



## Clustering Pemetaan Tingkat Kemiskinan di Provinsi Jawa Barat Menggunakan Algoritma K-Means

I Wayan Setya Adi Nugraha

Mahasiswa Program Studi Teknik Informatika Fakultas Ilmu Komputer  
Universitas Singaperbangsa

---

### Abstract

*Received: 21 Desember 2022* One of the problems Indonesia is still facing is the problem of poverty.  
*Revised: 23 Desember 2022* The issue of poverty is complex and multifaceted in society and has  
*Accepted: 26 Desember 2022* become a government development priority. The percentage of poverty  
in Indonesia at the beginning of the pandemic in March increased by 9.78% and in September it increased by 10.19%. West Java ranks first in  
the highest level of extreme poverty in Indonesia with a total of 17,856 cases, reported on [radarsukabumi.com](http://radarsukabumi.com). This study uses the clustering  
method with the k-means algorithm and mapping poverty-prone areas  
using QGIS. The results of grouping poverty-prone areas in West Java  
Province in 2015 to 2020 found 12 districts/cities not vulnerable, 14  
districts/cities prone and 1 district/city very vulnerable. The results of  
the clustering evaluation using the silhouette coefficient that is equal to  
0.55. The evaluation results fall into the category of medium structure  
with a reasonable interpretation of cluster placement.

**Keywords:** Clustering, Data Mining, K-Means, Python, Poverty

(\* Corresponding Author: [iwayan.16109@student.unsika.ac.id](mailto:iwayan.16109@student.unsika.ac.id)

**How to Cite:** Setya Adi Nugraha, I. W. (2023). Clustering Pemetaan Tingkat Kemiskinan di Provinsi Jawa Barat Menggunakan Algoritma K-Means. *Jurnal Ilmiah Wahana Pendidikan*, 9(2), 234-244. <https://doi.org/10.5281/zenodo.7567622>.

---

### PENDAHULUAN

Saat masa pandemi covid-19, Indonesia berada di fase krisis ekonomi dampak dari covid-19. Perusahaan bahkan banyak yang melakukan PHK massal kepada karyawannya, UMKM mengalami penurunan omset penghasilan karena adanya kebijakan PSBB dan PPKM yang membuat durasi berjualan menjadi terbatas. Dilansir dari [katadata.co.id](http://katadata.co.id) sebanyak 63,9% pelaku UMKM mengalami dampak penurunan omzet lebih dari 30%.

Dari salah satu faktor tersebut mengakibatkan meningkatnya angka kemiskinan akibat tidak adanya aktivitas ekonomi dan dampak negative yang ditimbulkannya bisa kita lihat pada grafik pada Gambar 1.2 dimana yang terdampak negative mencapai 82,9%. Perkembangan jumlah penduduk dari beberapa tahun sebelum pandemi mengalami penurunan sementara setelah terjadi pandemi mengalami kenaikan yang drastis. Persentase kemiskinan di Indonesia pada awal masa pandemi di bulan Maret mengalami kenaikan sebesar 9,78% dan di bulan september mengalami kenaikan sebesar 10,19%.

Provinsi Jawa Barat menjadi ranking pertama tingkat tertinggi kemiskinan ekstrim di Indonesia berdasarkan jumlah penduduk dengan jumlah kasus sebanyak 17.856 penduduk, dilansir di [radarsukabumi.com](http://radarsukabumi.com) (30/09/2021). Kemiskinan ekstrim ini merupakan persoalan yang besar karena tiap tahun selalu meningkat. Menurut Wakil Presiden Ma'ruf Amin saat memimpin rapat kerja membahas penanggulangan kemiskinan ekstrim meminta gubernur dan seluruh



bupati di wilayah prioritas bekerja keras memastikan agar seluruh rumah tangga miskin ekstrim di wilayahnya masing-masing mendapatkan seluruh program, baik program untuk pengurangan beban pengeluaran masyarakat maupun program pemberdayaan masyarakat. (setkab.go.id).

Menurut (Yacoub, 2012), menurut penelitiannya, kemiskinan adalah tentang pemenuhan kebutuhan hidup yang paling mendasar, dan kemiskinan merupakan masalah global yang dihadapi banyak negara, sehingga kemiskinan merupakan salah satu masalah mendasar. Menurut Bank Dunia, 2004, salah satu penyebab kemiskinan adalah kurangnya pendapatan dan kekayaan (income and wealth) untuk memenuhi kebutuhan dasar seperti pangan, sandang, papan, tingkat kesehatan dan pendidikan yang dapat diterima. Selain itu, kemiskinan juga dikaitkan dengan terbatasnya kesempatan kerja, masyarakat yang tergolong miskin (miskin) biasanya menganggur (unemployed), serta pendidikan dan kesehatan mereka umumnya tidak memadai.

Indikator-indikator kemiskinan yang digunakan secara umum adalah tingkat upah, pendapatan, konsumsi, mortalitas anak usia balita, imunisasi, kekurangan gizi anak, tingkat fertilitas, tingkat kematian ibu, harapan hidup rata-rata, tingkat penyerapan anak usia sekolah dasar, proporsi pengeluaran pemerintah untuk pelayanan kebutuhan dasar masyarakat, pemenuhan bahan pangan (kalori/protein), air bersih, perkembangan penduduk, melek huruf, urbanisasi, pendapatan per kapita, dan distribusi pendapatan. Tolok ukur kemiskinan bukan hanya hidup dalam kekurangan pangan dan tingkat pendapatan yang rendah, akan tetapi melihat tingkat kesehatan, pendidikan dan perlakuan adil dimuka hukum dan sebagainya (Adisasmita, 2005).

Pada penelitian sebelumnya pada tahun 2019 dilakukan penelitian *clustering* pada data informasi kemiskinan di Jawa Barat dengan menggunakan algoritma *k-means*. Pengujian yang dilakukan yaitu sebanyak 3 kali iterasi (Nugroho & Nico, 2019). Pada penelitian Achmad Bahauddin, et al. (2021) dilakukan analisis *clustering* kemiskinan menurut provinsi di Indonesia dengan menggunakan algoritma *k-means*. Data yang digunakan yaitu data kriminalitas dari Badan Pusat Statistik Nasional. Algoritma *k-means* merupakan metode yang sederhana yang memiliki kemampuan dalam mengelompokkan data dengan baik dan dalam jumlah yang cukup besar dengan waktu komputasi yang cepat dan efisien (Dewi et al., 2019). Pada penelitian Wahyu Saputro, et al (2020) dalam melakukan pengklasteran tindak pidana korupsi di Indonesia, dengan menggunakan algoritma *k-means* diperoleh 3 *cluster* berdasarkan kelompok tingkat kerawanan tindak korupsi yaitu tindak pidana korupsi paling tinggi, menengah dan cukup. Algoritma *k-means* digunakan karena memiliki kemampuan menggabungkan data yang besar dan dalam waktu yang relatif cepat dan efektif (Saputro et al., 2020).

Berdasarkan latar belakang di atas, maka penelitian ini akan melakukan pemetaan daerah kemiskinan di Provinsi Jawa Barat berdasarkan lokasi, jumlah data. Dengan melakukan analisis pemetaan daerah rawan kemiskinan di Provinsi Jawa Barat, diharapkan Pemerintah Daerah Provinsi Jawa Barat dapat menentukan keputusan yang tepat dalam mengatasi dan melakukan pencegahan kemiskinan yang terjadi. Berdasarkan dari akurasi dan nilai akhir yang baik dari metode yang digunakan pada penelitian sebelumnya, maka metode yang digunakan adalah metode *cluster analysis* menggunakan algoritma *k-means*,

algoritma juga mempunyai kemampuan ketelitian yang tinggi terhadap ukuran objek, dan relatif terukur dan efisien dalam pengolahan objek dalam jumlah yang besar (Bastian et al., 2018). Algoritma *k-means* ini digunakan untuk mengetahui jumlah kejadian di setiap kabupaten/kota. Dalam memvisualisasikan pemetaan pada penelitian ini akan dilakukan dengan bantuan QGIS.

## **METODE PENELITIAN**

Metodologi yang digunakan dalam penelitian ini yaitu metode CRISP-DM (*Cross Industry Standard Process for Data Mining*). Dalam metode CRISP-DM ini terdapat 6 tahapan yaitu pemahaman bisnis (*Business Understanding*), pemahaman data (*Data Understanding*), pengolahan data (*Data Preparation*), pemodelan (*Modeling*), evaluasi (*Evaluation*), dan yang terakhir yaitu penyebaran (*deployment*).

Proses clustering menggunakan algoritma k-means dilakukan pada fase Modelling. Berikut ini merupakan penjelasan dalam setiap tahapan dari metode CRISP-DM pada penelitian ini adalah:

### *1. Business Understanding*

Pada fase pertama dilakukan analisis masalah atau pemahaman terhadap masalah apa yang bisa diangkat dalam penelitian ini, fase ini terdapat tiga proses diantaranya:

- a. *Determine business objectives*, dilakukan pemahaman tujuan yang ingin dicapai dengan melakukan tahap wawancara kepada Badan Pusat Statistik.
- b. *Asses situation*, proses ini dilakukan dengan menganalisis fakta yang terjadi dilapangan, dengan mengetahui analisis apa yang sudah dilakukan di Pemerintah Provinsi Jawa Barat dan apakah sudah dilakukan pemetaan kemiskinan di Provinsi Jawa Barat.
- c. *Determine data mining goals*, proses ini yaitu dilakukan penentuan proses secara teknik yaitu menentukan metode *data mining* yang akan digunakan terhadap tujuan penelitian yang akan dicapai.

### *2. Data Understanding*

Pada fase kedua yaitu *data preparation* dilakukan pengumpulan data, mendeskripsikan atau menggambarkan data, kemudian explore data mana yang mungkin bermanfaat untuk Pemerintah Provinsi Jawa Barat, selanjutnya mengidentifikasi masalah yang akan dilakukan berkaitan dengan data yang dimiliki dan mempelajari data yang didapatkan yang akan digunakan dalam penelitian. Data didapatkan dari BPS mengenai data kemiskinan berdasarkan laporan yang ada di wilayah Provinsi Jawa Barat selama dua tahun yaitu 2015 sampai dengan 2020.

### *3. Data Preparation*

Pada tahap ini, dilakukan *preprocessing* dimana proses fase ini memiliki beberapa tahapan:

- a. *Data selection*, proses ini merupakan proses memilih data dan juga memilih atribut yang akan digunakan sesuai tujuan *data mining*.

- b. *Data preprocessing*, proses ini dilakukan pembersihan data atau *cleansing* data dengan mengatasi *outlier*, *noisy data* dan *missing value* yang ada. Pada tahap ini juga memastikan data memiliki kualitas yang baik.
- c. *Transformation*, proses ini yaitu proses pengelompokan atribut ke dalam data yang baru, kemudian mengintegrasikan data, mentransmisikan data yang sesuai dengan tujuan untuk kemudian diproses dalam *data mining*.

#### 4. *Modeling*

Fase *modelling* ini dilakukan proses penentuan model dan penerapan teknik *data mining* yang digunakan, algoritma dan juga *tools* yang akan digunakan yaitu menggunakan *jupyter notebook* dengan bahasa pemrograman *python*. Pada penelitian ini metode atau teknik pemodelan *data mining* yang akan digunakan yaitu metode *clustering* dengan algoritma *k-means* terhadap data kemiskinan di Provinsi Jawa Barat. Hasil pemodelan tersebut selanjutnya akan divisualisasikan dengan menggunakan QGIS.

#### 5. *Evaluation*

Pada fase *evaluation* ini, dilakukan analisis terhadap hasil dari proses pembelajaran data. Fase ini yaitu proses interpretasi hasil pemodelan *data mining* yang digunakan. Evaluasi dilakukan dengan menggunakan metode *Silhouette Coefficient*, yang dimana metode ini merupakan metode yang menguji kualitas dari klaster yang dihasilkan. Evaluasi ini dilakukan untuk mengetahui pemodelan yang diterapkan apakah sudah sesuai dan cocok diterapkan pada kasus penelitian ini serta sudah sesuai dengan tujuan yang ingin dicapai. Kemudian dari hasil evaluasi tersebut tentukan langkah-langkah berikutnya apakah bisa dilanjutkan ke tahap berikutnya atau diulang dari awal karena tidak sesuai dengan tujuan.

#### 6. *Deployment*

Pada fase keenam yaitu fase penyebaran dengan membuat laporan atau presentasi dari pengetahuan yang didapatkan dari hasil pemodelan dan evaluasi pada proses *data mining*. Hasil yang didapatkan diberikan kepada Pemerintah Provinsi Jawa Barat yang digunakan untuk menentukan keputusan yang tepat dalam mengatasi dan melakukan pencegahan kemiskinan terhadap hasil pemetaan daerah rawan kemiskinan di Provinsi Jawa Barat.

## HASIL DAN PEMBAHASAN

### Hasil Penelitian *Data Mining*

Hasil dari penelitian yang dilakukan adalah melakukan analisis pemetaan daerah rawan kemiskinan di Provinsi Jawa Barat dengan menggunakan algoritma *k-means clustering*. Hasil dari clustering tersebut dievaluasi dan divisualisasikan menggunakan *software* QGIS.

#### 1. *Pemahaman Bisnis (Business Understanding)*

Tahapan pemahaman bisnis berfokus pada pemahaman tujuan kebutuhan berdasarkan penilaian bisnis. Kemudian rencana tersebut diubah menjadi sebuah rencana awal *data mining* yang dirancang untuk mencapai tujuan. Tahapan pemahaman bisnis ada 3 yaitu:

- a. Menentukan Tujuan Bisnis (*Determine Business Objectives*)

Pada fase awal ini dilakukan observasi dengan melakukan pencarian data pada website *opendata* Jawa Barat untuk mencari tahu potensi yang dapat dilakukan dalam penelitian sesuai dengan kebutuhan agar dapat bermanfaat dan dapat digunakan. Fokus observasi ini pada data kemiskinan yang terjadi di Jawa Barat.

b. Menilai Situasi (*Asses the Situation*)

Setelah dilakukannya observasi, dihasilkan kesimpulan sebagai berikut:

1. Pemerintah Provinsi Jawa Barat memiliki data yang dapat dijadikan informasi ataupun pengetahuan. Data tersebut adalah data penyebaran jumlah penduduk miskin berdasarkan kabupaten/kota, dimana kedua data tersebut tersedia setiap tahunnya. Data yang diperoleh yaitu data dari tahun 2002 sampai dengan 2020. Tetapi atribut pada data tahun 2002 sampai dengan 2014 tidak sesuai dengan atribut pada tahun 2015 sampai dengan 2022. Maka, data yang akan digunakan yaitu data pada tahun 2015 sampai dengan 2020.
2. Kumpulan data setiap tahunnya belum diolah untuk dapat menghasilkan informasi yang lebih baik dan bermanfaat.
3. Bentuk penyampaian informasi mengenai kemiskinan di Jawa Barat yang ada di website Badan Pusat Statistika Jawa Barat hanya berupa list jumlah kemiskinan secara keseluruhan, belum ada bentuk visualisasi penyampaian informasi kepada masyarakat untuk mengetahui penyebaran data kemiskinan yang ada

c. Menentukan Tujuan Data Mining (*Determine the Data Mining Goals*)

Dalam mengatasi masalah hasil observasi dalam fakta yang terjadi di Pemerintah Provinsi Jawa Barat, maka tujuan *data mining* yang dapat dilakukan dalam penelitian ini yaitu melakukan pengklasteran daerah rawan kemiskinan dengan menggunakan metode *clustering* dengan menerapkan algoritma *k-means* yang kemudian untuk mempermudah masyarakat dalam mengetahui intensitas kejahatan yang ada maka akan divisualisasikan berupa peta dengan bantuan salah satu *software* sistem informasi geografi.

2. **Pemahaman Data (*Data Understanding*)**

Fase pemahaman data awal, mengenal dan memahami data yang dimiliki serta melakukan analisis apa yang bisa dilakukan pada data-data tersebut. Tahapan awal yaitu proses pengumpulan data yang didapatkan dari website portal *open data* Jawa Barat yaitu data kemiskinan 2015 sampai dengan 2020.

Data yang didapatkan dari portal open data Jawa Barat tersebut berupa data kemiskinan di Provinsi Jawa Barat berbentuk excel. Dimana terdapat 27 Kabupaten/Kota dalam 6 tahun. Dataset awal kemiskinan memiliki atribut nama provinsi, nama kabupaten kota, jumlah penduduk miskin, satuan dan tahun.

3. **Pengolahan Data (*Data Preparation*)**

*Data preparation* mencakup seluruh kegiatan dalam membangun *dataset* yang akan dimasukkan ke dalam pemodelan dari data mentah awal. *K-means* menjadi algoritma yang digunakan dalam mengolah pemodelan *dataset*. Adapun beberapa tahapan dalam *data preparation* yaitu *data selection*, *data preprocessing* dan *data transformation*.

#### a. *Data Selection*

Data yang diambil mengenai data kemiskinan, yaitu data jumlah penduduk miskin berdasarkan kabupaten/kota di Provinsi Jawa Barat. Dataset ini disimpan dalam bentuk *excel* atau *csv*. Lalu data dibersihkan dengan cara menghapus atribut-atribut yang tidak diperlukan pada penelitian ini. Data yang dipilih yaitu data yang berkaitan dengan kemiskinan di tahun 2015 sampai dengan 2020. Adapun atribut yang digunakan yaitu nama kabupaten/kota, jumlah penduduk miskin dan tahun.

#### b. *Data Preprocessing*

Pada proses ini dilakukan pembuatan dataset dari data mentah ke data yang siap digunakan pada proses pemodelan *data mining*. Hal yang dilakukan adalah menghapus atribut yang tidak digunakan seperti nama provinsi dan satuan. Kemudian melakukan perubahan pada jumlah penduduk miskin yang awalnya mempunyai satuan ribu jiwa menjadi satuan jiwa. Proses tersebut dapat dilihat dalam gambar berikut.

```
# Menghapus beberapa atribut yang tidak digunakan
clear_atribut = df.drop(['nama_provinsi', 'satuan'], axis=1)
```

#### c. *Data Transformation*

Tahap pengembangan set data yaitu proses transformasi data sesuai kebutuhan pada proses pemodelan. Pengembangan dataset baru ini dilakukan dengan mentransformasikan data dimana nilai-nilai tersebut dihitung dengan meringkas informasi. Pada proses pemodelan dibutuhkan jumlah data penduduk miskin dan dikelompokkan berdasarkan kabupaten/kota. Perubahan data dilakukan dengan mengubah yang tadinya data dalam satu kolom menunjukkan jumlah penduduk miskin tiap tahunnya, pada tahap ini mengubah jumlah penduduk miskin tersebut menjadi beberapa kolom yang dikelompokkan berdasarkan tahun. Kemudian menambahkan atribut jumlah yang menunjukkan jumlah penduduk miskin selama tahun 2015 s.d. 2020.

Transformasi data ini menghasilkan 27 data yang berisi nama kabupaten/kota di Jawa Barat yang terdapat pada kolom Kabupaten/Kota. Pada setiap kabupaten/kota terdapat data jumlah penduduk miskin sesuai tahunnya dari tahun 2015 hingga 2020 yang terletak pada kolom 2015, 2016, 2017, 2018, 2019 dan 2020. Pada kolom terakhir yaitu total merupakan nilai keseluruhan dari jumlah penduduk miskin di Jawa Barat. Pada tabel 3.3 berikut menunjukkan data jumlah penduduk miskin di Jawa Barat hasil dari proses transformasi data.

### 4. *Pemodelan (Modeling)*

*Modeling* merupakan fase yang secara langsung melibatkan teknik *data mining*, pemilihan teknik *data mining*, algoritma dan menentukan parameter dengan nilai yang optimal. *Clustering* merupakan teknik *data mining* yang digunakan pada penelitian ini. Adapun algoritma yang dipilih yaitu algoritma *k-means* yang nantinya akan dimodelkan pada data hasil transformasi. Hasil yang diperoleh dari teknik *clustering* dan algoritma *k-means* yaitu kelompok data yang digunakan untuk mengelompokkan daerah rawan penyebaran kemiskinan yang tidak rawan, rawan, hingga sangat rawan di Provinsi Jawa Barat. *Cluster* yang akan

dibuat adalah 3 cluster, penentuan *cluster* akan dibagi 3 yaitu *cluster* tidak rawan (C1), *cluster* rawan (C2), dan *cluster* sangat rawan (C3). Penentuan titik tengah atau *centroid* awal dilakukan dari data masing-masing atribut dengan mengambil nilai terkecil atau nilai minimum untuk *cluster* tidak rawan (C1), kemudian nilai rata-rata untuk *cluster* rawan (C2), dan nilai terbesar pada data atau maksimum untuk *cluster* sangat rawan (C3). Atribut yang menjadi acuan dalam pengklasteran yaitu keseluruhan dari kabupaten/kota.

Berikut merupakan proses *clustering k-means* menggunakan *python*.

```
# mulai proses kmeans nya
X = np.array(nilai_centroid_awal, np.float64)
km = KMeans(n_clusters=3, init=X, n_init=1).fit(clear_dataset)
centers = km.cluster_centers_
print(centers)

[[ 68075.          65575.          64825.          57108.33333333
   53900.          62203.33333333]
 [223685.71428571 210492.78571429 207392.78571429 179671.42857143
 168383.57142857 193437.14285714]
 [487100.          490800.          487300.          415000.
 395030.          465670.          ]]
```

Gambar Proses K-Means

Data diatas menunjukkan nilai centroid pada iterasi terakhir dalam proses pengklasteran k-means menggunakan *library* yang ada pada *python*. Untuk proses pembagian atau pengelompokan menjadi 3 cluster dengan menggunakan *python*.

```
# masukan nilai cluster ke dalam dataframe
df['cluster'] = km.labels_
a = km.labels_
mapping = {0:'Tidak Rawan', 1:'Rawan', 2:'Sangat Rawan'}
a = [mapping[i] for i in a]
df['keterangan'] = a

new_df = df[['kabupaten', 'cluster', 'keterangan']]
new_df
```

Gambar Proses Clustering

Hasil *clustering* data kemiskinan menggunakan *k-means* didapatkan untuk cluster daerah dengan tingkat kemiskinan tidak rawan atau C1 yaitu Kabupaten Ciamis, Kabupaten Sumedang, Kabupaten Purwakarta, Kabupaten Pangandaran, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Depok, Kota Cimahi, Kota Tasikmalaya, Kota Banjar. Kemudian untuk cluster daerah dengan tingkat kemiskinan rawan atau C2 yaitu Kabupaten Sukabumi, Kabupaten Cianjur, Kabupaten Bandung, Kabupaten Garut, Kabupaten Tasikmalaya, Kabupaten Kuningan, Kabupaten Cirebon, Kabupaten Majalengka, Kabupaten Indramayu, Kabupaten Subang, Kabupaten Karawang, Kabupaten Bekasi, Kabupaten Bandung Barat dan Kota Bekasi. Cluster ketiga yaitu daerah dengan tingkat

kemiskinan sangat rawan atau C3 yaitu Kabupaten Bogor.

### 5. *Evaluasi (Evaluation)*

Pada tahap evaluasi ini dilakukannya fase interpretasi terhadap hasil *data mining* yang dilakukan secara mendalam dengan tujuan agar hasil pada tahap *modeling* sesuai dengan sasaran yang ingin dicapai dalam tahap *business understanding* sebelumnya dan mengetahui sejauh mana tingkat kualitas dari pemodelan yang diterapkan.

#### a. *Evaluation Result*

Hasil pemodelan selanjutnya dilakukan evaluasi dengan menggunakan *silhouette coefficient*. Pengujian ini dilakukan untuk mengetahui kedekatan relasi antar objek dan seberapa jauh antar cluster terpisah, sehingga diketahui kualitas dari sebuah *cluster*. Jika nilai mendekati 1 atau bernilai positif dan  $ai$  mendekati 0, hal tersebut menghasilkan nilai maksimal yaitu 1 dengan  $ai=0$ , maka struktur dari klaster masuk kedalam kategori baik atau klaster berada pada klaster yang tepat. Kemudian jika nilai *silhouette coefficient* = 0 maka struktur klaster masuk kedalam kategori tidak jelas. Sedangkan jika nilai *silhouette coefficient* = -1 maka struktur dari klaster masuk kedalam kategori *overlapping*.

Berikut pengujian yang dilakukan dengan menggunakan *library* yang ada pada *python*.

```
from sklearn.metrics import silhouette_score

silhouette_score(clear_dataset, km.labels_)

0.5539929079261192
```

Gambar Proses evaluasi

Hasil pengujian menggunakan *silhouette coefficient* atau *silhouette score* yang ditunjukkan pada Gambar 4.5 menghasilkan nilai *index* sebesar 0,55. Kualitas *cluster* yang dihasilkan dari algoritma *k-means* ini termasuk ke dalam kategori *medium structure* dengan interpretasi penempatan *cluster* wajar.

#### b. *Determine Next Step*

Pada tahap ini yaitu tahap menentukan langkah selanjutnya yang akan diambil berdasarkan hasil pengujian. Terdapat dua pilihan dalam menentukan langkah selanjutnya yaitu kembali kepada tahapan pemahaman bisnis (*business understanding*) atau melanjutkan ke tahap akhir dalam metodologi *crisp-dm* adalah *deployment*. Tahap ini bisa dilanjutkan apabila hasil pengujian sudah sesuai dengan tujuan, jika tidak maka kembali kepada tahap awal (*business understanding*). Pada penelitian ini hasil pengujian didapatkan bahwa *Cluster* masuk kedalam kategori *medium structure*. Kategori ini dapat dikatakan merupakan klaster yang baik dan hasil penelitian sudah sesuai dengan tujuan awal, maka dapat dilanjutkan ke langkah selanjutnya.



*coefficient* didapatkan hasil yaitu 0,55 dan masuk dalam kategori medium strucutre dengan interpretasi penempatan klaster yang wajar.

3. Hasil dari pengklasteran pada proses *modeling* dilakukan visualiasi berupa peta tingkat kemiskinan yang divisualisasikan dengan perbedaan warna masing-masing daerah yang mana merupakan tingkat kerawanan kemiskinan

## DAFTAR PUSTAKA

- Astuti, D. (2019). Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma K-Means Clustering. *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, 1(2), 60–72. <https://doi.org/10.20895/inista.v1i2.71>
- Bahauddin, Achmad & Fatmawati, Agustina & Sari, Febrianti. (2021). Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means. *Jurnal Manajemen Informatika dan Sistem Informasi*. 4. 1. 10.36595/misi.v4i1.216.
- Bastian, A., Sujadi, H., & Febrianto, G. (n.d.). Penerapan Algoritma K-Means Clustering Analysis Pada Penyakit Menular Manusia (Studi Kasus Kabupaten Majalengka). 1, 26–32.
- Eka, R., Putera, P., Ramdani, F., & Rokhmawati, R. I. (2019). Evaluasi Tampilan Antarmuka QGIS dan MapWindow dengan Menggunakan Pendekatan Heuristic Evaluation ( Studi Kasus : Fungsi Geoprocessing Tools ). 3(11), 2850–2857.
- Fahri, M. U. (2020). Melihat Peta Penyebaran Pasien Covid-19 Dengan Kombinasi Qgis Dan Framework Laravel. *Jurnal Teknologi Terpadu*, 6(1), 25–30.
- Feblian, D., & Daihani, D. U. (2017). Implementasi Model Crisp-Dm Untuk Menentukan Sales Pipeline Pada Pt X. *Jurnal Teknik Industri*, 6(1), 1–12. <https://doi.org/10.25105/jti.v6i1.1526>
- Hapsari, D. P. T., & Widodo, E. (2017). Pengelompokan Daerah Rawan Kriminalitas di Indonesia Menggunakan Analisis K-Means Clustering. *Prosiding SI MaNIs (Seminar Nasional Integrasi Matematika Dan Nilai Islami)*, 1(1), 147–153.
- Irfiani, E., & Rani, S. S. (2018). Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita. *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)*, 6(4), 161. <https://doi.org/10.26418/justin.v6i4.29024>
- Muliono, R., & Sembiring, Z. (2019). Data Mining Clustering Menggunakan Algoritma K-Means Untuk Klasterisasi Tingkat Tridarma Pengajaran Dosen. 4(2), 2502–2714.
- Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square. *Jurnal Informatika: Jurnal Pengembangan IT*, 4(1), 20–24. <https://doi.org/10.30591/jpit.v4i1.1253>
- Paramartha, G. N. W., Ratnawati, D. E., & Widodo, A. W. (2017). Analisis Perbandingan Metode K-Means Dengan Improved Semi-Supervised Analisis Perbandingan Metode K-Means Dengan Improved Semi-Supervised K-

- Means Pada Data Indeks Pembangunan Manusia (IPM). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, Vol. 1(9), 813–824. <http://j-ptiik.ub.ac.id>
- Pramesti, D. F., Lahan, Tanzil Furqon, M., & Dewi, C. (2017). Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(9), 723–732. <https://doi.org/10.1109/EUMC.2008.4751704>
- Ratna Sari, Y., Sudewa, A., Ayu Lestari, D., & Ika Jaya, T. (2020). Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer.
- Supriyatna, A., & Mustika, W. P. (2018). Komparasi Algoritma Naive bayes dan SVM Untuk Memprediksi Keberhasilan Imunoterapi Pada Penyakit Kutil. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 2(2), 152. <https://doi.org/10.30645/j-sakti.v2i2.78>
- Suryadi, U. T., & Supriatna, Y. (2019). Sistem Clustering Tindak Kejahatan Pencurian Di Wilayah Jawa Barat Menggunakan Algoritma K-Means. *Jurnal Teknologi Dan Komunikasi STMIK Subang*, 12(1), 15–27. <https://doi.org/10.47561/a.v12i1.147>
- Vita Ferezagia, D. (2018). Analisis Tingkat Kemiskinan di Indonesia.
- Wahyudi, D., & Rejekingsih, T. (2013). Analisis Kemiskinan Di Jawa Tengah. <https://www.kompas.com> (2021, 18 Februari). Angka Kemiskinan Indonesia Naik, Ini Data per Provinsi. Diakses pada 07 Mei 2022, dari <https://www.kompas.com/tren/read/2021/02/18/110300865/angka-kemiskinan-indonesia-naik-ini-data-per-provinsi>
- <https://katadata.co.id> . Digitalisasi UMKM di Tengah Pandemi Covid-19. Diakses pada 08 Mei 2022, dari <https://katadata.co.id/umkm>
- <https://radarsukabumi.com> . (2021, 30 September). 5 Kabupaten di Jawa Barat Masuk Miskin Ekstrem. Diakses pada 08 Mei 2022, dari <https://radarsukabumi.com/jawa-barat/bandung/5-kabupaten-di-jawa-barat-masuk-miskin-ekstrem/>