

Pengelompokan Kabupaten/Kota di Jawa Barat Tahun 2018 Berdasarkan Indikator Kemiskinan dengan *Polythetic Divisive Method*

Clarita Simar^{1, a)}, Nurul Gusriani^{1, b)} dan lin Irianingsih^{1, c)}

¹Departemen Matematika Fakultas MIPA – Universitas Padjadjaran

^{a)}claritasimar57@gmail.com

^{b)}nurul.gusriani@unpad.ac.id

^{c)}iin.irianingsih@gmail.com

Abstrak

Masalah kemiskinan merupakan salah satu masalah yang bersifat multidimensi. Faktor kemiskinan pada setiap wilayah berbeda, dan hal tersebut dipengaruhi oleh banyak indikator kemiskinan. Pengukuran dan penentuan indikator kemiskinan akan memudahkan pemerintah membedakan tingkat kemiskinan pada suatu wilayah, sehingga pemerintah dapat membuat kebijakan yang lebih tepat untuk menanggulangi kemiskinan di wilayah tersebut. Penelitian ini bertujuan untuk mengelompokkan karakteristik suatu wilayah ke dalam beberapa kelompok berdasarkan variabel penjelas kemiskinan. Penelitian ini menggunakan data kemiskinan kabupaten/kota se-Jawa Barat pada tahun 2018 yang diperoleh dari BPS. Terkait dengan pengelompokan wilayah kemiskinan ke beberapa kategori, metode yang digunakan adalah metode hierarki pada analisis *cluster* yaitu *polythetic divisive method*. Banyaknya *cluster* optimal dipilih dengan menggunakan Dunn Index. Hasil yang diperoleh adalah *polythetic divisive method* menghasilkan tiga kelompok wilayah kemiskinan dengan Dunn Index sebesar 0,4490613. Karakteristik wilayah masing-masing *cluster* diharapkan dapat membantu pemerintah menentukan kebijakan yang sesuai untuk menanggulangi tingkat kemiskinan di Jawa Barat.

Kata kunci: Kemiskinan, Analisis Cluster, Polythetic Divisive Method, Dunn Index

Abstract

Poverty is one of the multidimensional problems. Poverty's factor in each region is different and is influenced by many poverty indicators. Measurement and determination of poverty indicators will make it easier for the government to distinguish poverty levels in an area, so that the government could make appropriate policies to tackle poverty in the region. This research is used to group the characteristics of an area into several groups of poverty based on poverty explanatory variables. This study used district/city poverty data throughout West Java in 2018 obtained from BPS. Associated with the classification of poverty areas into several categories, the method used was polythetic divisive method. Dunn Index was used to choose the optimal number of clusters. The result is polythetic divisive method produced the three best poverty area cluster with a Dunn Index of 0.4490613. The regional characteristics of each cluster are expected to help the government determine the appropriate policies to tackle poverty levels in West Java.

Keywords: Poverty, Cluster Analysis, Polythetic Divisive Method, Dunn Index

Pendahuluan

Kemiskinan merupakan masalah multidimensi yang tidak hanya tergantung pada variabel keuangan, tetapi juga non keuangan [1]. Faktor kemiskinan setiap wilayah berbeda, dan hal tersebut dipengaruhi oleh banyak indikator kemiskinan. Pengukuran dan penentuan indikator kemiskinan akan memudahkan pemerintah membedakan tingkat kemiskinan dan karakteristik daerah tersebut, sehingga pemerintah dapat membuat kebijakan yang lebih tepat untuk menanggulangi kemiskinan di wilayah tersebut. Oleh karena itu, pemerintah perlu mengelompokkan wilayah-wilayah ke beberapa kelompok kemiskinan berdasarkan variabel penjabar kemiskinan untuk mengetahui karakteristik wilayah. Metode hierarki pada analisis *cluster* merupakan salah satu metode yang dapat digunakan untuk mengelompokkan data yang belum ditentukan banyaknya kelompok yang ingin dibentuk.

Analisis *cluster* merupakan salah satu metode statistika yang dapat digunakan untuk melakukan proses pengelompokan dengan menggunakan ukuran jarak. Analisis *cluster* dibagi menjadi dua metode, yaitu metode hierarki dan metode non hierarki. Metode hierarki dibagi menjadi *agglomerative method* (metode penggabungan) dan *divisive method* (metode pemisahan). Penelitian mengenai analisis *cluster* pernah dilakukan oleh Roux (2018), yaitu mengenai perbandingan algoritma *agglomerative* dan *divisive method* pada data leukemia, tembakar, dan iris [2]. Hasil penelitian Roux menyimpulkan bahwa algoritma *divisive method* lebih baik dibandingkan *agglomerative method*. Meskipun *divisive method* lebih jarang diaplikasikan daripada *agglomerative method*, tetapi *divisive method* memiliki kelebihan di struktur utama data. Berbeda dengan *agglomerative method*, pada *divisive method* semua informasi pada tahap awal digunakan sehingga dapat menghindari keputusan yang tidak tepat pada tahap sebelumnya [3].

Pada penelitian ini dilakukan *clustering* dengan menggunakan *divisive method* yaitu *polythetic divisive method* pada data kemiskinan seluruh kabupaten/kota di Jawa Barat tahun 2018 berdasarkan sepuluh indikator kemiskinan. Penelitian ini harus memenuhi asumsi bahwa tidak terdapat multikolinearitas antar variabel penjabar kemiskinan agar tidak merepresentasikan konsep yang sama dua kali atau lebih pada solusi akhir [4]. Banyaknya *cluster* wilayah kemiskinan pada penelitian ini akan lebih mudah ditangani pemerintah jika kurang dari lima karena kebijakan pengentasan kemiskinan yang dikeluarkan berwujud ke dalam empat kategori kelompok masyarakat [5]. Pemilihan *cluster* optimal dan metode analisis *cluster* terbaik menggunakan Dunn Index.

Metode

1. Analisis Cluster

Analisis *cluster* merupakan teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan kemiripan karakteristik yang dimilikinya. Secara garis besar, analisis *cluster* memiliki dua macam metode yaitu metode hierarki dan metode non hierarki. Metode hierarki terbagi menjadi dua, yaitu *agglomerative method* dan *divisive method*. Pada tahap awal *agglomerative method*, setiap objek berada pada satu *cluster* sehingga jumlah *cluster* sama dengan jumlah objek. Objek yang memiliki kemiripan paling tinggi digabung menjadi satu *cluster*. Proses yang sama akan berlanjut hingga akhirnya terbentuk satu *cluster* yang terdiri dari keseluruhan objek. Berbeda dengan *agglomerative method*, *divisive method* dimulai dengan satu *cluster* dan kemudian objek-objek pada *cluster* tersebut berpisah dan membentuk *cluster* baru hingga pada setiap *cluster* hanya terdapat satu objek.

Analisis *cluster* menggunakan ukuran jarak dalam pengelompokannya. Jarak Euclidean adalah ukuran jarak yang paling umum dikenal. Jarak Euclidean adalah jarak geometris antar dua objek data. Semakin

dekat jarak maka semakin mirip suatu objek data tersebut [6]. Rumus yang digunakan untuk menghitung jarak Euclidean antara objek i dengan j dengan variabel k terdapat pada persamaan berikut.

$$d_{i,j} = \left[\sum_{k=1}^p (x_{i,k} - x_{j,k})^2 \right]^{1/2} \quad (1)$$

untuk $i = j = 1, 2, \dots, n$; $k = 1, 2, \dots, p$.

$d_{i,j}$: jarak antara objek i dengan j

$x_{i,k}$: data ke i variabel ke k

$x_{j,k}$: data ke j variabel ke k

2. Polythetic Divisive Method

Polythetic divisive method merupakan salah satu *divisive method* yang menggunakan algoritma pemisahan untuk mengelompokkan objek-objek yang memiliki kemiripan karakteristik yang tinggi ke dalam satu kelompok. Pendekatan *polythetic divisive method* menggunakan semua variabel p untuk melakukan pemisahan.

Metode ini bekerja dengan cara menemukan jarak rata-rata objek yang paling jauh dari objek lainnya dalam satu kelompok, kemudian menggunakan objek itu sebagai benih untuk *splinter group* yaitu kelompok baru hasil pemisahan tersebut. Setelah itu, akan dihitung *difference* setiap objek di luar *splinter group* dengan rumus sebagai berikut:

$$diff(i) = [\bar{d}_{i,j}, j \notin sg_l] - [\bar{d}_{i,j}, j \in sg_l] \quad (2)$$

dengan:

$diff(i)$: *difference* untuk objek ke i di luar *splinter group*

$\bar{d}_{i,j}$: jarak rata-rata antara objek i dengan j

sg_l : *splinter group* ke l

Jika nilai *difference* terbesar positif, objek ke i akan bergabung dengan *splinter group*. Jika *difference* terbesar negative, prosedur akan dihentikan dan anggota *splinter group* ke l telah lengkap.

3. Dunn Index (DI)

Masalah utama dalam analisis *cluster* adalah menentukan berapa banyaknya *cluster*. Salah satu metode yang dapat digunakan untuk menentukan banyaknya *cluster* optimal adalah dengan menggunakan Dunn Index. Metode ini dapat menentukan banyaknya *cluster* optimal dengan prinsip meminimumkan jumlah kuadrat *error*.

Cluster yang datanya terpisah dengan baik akan memiliki nilai *Dunn Index* yang besar. Semakin besar nilai *Dunn Index* suatu *cluster*, semakin bagus *cluster* tersebut. Indeks ini bekerja dengan cara memaksimumkan jarak *intercluster* dan meminimumkan jarak *intracluster*. *Dunn Index* dapat dihitung dengan menggunakan rumus pada persamaan berikut.

$$DI_c = \min_{1 \leq r \leq c} \left\{ \min_{r+1 \leq s \leq c} \left(\frac{d(c_r, c_s)}{\max_{1 \leq q \leq c} diam(c_q)} \right) \right\} \quad (3)$$

dengan:

DI_c merupakan *Dunn Index* pada jumlah *cluster* $c = 1, 2, \dots, n$.

$d(c_r, c_s)$ merupakan jarak antara *cluster* c_r dan c_s yang dicari dengan menentukan jarak minimum antara objek i di *cluster* c_r dan objek j di *cluster* c_s .

$$d(c_r, c_s) = \min_{i \in c_r, j \in c_s} d_{i,j} \quad (4)$$

$diam(c_q)$ merupakan diameter dari *cluster* c_q yang dapat dihitung dengan rumus berikut

$$diam(c_q) = \max_{u,v \in c_q} d_{u,v} \quad (5)$$

Objek pada penelitian ini adalah data kemiskinan kabupaten/kota di Jawa Barat pada tahun 2018. Data tersebut berjumlah 27 kabupaten/kota di Jawa Barat yang diperoleh dari hasil Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2018 melalui website www.jabar.bps.go.id [8].

Sepuluh variabel yang digunakan dalam penelitian ini adalah persentase kepala rumah tangga laki-laki (X_1), persentase penduduk miskin usia 15 tahun ke atas yang tamat SLTA+ (X_2), persentase penduduk miskin usia 15 tahun ke atas yang tidak bekerja (X_3), persentase pengeluaran non makanan rumah tangga miskin (X_4), persentase bangunan tempat tinggal milik sendiri (X_5), persentase rumah tangga dengan jenis lantai tanah (X_6), persentase rumah tangga dengan jenis dinding bambu (X_7), persentase rumah tangga dengan jenis atap beton (X_8), persentase rumah tangga yang menggunakan jamban sendiri (X_9), dan persentase rumah tangga yang menggunakan air layak (X_{10}).

Penelitian ini menggunakan analisis *cluster* dengan *ward's method*, *average linkage method*, dan *polythetic divisive method*. Data penelitian diolah menggunakan Microsoft Excel dan RStudio 1.2.1335.

Tahapan analisis data yang digunakan adalah sebagai berikut:

1. Mengumpulkan data penelitian.
2. Melakukan pengujian tidak terdapat multikolinieritas antar semua variabel dengan *Variance Inflation Factor* (VIF). Multikolinieritas terindikasi apabila nilai $VIF_h > 10$ [7]. Rumus yang digunakan untuk menghitung nilai VIF terdapat pada persamaan berikut.

$$VIF_h = \frac{1}{1-R_h^2} \quad (6)$$

3. Melakukan *clustering* dengan *polythetic divisive method*.
4. Menghitung Dunn Index dari masing-masing jumlah *cluster*.
5. Menginterpretasikan hasil analisis *cluster*.

Langkah-langkah *clustering* dengan *polythetic divisive method* adalah sebagai berikut:

1. Menyatakan semua data dalam satu *cluster* yaitu $c = 1$ dan banyak *splinter group* $l = 0$.
2. Menghitung jarak Euclidean antar objek.
3. Memilih objek yang memiliki *highest average dissimilarity* atau jarak rata-rata terjauh ke objek lainnya.
4. Banyak *splinter group* bertambah satu yaitu $l = l + 1$.
5. Objek yang terpilih pada langkah 3 memisahkan diri dari *cluster* dan membentuk *cluster* baru atau *splinter group* ke l yaitu sg_l .
6. Banyak *cluster* bertambah satu menjadi $c = c + 1$.
7. Menghitung *difference* untuk setiap objek i di luar *splinter group*.
8. Jika $diff(i)$ terbesar positif, maka objek i tersebut masuk ke sg_l dan ulangi langkah 7. Jika negatif, lanjutkan ke langkah 10.
9. Melakukan pengujian jumlah *cluster* yaitu jika $c = n$, maka iterasi untuk *polythetic divisive method* selesai. Jika $c \neq n$, maka dilanjutkan ke langkah 11.

10. Memilih *cluster* dengan diameter terbesar atau *highest dissimilarity* antar objeknya. Kemudian *cluster* tersebut akan dipisah dengan mengulangi langkah 3-10.

Hasil dan Diskusi

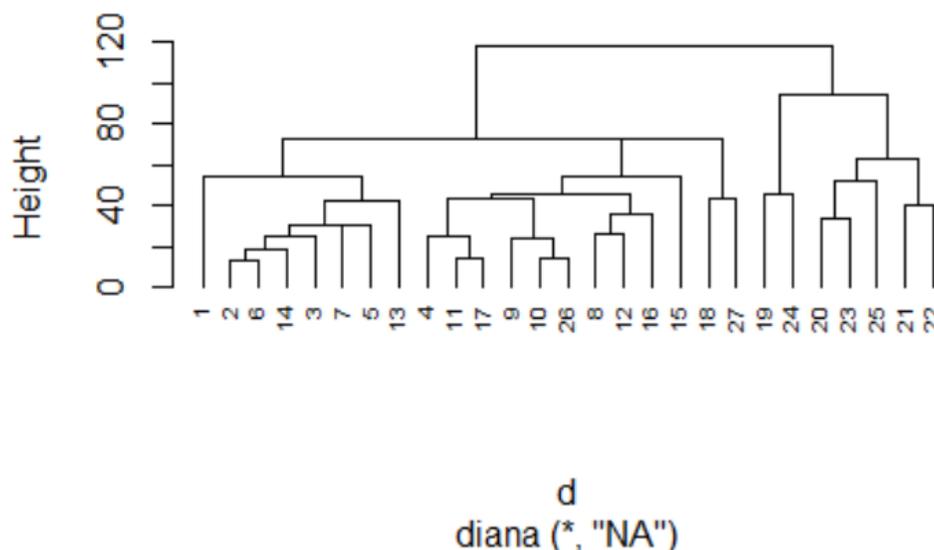
Sebelum dilakukan proses *clustering*, dilakukan pengujian terhadap asumsi multikolinearitas yaitu ada atau tidaknya multikolinearitas antar setiap variabel. Multikolinearitas terindikasi apabila $VIF_h > 10, \forall h$. Hasil dari pengolahan uji asumsi tidak ada multikolinearitas adalah sebagai berikut:

Tabel 1. Nilai VIF masing-masing variabel

h	R_h^2	VIF_h
1	0,6036	2,522704
2	0,7985	4,962779
3	0,2679	1,365934
4	0,5388	2,168257
5	0,4071	1,686625
6	0,7134	3,489184
7	0,6273	2,683123
8	0,6546	2,895194
9	0,6975	3,305785
10	0,577	2,364066

Berdasarkan Tabel 1, diketahui bahwa semua variabel penjelas kemiskinan tersebut tidak terindikasi multikolinearitas atau memenuhi asumsi analisis *cluster*.

Tahapan *clustering* dilanjutkan menggunakan *polythetic divisive method* dengan bantuan RStudio yang digambarkan dengan menggunakan dendrogram. Dendrogram memiliki sumbu x yang menyatakan semua objek *clustering*, dan sumbu y menyatakan jarak dimana *cluster* dibentuk. Dendrogram *polythetic divisive method* dapat dilihat pada Gambar 1.



Gambar 1. Dendrogram *Polythetic Divisive Method*

Berdasarkan hasil *clustering*, dilakukan pemilihan jumlah cluster optimum menggunakan *Dunn Index* dengan membatasi banyaknya *cluster* kurang dari lima. Hasil perbandingan nilai *Dunn Index* pada ketiga metode *clustering* yang digunakan terdapat pada Tabel 2.

Tabel 2. Nilai Dunn Index

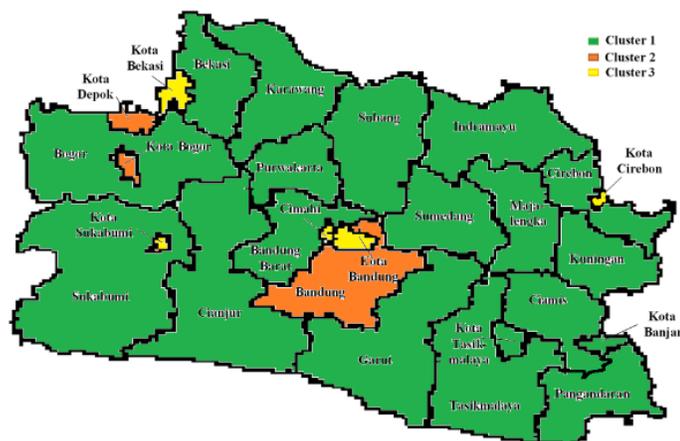
Banyak Cluster	<i>Dunn Index</i>
2	0,3458241
3	0,4490613
4	0,3723184

Berdasarkan Tabel 2, diketahui bahwa nilai *Dunn Index* tertinggi untuk *polythetic divisive method* sebesar 0,4490613 dengan banyak *cluster* tiga. Oleh karena itu, *cluster* optimum untuk *polythetic divisive method* adalah tiga *cluster*.

Hasil *clustering* dengan *polythetic divisive method* untuk tiga *cluster* adalah sebagai berikut:

1. *Cluster 1* terdiri dari 20 kabupaten/kota, yaitu Kabupaten Bogor, Kabupaten Sukabumi, Kabupaten Cianjur, Kabupaten Bandung, Kabupaten Garut, Kabupaten Tasikmalaya, Kabupaten Ciamis, Kabupaten Kuningan, Kabupaten Cirebon, Kabupaten Majalengka, Kabupaten Sumedang, Kabupaten Indramayu, Kabupaten Subang, Kabupaten Purwakarta, Kabupaten Karawang, Kabupaten Bekasi, Kabupaten Bandung Barat, Kabupaten Pangandaran, Kota Tasikmalaya, dan Kota Banjar.
2. *Cluster 2* terdiri dari 2 kota, yaitu Kota Bogor, dan Kota Depok.
3. *Cluster 3* terdiri dari 5 kota, yaitu Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, dan Kota Cimahi.

Daerah *clustering* dengan *polythetic divisive method* dapat digambarkan dengan ilustrasi peta di Gambar 2.



Gambar 2. Peta *Clustering* dengan *Polythetic Divisive Method*

Berdasarkan setiap *cluster* yang diperoleh, dilakukan perhitungan rata-rata dari setiap indikator kemiskinan atau variabel penjas kemiskinan. Nilai rata-rata untuk setiap variabel pada setiap *cluster* terdapat pada Tabel 3 berikut:

Tabel 3. Variabel Tertinggi dan Terendah dengan *Polythetic Divisive Method*

Variabel	Cluster 1	Cluster 2	Cluster 3
X_1	84.0305	67.665	90.94
X_2	10.807	39.49	24.724
X_3	47.8815	52.715	51.222
X_4	28.035	38.08	31.768
X_5	7.397	0.675	2.2
X_6	82.0385	75.07	42.47
X_7	27.6245	0	3.72
X_8	91.613	33.61	66.698
X_9	71.1775	91.86	92.344
X_{10}	63.0385	72.975	69.88

Berdasarkan perhitungan rata-rata tersebut, variabel yang memiliki nilai tertinggi dan terendah dari rata-rata setiap variabel pada masing-masing *cluster* dengan Tabel 4 berikut.

Tabel 4. Variabel Tertinggi dan Terendah dengan *Polythetic Divisive Method*

Cluster	Indikator Tertinggi	Indikator Terendah
1	X_5, X_6, X_7, X_8	$X_2, X_3, X_4, X_9, X_{10}$
2	X_2, X_3, X_4, X_{10}	X_1, X_5, X_7, X_8
3	X_1, X_9	X_6

Kesimpulan

Kesimpulan berdasarkan analisis dan uraian pada bab sebelumnya adalah *polythetic divisive method* menghasilkan tiga kelompok kemiskinan dengan karakteristik masing-masing kelompok sebagai berikut:

1. *Cluster 1* terdiri dari 20 kabupaten/kota di Jawa Barat yaitu Kabupaten Bogor, Kabupaten Sukabumi, Kabupaten Cianjur, Kabupaten Bandung, Kabupaten Garut, Kabupaten Tasikmalaya, Kabupaten Ciamis, Kabupaten Kuningan, Kabupaten Cirebon, Kabupaten Majalengka, Kabupaten Sumedang, Kabupaten Indramayu, Kabupaten Subang, Kabupaten Purwakarta, Kabupaten Karawang, Kabupaten Bekasi, Kabupaten Bandung Barat, Kabupaten Pangandaran, Kota Tasikmalaya, dan Kota Banjar. Kelompok wilayah ini memiliki nilai tertinggi di karakteristik bangunan tempat tinggal milik sendiri, rumah tangga dengan jenis lantai tanah, rumah tangga dengan jenis dinding bambu, dan rumah tangga dengan jenis atap beton. Nilai terendah terdapat di karakteristik penduduk miskin usia 15 tahun ke atas yang tamat SLTA+, penduduk miskin usia 15 tahun ke atas yang tidak bekerja, pengeluaran non makanan rumah tangga miskin, rumah tangga yang menggunakan jamban sendiri, dan rumah tangga yang menggunakan air layak.
2. *Cluster 2* terdiri dari dua kota di Jawa Barat yaitu Kota Bogor, dan Kota Depok dengan nilai tertinggi di karakteristik penduduk miskin usia 15 tahun ke atas yang tamat SLTA+, penduduk miskin usia 15 tahun ke atas yang tidak bekerja, pengeluaran non makanan rumah tangga miskin, dan rumah tangga yang menggunakan air layak. Nilai terendah di karakteristik kepala rumah tangga laki-laki, bangunan tempat tinggal milik sendiri, rumah tangga dengan jenis dinding bambu, dan rumah tangga dengan jenis atap beton.

3. *Cluster* 3 terdiri dari lima kota di Jawa Barat yaitu Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, dan Kota Cimahi dengan nilai tertinggi di karakteristik kepala rumah tangga laki-laki, dan rumah tangga yang menggunakan jamban sendiri, sedangkan nilai terendah di karakteristik rumah tangga dengan jenis lantai tanah.

Ucapan Terima Kasih

Penulis dapat mengucapkan terima kasih kepada pihak yang ikut berperan serta dalam penelitian ini, yang telah memberikan bantuan finansial dan atau fasilitas kepada penulis.

Referensi

- [1] Badan Pusat Statistik Provinsi DIY. 2017. *Laporan Akhir Analisis Kriteria dan Indikator Kemiskinan Multidimensi untuk Diagnostik Kemajuan Daerah di Daerah Istimewa Yogyakarta*. Diakses di: www.bappeda.jogjapro.go.id pada 7 Januari 2020.
- [2] Roux, M. 2019. *A Comparative Study of Divisive and Agglomerative Hierarchical Clustering Algorithms*. *Journal of Classification*, 35(2), 345-366.
- [3] Kaufman, L. dan Rousseeuw, P.J. 2005. *Finding Groups in Data: An Introduction to Cluster Analysis*. New Jersey: John Wiley & Sons, Inc.
- [4] Valovcin, S. 2013. *Assessing Residential Building Energy Simulation Accuracy Through the Use of Clustering*. Colorado: Colorado School of Mines. 8
- [5] Santoso, D. 2018. *Penduduk Miskin Transient*. Jakarta: Pustaka Obor.
- [6] Johnson, R.A. dan Wichern, D.W. 2007. *Applied Multivariate Statistical Analysis Sixth Edition*. New Jersey: Pearson Prentice Hall.
- [7] Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. 2014. *Multivariate Data Analysis Seventh Edition*. Harlow: Pearson Education Limited.
- [8] Badan Pusat Statistik Provinsi Jawa Barat. 2019. *Kemiskinan Kabupaten/Kota di Jawa Barat 2013-2018*. Diakses di: www.jabar.bps.go.id pada 28 Desember 2019