Positif	**155,8889**	113,6	138,88889	93
Jumlah desa dengan kemudahan akses ke pelayanan keamanan	Negatif	**52,666667**	63,3	53,444444	67,2
Jumlah desa dengan permukaan jalan terluas aspal	Negatif	200,33333	**187,9**	199,22222	238,3
Jumlah desa dengan permukaan jalan terluas diperkeras	Positif	6,6666667	**29,1**	9,8888889	15,2
Jumlah desa dengan permukaan jalan terluas tanah	Positif	0,4444444	0,8	**1**	0,5
Jumlah desa dengan permukaan jalan terluas lainnya	Negatif	**0**	0,1	0,2222222	0,1
Jumlah desa mempunyai pasar tanpa bangunan permanen	Positif	22,666667	28,5	25,777778	**38,5**
Jumlah prasarana kesehatan per 1000 penduduk	Negatif	0,0086667	0,0087	**0,004**	0,0172

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.64 Statistik Deskriptif (Rata-rata) Indikator Skala Numerik Anggota Tiap Klaster (Lanjutan)

Indikator	Arah Ketertinggalan	Klaster 1	Klaster 2	Klaster 3	Klaster 4
Jumlah dokter per 1000 penduduk	Negatif	0,1101111	0,1124	**0,064**	0,1684
Jumlah SMA per 1000 penduduk	Negatif	**0,027556**	0,0328	0,048	0,0525
Presentase rumah tangga pengguna listrik (persen)	Negatif	99,718889	99,759	**99,27333**	99,827
Presentase rumah tangga pengguna telepon (persen)	Negatif	88,256667	88,82	**85,37778**	90,004
Presentase rumah tangga pengguna air bersih	Negatif	17,960883	**5,628601**	6,1073584	21,428628
Derajat Desentralisasi Fiskal	Negatif	18,724793	19,657659	**13,36464**	25,314406
Presentase Desa Tertinggal	Positif	0,6277778	1,911	**5,223333**	1,353

Keterangan : = Berpotensi tertinggal

Selanjutnya, untuk modus dari data bertipe kategorik untuk tiap-tiap klaster dapat dilihat pada tabel 4.65.

Tabel 4.65 Statistik Deskriptif (Modus) Indikator Skala Kategorik Anggota Tiap Klaster

Indikator	Klaster 1	Klaster 2	Klaster 3	Klaster 4
Daerah perbatasan	Tidak	Ya	Ya	Ya
Kepemilikan pulau	Tidak	Tidak	Tidak	Tidak
Daerah pasca konflik	Tidak	Tidak	Ya	Tidak
Daerah rawan pangan	Ya	Ya	Ya	Ya
Daerah rawan bencana tanah longsor	Tidak	Tidak	Tinggi	Tidak
Daerah rawan bencana banjir	Tinggi	Tinggi	Tinggi	Tinggi
Daerah rawan bencana gempa bumi	Ya	Tidak	Ya	Tidak
Daerah rawan bencana tsunami	Tidak	Tidak	Tidak	Tidak
Daerah rawan bencana letusan gunung	Tidak	Tidak	Tidak	Tidak

Berdasarkan Tabel 4.64, diketahui bahwa rata-rata indikator presentase penduduk miskin memiliki arah ketertingalan positif. Oleh karena itu, semakin tinggi presentase penduduk miskin suatu daerah, maka akan semakin buruk juga kondisi Kabupaten/Kota dari anggota klaster. Sehingga Kabupaten/Kota yang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

tergolong kedalam klaster tersebut memiliki potensi yang tinggi menjadi daerah tertinggal. Klaster 3 memiliki rata-rata presentase penduduk miskin sebesar 16,10839% yang menunjukkan angka paling tinggi jika dibandingkan klaster lainnya. Sehingga, klaster 3 memiliki potensi sebagai daerah tertinggal yang tinggi. Selanjutnya untuk indikator pengeluaran penduduk perkapita memiliki arah ketertinggalan negatif. Oleh karena itu, semakin rendah pengeluaran penduduk perkapita suatu daerah, maka akan semakin buruk juga kondisi Kabupaten/Kota dari anggota klaster. Klaster 3 memiliki rata-rata pengeluaran penduduk perkapita sebesar 9382,889 ribu yang menunjukkan angka paling redah jika dibandingkan klaster lainnya. Sehingga, klaster 3 memiliki potensi sebagai daerah tertinggal yang tinggi, dan seterusnya hingga indikator presentase desa tertinggal.

Berdasarkan Tabel 4.64, dapat diketahui bahwa klaster 1 memiliki potensi ketertinggalan sebanyak 6 indikator. Klaster 2 memiliki potensi ketertinggalan sebanyak 5 indikator. Klaster 3 memiliki potensi ketertinggalan sebanyak 14 indikator. Sedangkan klaster 4 memiliki potensi ketertinggalan sebanyak 2 indikator. Selanjutnya untuk potensi ketertinggalan ditinjau dari variabel yang bertipe kategorik berdasarkan Tabel 4.51, terlihat bahwa untuk klaster 1 memiliki potensi ketertinggalan sebanyak 2 indikator. Klaster 2 memiliki potensi ketertinggalan sebanyak 2 indikator. Klaster 3 memiliki potensi ketertinggalan sebanyak 4 indikator. Sedangkan klaster 4 memiliki potensi ketertinggalan sebanyak 2 indikator. Oleh karena itu, urutan klaster untuk Kabupaten/Kota tertinggal, berkembang, mandiri dan maju berturut-turut yaitu klaster 3, klaster 1,

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

klaster 2 dan klaster 4. Tabel 4.66 adalah anggota dari masing-masing klaster mulai dari tertinggal hingga maju.

Tabel 4.66 Anggota klaster daerah tertinggal, berkembang, mandiri dan maju

Kelompok	Hasil Klaster
Tertinggal	Trenggalek, Bondowoso, Situbondo, Probolinggo, Tuban, Bangkalan, Pamekasan, Sumenep, Sampang.
Berkembang	Pacitan, Tulungagung, Blitar, Jember, Mojokerto, Jombang, Magetan, Kota Blitar, Kota Madiun, Kota Probolinggo.
Mandiri	Kediri, Banyuwangi, Nganjuk, Madiun, Bojonegoro, Gresik, Kota Kediri, Kota Pasuruan, Kota Mojokerto.
Maju	Ponorogo, Malang, Lumajang, Pasuruan, Sidoarjo, Ngawi, Lamongan, Kota Malang, Kota Surabaya, Kota Batu.

Selanjutnya, untuk visualisasi dan pemetaan penetapan daerah tertinggal pada peta dapat dilihat pada Gambar 4.2.

Gambar 4.2 Prediksi Pemetaan Daerah Tertinggal Provinsi Jawa Timur Tahun 2020

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

I. Evaluasi Prediksi Penetapan Daerah Tertinggal Tahun 2020

Berdasarkan Tabel 4.63, daerah yang diprediksi masuk kedalam daerah tertinggal pada penetapan tahun 2020 mendatang yaitu Trenggalek, Bondowoso, Situbondo, Probolinggo, Tuban, Bangkalan, Pamekasan, Sumenep, dan Sampang.

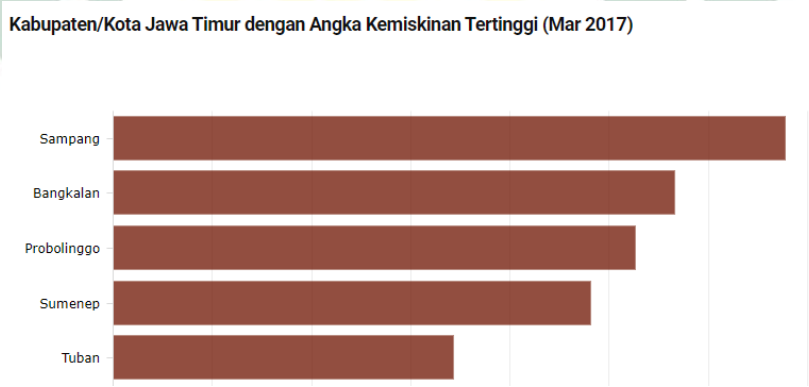
Kabupaten Banyuwangi diprediksi terlepas dari penetapan sebagai daerah tertinggal pada Tahun 2020 mendatang hal ini dipengaruhi oleh presentase penduduk miskin yang mengalami penurunan cukup signifikan dari tahun 2014 ke tahun 2017 yaitu sebesar 0,65%. Tidak hanya itu pengeluaran penduduk per kapita juga mengalami kenaikan yang cukup besar 10,20 %. Didukung lagi dengan kenaikan setiap indikator pada kriteria sumber daya manusia, yaitu 0,37% untuk angka harapan hidup, 3,49% untuk rata-rata lama sekolah dan 7,37% untuk angka melek huruf. Selain itu, Kabupaten Banyuwangi telah melakukan perbaikan infrastruktur dengan cukup baik. Dibuktikan dengan meningkatnya kualitas infrastruktur setiap indikator dari tahun 2014 ke tahun 2018. Dilihat dari kemampuan keuangan daerah yang diwakili dengan tingkat desentralisasi fiskal, menunjukkan bahwa Kabupaten Banyuwangi mengalami kenaikan keuangan daerah yang cukup tinggi, yaitu sebesar 2,67% dari kenaikan sebelumnya.

Selain Kabupaten Banyuwangi, Kabupaten yang keluar dari penetapan daerah tertinggal adalah Kabupaten Jember, hal ini disebabkan karena presentase penduduk miskin mengalami penurun sebesar 0,28 %. Disamping itu, rata-rata pengeluaran penduduk perkapita mangalami kenaikan yang cukup tinggi dari tahun 2014 yaitu sebesar 5,73%. Kenaikan yang signifikan juga terdapat pada kriteria sumber daya manusia, yaitu 1,09% untuk indikator angka harapan hidup,

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

7,64% untuk indikator rata-rata lama sekolah, 6,58% untuk indikator angka melek huruf. Kenaikan yang cukup signifikan pula terjadi pada indikator desentralisasi fiskal yang mengalami peningkatan sebesar 5,396 %.

Pada prediksi penetapan daerah tertinggal tahun 2020, diketahui bahwa Tuban masuk kedalam daerah tertinggal. Hal ini dapat dilihat dari Kabupaten Tuban yang mengalami peningkatan presentase penduduk miskin dari tahun penetapan sebelumnya. Berdasarkan berita yang termuat dalam SuaraBanyuurip.com (2019), tingkat kemiskinan Kabupaten Tuban menduduki peringkat kelima dari seluruh Kabupaten/Kota di Jawa Timur (Samian, 2019) . Berikut adalah grafik 5 Kabupaten/Kota yang menjadi kantong kemiskinan di Jawa Timur.



Gambar 4.3 Grafik 5 Peringkat teratas Kabupaten/Kota Miskin di Jawa Timur
Sumber : (KataData.co.id, 2018)

Presentase rumah tangga pengguna air bersih di Kabupaten Tuban menduduki angka yang sangat rendah sama seperti Kabupaten Trenggalek dan Situbondo.

Trenggalek, Bondowoso, Situbondo, Probolinggo, Pamekasan, Sumenep, Kota Probolinggo dan Sampang diprediksi masih menjadi daerah yang tergolong sebagai Kabupaten/Kota tertinggal untuk penetapan daerah tertinggal selanjutnya.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Karakteristik yang dimiliki masing-masing Kabupaten/Kota tersebut diantaranya yaitu, Kabupaten Trenggalek memiliki jumlah sarana prasarana kesehatan yang paling rendah diantara 37 Kabupaten/Kota lainnya. Kemudian Kabupaten Bondowoso memiliki angka harapan hidup dan pengguna telepon yang paling rendah dari 38 Kabupaten/Kota yang ada di Jawa Timur. Tidak hanya itu, Kabupaten Bondowoso memiliki jalan terluas tanah yang cukup besar. Kabupaten Sampang menjadi Kabupaten di Jawa Timur yang memiliki tingkat presentase penduduk miskin tertinggi. Selain itu, rata-rata lama sekolah yang sangat rendah menjadikan Kabupaten Sampang masuk kedalam daerah tertinggal. Kabupaten Pamekasan memiliki derajat desentralisasi fiskal yang paling rendah dari 38 Kabupaten/Kota di Jawa Timur. Kabupaten Sumenep merupakan Kabupaten yang memiliki Jumlah pengguna listrik dan air bersih terendah se-Jawa Timur dan Kota Probolinggo memiliki kawasan hutan lindung yang sangat kecil serta akses keamanan yang sulit menjadikan Kota Probolinggo masuk kedalam daerah tertinggal.

Seperti yang telah dijelaskan sebelumnya terkait karakteristik masing-masing daerah yang diprediksi tergolong sebagai daerah tertinggal untuk penetapan daerah tertinggal tahun 2020 mendatang, diharapkan bagi pihak Pemerintah untuk mengambil kebijakan terhadap daerahnya masing-masing yang disesuaikan dengan karakteristik daerahnya masing-masing.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB V
PENUTUP

A. Simpulan

Berdasarkan hasil dan pembahasan yang telah dijelaskan pada bab sebelumnya, simpulan yang dapat diambil adalah sebagai berikut.

1. Hasil *clustering* menggunakan metode *ensemble* ROCK didapatkan hasil klaster penetapan daerah tertinggal tahun 2015 terdiri dari Trenggalek, Jember, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Bangkalan, Sampang, Pamekasan, Sumenep, dan Kota Probolinggo. Selanjutnya untuk prediksi penetapan daerah tertinggal untuk tahun 2020 terdiri dari Trenggalek, Bondowoso, Situbondo, Probolinggo, Tuban, Pamekasan, Sumenep, Bangkalan dan Sampang.
2. Hasil klaster terbaik untuk penetapan daerah tertinggal tahun 2015 terdiri dari 4 klaster dengan nilai $ratio\ sw\ dan\ sb$ terkecil sebesar 0,3917473 dan nilai *threshold* optimum sebesar 0,02. Sedangkan hasil klaster terbaik untuk prediksi penetapan daerah tertinggal tahun 2020 terdiri dari 4 klaster dengan nilai $ratio\ sw\ dan\ sb$ terkecil sebesar 0,3873984 dan nilai *threshold* optimum sebesar 0,04.
3. Karakteristik dari masing-masing daerah yang ditetapkan sebagai daerah tertinggal yang perlu untuk dilakukan peningkatan untuk mengentaskan daerah tersebut dari ketertinggalan diantaranya yaitu :
   a. Kriteria kemampuan keuangan daerah, dengan indikator tingkat desentralisai fiskal yang rendah.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

b. Kriteria ekonomi, dengan indikator presentase penduduk miskin yang tinggi.
c. Kriteria sumber daya manusia, dengan indikator angka harapan hidup dan rata-rata lama sekolah yang rendah.
d. Kriteria infrastruktur, dengan indikator jumlah prasarana kesehatan, presentase rumah tangga pengguna listrik, presentase rumah tangga pengguna telepon dan presentase rumah tangga pengguna air bersih yang rendah.
e. Kriteria aksesibilitas, dengan kriteria akses kemudahan kepelayanan terdekat.
f. Kriteria karakteristik daerah, dengan kriteria presentase kawasan hutan lindung yang rendah

B. Saran

Saran yang dapat disampaikan untuk peneliti yang ingin melanjutkan kepenelitian selanjutnya adalah sebagai berikut.

a. Untuk penelitian selanjutnya, peneliti dapat menggunakan objek pengamatan yang lebih kecil dari penelitian ini, misalnya desa di Jawa Timur. Dengan demikian akan didapatkan informasi yang lebih detail terkait desa yang tergolong tertinggal di masing-masing Kabupaten/Kota di Jawa Timur, sehingga target pencapaian untuk pembangunan daerah akan tercapai lebih optimal.
b. Untuk penelitian selanjutnya, peneliti dapat menggunakan metode *ensemble* lainnya untuk membandingkan dengan metode *ensemble* ROCK.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Saran yang dapat disampaikan untuk Pemerintah Kabupaten/Kota yaitu hasil prediksi penetapan daerah tertinggal tahun 2020, dapat memberikan gambaran bagi Pemerintah Kabupaten/Kota sehingga mampu segera mengambil kebijakan yang disesuaikan dengan karakteristik masing-masing daerah guna untuk mengentaskan daerah tersebut dari ketertinggalan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR PUSTAKA

Agresti, A. (2009). An introduction to categorical data analysis (Second Edition). In *Statistics in Medicine* (Vol. 28). https://doi.org/10.1002/sim.3564

Alvionita, Sutikno, & Suharsono. (2017). Ensemble ROCK Methods and Ensemble SWFM Methods for Clustering of Cross Citrus Accessions Based on Mixed Numerical and Categorical Dataset. *IOP Publishing*. https://doi.org/10.1088/1755-1315/5

Arifin, J. (2014). 2 Ribu Warga Miskin Kota Probolinggo Belum Ter-Cover BPNT. Retrieved May 31, 2019, from https://radarbromo.jawapos.com/2014/01/26/2-ribu-warga-miskin-kota-probolinggo-belum-ter-cover-bpnt/

Arsham, H., & Lovric, M. (2011). Bartlett's Test. *International Encyclopedia of Statistical Science*, (December). https://doi.org/10.1007/978-3-642-04898-2

Badan Penelitian dan Pengembangan Pendidikan dan Pelatiihan dan Informasi Kementrian Desa Pembangunan Daerah Tertinggal dan Transformasi. (2017). *Buku Perkembangan Daerah Tertentu Daerah Rawan Pangan*. Jakarta: Badan Penelitian dan Pengembangan Pendidikan dan Pelatiihan dan Informasi Kementrian Desa Pembangunan Daerah Tertinggal dan Transformasi.

Badan Perencanaan Pembangunan Daerah Provinsi Jawa Timur. (2017). *Dokumentasi Hasil Pelaksanaan Pembangunan Provinsi Jawa Timur*. Surabaya: Badan Perencanaan Pembangunan Daerah Provinsi Jawa Timur.

Bagus, R. U. I. G. (2016). Mengapa memilih analisis faktor. *Universitas Dhyana Pura Bali*. https://doi.org/10.13140/RG.2.2.35757.28649

BAPPEDA. (2018). *Dokumentasi Hasil Pelaksanaan Pembangunan Provinsi Jawa Timur 2018*. Surabaya: Badan Perencanaan Pembangunan Daerah Provinsi Jawa Timur.

Ben-Hur, A., & Guyon, I. (2003). Detecting Stable Clusters Using Principal Component Analysis. *Functional Genomics*, (1), 159–182. https://doi.org/10.1385/1-59259-364-x:159

BPS. (n.d.-a). Sirusa - Sistem Rujukan Statistik - Angka Harapan Hidup. Retrieved April 15, 2019, from https://sirusa.bps.go.id/index.php?r=indikator/view&id=48

BPS. (n.d.-b). Sirusa - Sistem Rujukan Statistik - Angka Melek Huruf. Retrieved April 15, 2019, from https://sirusa.bps.go.id/index.php?r=indikator/view&id=313

BPS. (n.d.-c). Sirusa - Sistem Rujukan Statistik - Pengeluaran Penduduk Per Kapita. Retrieved April 15, 2019, from Badan Pusat Statistika website: https://sirusa.bps.go.id/index.php?r=indikator/view&id=197

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BPS. (n.d.-d). Sirusa - Sistem Rujukan Statistik - Presentase Penduduk Miskin. Retrieved April 15, 2019, from Badan Pusat Statistika website: https://sirusa.bps.go.id/index.php?r=indikator/view&id=18

BPS. (n.d.-e). Sirusa - Sistem Rujukan Statistik - Rata-rata Lama Sekolah. Retrieved April 15, 2019, from https://sirusa.bps.go.id/index.php?r=indikator/view&id=11

BPS Kabupaten Banyuwangi. (2018). *Kabupaten Banyuwangi Dalam Angka 2018*. Banyuwangi: Badan Pusat Statistika Kabupaten Banyuwangi.

BPS Kabupaten Lamongan. (2018). *Kabupaten Lamongan Dalam Angka 2018*. Lamongan: Badan Pusat Statistika Kabupaten Lamongan.

BPS Kabupaten Malang. (2018). *Kabupaten Malang Dalam Angka 2018*. Malang: Badan Pusat Statistika Kabupaten Malang.

BPS Kabupaten Mojokerto. (2018). *Kota Mojokerto Dalam Angka 2018*. Mojokerto: Badan Pusat Statistika Kota Mojokerto.

Bunkers, & James. (1996). Definition of Climate Regions in the Northern Plains Using an Objective Cluster Modification Technique. *J.Climate*, 130–146.

David, W., & Djamaris, A. R. (2018). *Metode Statistik Untuk Ilmu dan Teknologi Pangan*. Retrieved from http://repository.bakrie.ac.id/id/eprint/1255%0A

Djuraidah, A. (2009). Evaluasi Status Ketertinggalan Daerah Dengan Analisis Diskriminan. *Seminar Nasional Matematika Dan Pendidikan Matematika Jurusan Pendidikan Matematika FMIPA UNY*, 1–16. Retrieved from http://eprints.uny.ac.id/7073/1/S.21 Anik Djuraidah.pdf

Dutta, M., Mahanta, A. K., & Pujari, A. K. (2005). QROCK: A quick version of the ROCK algorithm for clustering of categorical data. *Pattern Recognition Letters*, *26*(15), 2364–2373. https://doi.org/10.1016/j.patrec.2005.04.008

Farchan, Fauzi, SE., M. (2016). Teknikal Manajemen Sumber Daya Manusia Strategik Sebuah Paradigma Pengukuran Kinerja. *Jurnal Pendidikan Dan Studi Islam*, *1*(1), 42–62.

Guha, S., Rastogi, R., & Shim, K. (2000). Rock: a robust clustering algorithm for categorical attributes. *Information Systems*, *25*(5), 345–366. https://doi.org/10.1016/S0306-4379(00)00022-3

Hair, J. F., Black, J. W. C., Babin, B. J., & Anderson, R. E. (2013). Multivariate Data Analysis Sevent Edition. In *Exploratory Data Analysis in Business and Economics*. https://doi.org/10.1007/978-3-319-01517-0_3

Haryono. (2017). *Metode SEM untuk Penelitian Manajemen dengan AMOS 22.00, LISREL 8.80 dan SMART-PLS 3.0*. Jakarta: PT. Luxima Metro Media.

Hatta, M. (2014). Angka Kemiskinan di Jember Tinggi, Akibat Buruknya Pengelolaan DD. Retrieved May 31, 2019, from https://faktualnews.co/2014/09/07/angka-kemiskinan-di-jember-akibat-buruknya-pengelolaan-dd/98159/

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

He, Z., Xu, X. i, & Deng, S. (2005). Clustering Mixed Numeric and Categorical Data : A Cluster Ensmable Approach. *Departement of Computer Science and Engineering Harbin Institute of Technology*.

Huang, Z. (1998). Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values. *Data Mining and Knowledge Discovery*, *2*(3), 283–304. https://doi.org/10.1023/A:1009769707641

Humas. (2015). 122 Daerah Ini Ditetapkan Pemerintah Sebagai Daerah Tertinggal 2015 – 2019. Retrieved May 31, 2019, from setkab.go.id website: https://setkab.go.id/122-daerah.ini-ditetapkan-pemerintah-sebagai-daerah-tertinggal-2015-2019/

Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis Sixth Edition* (Six). New Jearsey: Person Prentice Hall.

KataData.co.id. (2018). Di Mana Kantong Kemiskinan Jawa Timur? Retrieved May 31, 2019, from https://databoks.katadata.co.id/datapublish/2018/01/24/di-mana-kantong-kemiskinan-jawa-timur

Kota, A. R., Setiawan, A., & Linawati, L. (2012). Uji Independensi Statistik Bartlett Terhadap Nilai Saham Untuk Mengetahui Uji Independensi Statistik Bartlett Terhadap Nilai Saham Untuk Mengetahui Kebergantungan Saham-Saham Pada Beberapa Sektor Saham di BEJ. *Seminar Nasional Matematika Dan Pendidikan Matematika Jurusan Pendidikan Matematika FMIPA UNY*, (April).

Lumanauw, N. (2015). Ini 122 Kabupaten yang Ditetapkan sebagai Daerah Tertinggal. Retrieved May 31, 2019, from https://www.beritasatu.com/nasional/328882/ini-122-kabupaten-yang-ditetapkan-sebagai-daerah-tertinggal

Ningstyas, I. (2011). Seratus Desa Terpencil Banyuwangi Tanpa Listrik. Retrieved May 31, 2019, from https://nasional.tempo.co/read/136476/seratus-desa-terpencil-banyuwangi-tanpa-listrik/full&view=ok

Orpin, A. R., & Kostylev, V. E. (2006). Towards a statistically valid method of textural sea floor characterization of benthic habitats. *Marine Geology*, *225*(1–4), 209–222. https://doi.org/10.1016/j.margeo.2005.09.002

Prosalina. (2013). 20 Kecamatan Di Jember, Masuk Kategori Daerah Relatif Tertinggal. Retrieved May 31, 2019, from Prosalina.com website: http://www.prosalinaradio.com/2013/03/23/20-kecamatan-di-jember-masuk-kategori-daerah-relatif-tertinggal/

Reddy, M. . J., & B.Kavitha. (2012). Clustering the Mixed Numerical and Categorical Datasets Using Similarity Weight and Filter Method. *International Journal of Database Theory and Application*, *5*(1), 121–133.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Republik Indonesia. (2014). *Peraturan Pemerintah Republik Indonesia Nomor 78 Tahun 2014 Tentang Percepatan Pembangunan Daerah Tertinggal*.

Republik Indonesia. (2015). *Peraturan Presiden Republik Indonesia Nomor 131 Tahun 2015 Tentang Penetapan Daerah Tertinggal Tahun 2015-2019*.

Republik Indonesia. (2016). *Peraturan Menteri Desa, Pembangunan Daerah Tertinggal, Dan Transmigrasi Nomor 3 Tahun 2016 Tentang Petunjuk Teknis Penentuan Indikator Dalam Penetapan Daerah Tertinggal Secara Nasional*.

Rifqi. (2016). Serapan APBD 2016 di Pamekasan Rendah. Retrieved May 31, 2019, from https://mediamadura.com/2017/01/09/serapan-apbd-2016-di-pamekasan-rendah/

Rossa. (2016). Jumlah Penduduk Miskin di Trenggalek Meningkat. Retrieved May 31, 2019, from https://www.bangsaonline.com/berita/19698/jumlah-penduduk-miskin-di-trenggalek-meningkat

Samian. (2019). Kemiskinan Tuban Masih Bertengger Lima Besar di Jatim. Retrieved May 30, 2019, from Suara Banyuurip.com website: http://m.suarabanyuurip.com

Setiadi, D. H. (2018). *Pemetaan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Daerah Tertinggal Dengan Metode Data Campuran Ensamble ROCK dan SWFM*. Institut Teknologi Sepuluh Nopember.

Setiawan, B., & Muntaha, A. (2014). *Metode Penelitian Komunikasi*.

Sharma, S. (1996). *Applied multivariate tecnhiques*. New York: John Wiley & Sons Inc.

Sharma, S., & Yadav, R. L. (2013). Comparative Study of K-means and Robust Clustering. *International Journal of Advanced Computer Research ISSN : 2249-7277*, (3), 1–4.

Simarmata, I., Arma, A. J. A., & Arnita. (2013). *Aplikasi Analisis Faktor dengan Metode Principal Component Analysis dan Maximum Likelihood dalam Faktor-faktor yang Memengaruhi Pemberian Makanan Tambahan Pada Bayi Usia 0-6 Bulan di Desa Pematang Panjang Kecamatan Air Putih Kabupaten Batubara Tahun 2013*. *5*(21076041), 8–13.

Subekti, R. (2015). Anak-anak Banyuwangi Pertaruhkan Nyawa Demi Sekolah. Retrieved May 31, 2019, from https://pojoksatu.id/pojok-bibir/2015/04/01/anak-anak-banyuwangi-pertaruhkan-nyawa-demi-sekolah/

Sudipo Guha, Rajeev Rastogi, K. S. (1999). *A robust clustering algorithm for categorical.pdf.pdf*.

Surabayapagi.com. (2012). Probolinggo Termiskin Nomor 3 Di Jatim. Retrieved May 31, 2019, from www.Surabayapagi.com website: http://www.surabayapagi.com/read/92663/2012/12/28/Probolinggo_Termiskin_Nomor_3_Di_Jatim.html

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Susanti, E. L. (2016). Modul Metode Penelitian. In *Penelitian*. https://doi.org/10.1021/ol7029646

Syaiful Anwar, & Arif, M. (2017). Janji Politik Gus Ipul : Perbaiki Permasalahan Kemiskinan dan Infrastruktur Di Madura. Retrieved May 31, 2019, from http://pojokpitu.com/baca.php?idurut=57124&&top=1&ktg=J

Wiratama, S., Diartho, H. C., & Prianto, F. W. (2018). Analisis Pembangunan Wilayah Tertinggal di Provinsi Jawa Timur. *E-Journal Ekonomi Bisnis Dan Akuntansi*, *5*(1), 16. https://doi.org/10.19184/ejeba.v5i1.7726

Wulan, A. R. (2016). Pengertian Dan Esensi Konsep. Retrieved April 15, 2019, from http://file.upi.edu/Direktori/FPMIPA/JUR._PEND._BIOLOGI/ANA_RATNAWULAN/pengertian_asesmen.pdf

Yanto, N. (2018). Berikut Jumlah Penduduk Miskin Di Sumenep Selama 5 Tahun Terakhir. Retrieved May 31, 2019, from InfoPublik.id website: http://infopublik.id/kategori/nusantara/315716/berikut-jumlah-penduduk-miskin-di-sumenep-selama-5-tahun-terakhir?video=

Yunita, N. W. (2015). Jokowi Tetapkan 122 Kabupaten ini Daerah Tertinggal 2015-2019. Retrieved May 31, 2019, from Detiknews website: https://news.detik.com/berita/d-3092196/jokowi-tetapkan-122-kabupaten-ini-daerah-tertinggal-2015-2019

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id