

& Community Services & Social Work Bulletin

Community Services and
Social Work Bulletin

Volume 3

Nomor 1

Page 1-74

Tahun 2023

e-ISSN 2828-027X



Community Services & Social Work Bulletin

IMPLEMENTASI METODE K-MEANS DAN K-MEDOIDS PADA PENGELOMPOKAN PROVINSI INDONESIA BERDASARKAN ASPEK PENDIDIKAN PEMUDA

Halima Tussyakdiah¹, Insani Hasanah¹, Sri Arista Panggola¹, Tiara Ramdhanti¹, Retno Permatasari¹, Cusanti¹, Edy Widodo¹

¹ Universitas Islam Indonesia, Indonesia

1-10

HUBUNGAN POLA HIDUP WANITA LANSIA DIMASA MUDA DENGAN DIAGNOSA PENYAKIT DEGENERATIF WANITA LANSIA

Irwanti Gustina¹, Putri Sarah Dita¹

¹ Universitas Binawan, Indonesia

11-19

PARTISIPASI MASYARAKAT DALAM PENGEMBANGAN DESA WISATA KEMIRI KECAMATAN PANTI KABUPATEN JEMBER

Firdasari Khoirun Nisa¹, Hety Mustika Ani¹, Wiwin Hartanto¹

¹ Universitas Jember, Indonesia

20-38

PEMBERDAYAAN WARGA DALAM PENANGGULANGAN KEBAKARAN DI KECAMATAN SERPONG UTARA KOTA TANGERANG SELATAN

Jukri¹, Tito Inneka¹

¹ Universitas Yuppentek, Indonesia

39-41

WEBSITE HIV/AIDS BAGI KADER DI PIMPINAN RANTING 'AISYIYAH BALECATUR YOGYAKARTA

Dhesi Ari Astuti^{1*}, Nurul Kurniati¹, Mega Ardina¹, Erwin Rasyid¹

¹ Universitas 'Aisyiyah Yogyakarta, Indonesia

42-68

THE EXISTENCE OF SURABI BANDUNG WITH A NEW TASTE

Rizky Asyahri Setiabudiarto^{1*}, Almira Amalia¹, Nurmala¹

¹ Universitas Al-Azhar, Indonesia

59-65

SOSIALISASI PENTINGNYA DIGITAL MARKETING UNTUK PEMASARAN PRODUK UMKM MASYARAKAT DESA KARANG SARI KOTA TANGERANG

Syepry Maulana Husain^{1*}, Muhammad Lutfhi Aksani¹, Firdiansyah Firdaus Basri¹, Muhammad Rizka Saputra¹

¹ Universitas Muhammadiyah Tangerang, Indonesia

66-74

IMPLEMENTASI METODE K-MEANS DAN K-MEDOIDS PADA PENGELOMPOKAN PROVINSI INDONESIA BERDASARKAN ASPEK PENDIDIKAN PEMUDA

Halima Tussyakdiah^{1*}, Insani Hasanah¹, Sri Arista Panggola¹, Tiara Ramdhanti¹, Retno Permatasari¹, Cusanti¹, Edy Widodo¹

¹ Universitas Islam Indonesia, Kabupaten Sleman, Indonesia

*Correspondence email: halima23121998@gmail.com

Received: 23 March 2023; Accepted: 10 May 2023; Published: 14 June 2023

doi: 10.31000/cswb.v3i1.10153

Abstract: The quality of education in Indonesia is still a concern, seen from a number of problems that become obstacles to improving the quality of education as well as affecting the quality of Indonesian youth. This study aims to group provinces in Indonesia based on the aspect of youth education using the K-Means and K-Medoids methods. To determine the optimum k, the average silhouette method is used and the SW and SB ratio is used to evaluate the cluster results. The results obtained are 2 clusters optimum. For the K-Means method, cluster 1 consists of 19 provinces and cluster 2 consists of 14 provinces. Whereas in the K-Medoids method, cluster 1 consists of 22 provinces and cluster 2 consists of 11 provinces. The K-Means method is better than the K-Medoids method because it has a ratio value of 0.527941 which is smaller than the K-Medoid ratio value of 0.5612719.

Keyword: K-Means; K-Medoid; Education; Average Silhouette; Standard Deviation.

PENDAHULUAN

Pemuda menurut Undang-Undang No.40 tahun 2009 adalah warga negara Indonesia yang berusia 16 sampai 30 tahun yang merupakan periode penting usia pertumbuhan dan perkembangan. Menurut hasil Susenas Tahun 2019, perkiraan jumlah pemuda sebesar 64,19 juta jiwa atau seperempat dari total penduduk Indonesia (Statistik, 2019). Untuk melihat kualitas pemuda, biasanya dilihat berdasarkan capaian pendidikan dan kesehatannya. Berdasarkan Undang-undang Dasar 1945 dengan misi “mencerdaskan kehidupan bangsa” memberikan pendidikan dalam arti yang seluas-luasnya. Namun demikian, pendidikan di Indonesia masih menghadapi sejumlah permasalahan yang hingga saat ini belum terpecahkan.

Kualitas pendidikan di Indonesia masih menjadi perhatian karena banyaknya kendala yang memengaruhi peningkatan kualitas pendidikan Indonesia seperti menurut ahli pendidikan Soedijarto mengatakan bahwa



“rendahnya kualitas pendidikan di samping disebabkan oleh pemberian peranan kurang proporsional terhadap sekolah, kurang memadainya perencanaan, pelaksanaan, pengelolaan sistem kurikulum, dan penggunaan prestasi hasil belajar secara kognitif sebagai satu-satunya indikator keberhasilan pendidikan, juga disebabkan oleh sistem evaluasi secara tidak berencana didudukkan sebagai alat pendidikan dan bagian terpadu dari sistem kurikulum.

Analisis cluster merupakan suatu teknik analisis statistik yang ditujukan untuk menempatkan sekumpulan obyek ke dalam dua atau lebih grup berdasarkan kesamaan-kesamaan obyek atas dasar berbagai karakteristik (D. A. S. Simamora, 2017). Salah satu pengelompokan non-hierarki adalah metode *K-Means* dan *K-Medoids* clustering. Perbedaan dari kedua algoritma ini yaitu algoritma *K-Medoids* menggunakan objek sebagai perwakilan (medoids) pusat cluster untuk setiap cluster, sedangkan *K-Means* menggunakan nilai rata-rata (mean) sebagai pusat cluster (N. K. Kaur, 2018). Pada metode *K means* memiliki ketelitian yang cukup tinggi terhadap ukuran objek, sehingga algoritma ini relatif lebih terukur dan efisien untuk pengolahan objek dalam jumlah yang besar. Sedangkan metode *K-Medoids* cocok digunakan untuk mengatasi kelemahan pada algoritma *K-Means* yang sensitif terhadap noise dan outlier, dimana objek dengan nilai yang besar yang memungkinkan menyimpang dari distribusi data (D. F. Pramesti, 2017).

Penelitian analisis cluster dengan membandingkan metode *K-Means* dan *K-Medoids* pernah dilakukan oleh (D. Marlina, 2018; Statistik, 2019) dengan judul “Implementasi Algoritma *K-Medoids* dan *K-Means* untuk Pengelompokan Wilayah Sebaran Cacat pada Anak”. Pada penelitiannya menunjukkan metode *K-Medoids* lebih baik dalam melakukan pengelompokan. Penelitian lainnya juga pernah dilakukan oleh (Fuady & Nugraha, 2017), hasil dari penelitiannya diperoleh bahwa metode *K-Means* lebih baik daripada metode *K-Medoids* dalam mengelompokan 82 Kota di Indonesia Berdasarkan Indeks Harga Konsumen dengan hasil cluster yang terbentuk sebanyak 3 cluster.

Berdasarkan permasalahan diatas, peneliti akan membandingkan metode *K-Means* dan metode *k-medoid clustering* untuk mengelompokkan provinsi-provinsi yang ada di Indonesia berdasarkan aspek pendidikan pemuda di Indonesia. Pengelompokan akan membentuk beberapa kelompok dari daerah yang mirip atau memiliki kesamaan satu sama lain, sehingga hasil analisa perhitungan diharapkan dapat dijadikan bahan untuk pengambilan kebijakan dalam meningkatkan kualitas dan kredabilitas pemuda di Indonesia.

Jarak Euclidean

Menurut (Wichern, 2002), jarak Euclid merupakan tipe pengukuran jarak dalam analisis cluster yang paling umum digunakan untuk mengukur jarak dari

obyek data ke pusat cluster. Semakin dekat jarak maka semakin mirip suatu obyek data tersebut. Bentuk umum *Euclidean Distance* (d) pada dua dimensi, misalkan ada dua titik koordinat (x_{11}, x_{21}) dan (x_{12}, x_{22}) dapat diperoleh seperti berikut (Jannah, 2009):

$$d(P_1, P_2) = \sqrt{(x_{21} - x_{11})^2 + (x_{22} - x_{12})^2} \quad (1)$$

Keterangan: d_{ij} = jarak *Euclidean* objek data ke- i dan objek data ke- j , m = banyaknya peubah/parameter yang digunakan, x_{ik} = objek data ke- i pada peubah ke- k , x_{jk} = objek data ke- j pada peubah ke- k .

K-Means Clustering

Metode *K-Means* bertujuan untuk meminimalisasikan *object function* yang diatur dalam proses *clustering* dengan cara meminimalkan variasi antar data yang ada dalam suatu *cluster* lainnya. Adapun tahap-tahap untuk melakukan *K-Means clustering* antara lain sebagai berikut (I. Kamila, 2019):

1. Menentukan nilai k sebagai pusat kluster awal atau jumlah *cluster* yang akan dibentuk secara acak
2. Setiap data dalam *dataset* dibagi kedalam beberapa kelompok k *cluster* antara setiap titik dan pusat *cluster* yang didapatkan berdasarkan jarak *Euclidean Distance* seperti yang terdapat pada persamaan (1)
3. Setiap pusat *cluster* dihitung ulang berdasarkan dari nilai rata-rata dalam *cluster* yang didapatkan.
4. Langkah 2 dan 3 diulangi samapai kelompok tersebut sesuai. Perulangan dapat didefinisikan secara berbeda tergantung pada implementasi, tetapi biasanya pada proses langkah 2 dan 3 dapat diulangi apabila kelompok *cluster* masih mengalami perubahan dan proses akan terhenti apabila tidak terjadi perubahan anatar material dalam *cluster*.

K-medoid Clustering

Adapun tahapan analisis *cluster* dengan menggunakan metode *K-Medoids* adalah sebagai berikut (D. F. Pramesti, 2017):

1. Inisiasi pusat *cluster* sesuai dengan jumlah klaster (k).
2. Alokasikan setiap data atau objek ke medoid terdekat dengan menggunakan ukuran jarak *Euclidean Distance* seperti pada persamaan (1).
3. Pilih objek atau data secara acak dari masing-masing cluster sebagai perwakilan *medoid* baru.
4. Alokasikan setiap data ke medoid baru dengan menggunakan *Euclidean Distance* seperti pada langkah b.
5. Hitung total simpangan (S) dengan cara mengurangi nilai total jarak baru dengan nilai total jarak lama. Jika diperoleh, maka tukar objek (data) lama dengan data baru sebagai *medoid* untuk membentuk *cluster* baru.

6. Ulangi langkah c sampai dengan e hingga tidak ada pergantian *medoid* atau (*medoid* tetap) sehingga diperoleh *cluster* beserta anggotanya masing-masing.

Silhouette Coefficient Index (SI)

Silhouette Coefficient Index adalah salah satu metode analisa untuk mendapatkan nilai validasi pada sebuah metode *clustering*. Jika $si=1$, berarti objek pada 1 sudah berada dalam *cluster* yang tepat (Santosa, 2012). Berikut adalah rumus untuk menghitung a_{ij} .

$$a_{ij} = \frac{1}{m_j - i} - \sum_{r=1, r \neq i}^{m_j} d((x_{ij}, X_{jr} = 1r \neq 1, x_{rj}), i = 1, 2, \dots, m_j) \quad (2)$$

dimana $j = \text{cluster}$; $i = \text{index data}$; a_{ij} = rata-rata jarak data ke- i terhadap semua data lainnya dalam satu *cluster*; m_j = jumlah data dalam *cluster* ke- j . Nilai SI adalah semua rata-rata pada semua kumpulan data didapat cara berikut persamaannya:

$$SI_j = \frac{1}{m_j} \sum_{i=1}^{m_j} SI_{ij} \quad (3)$$

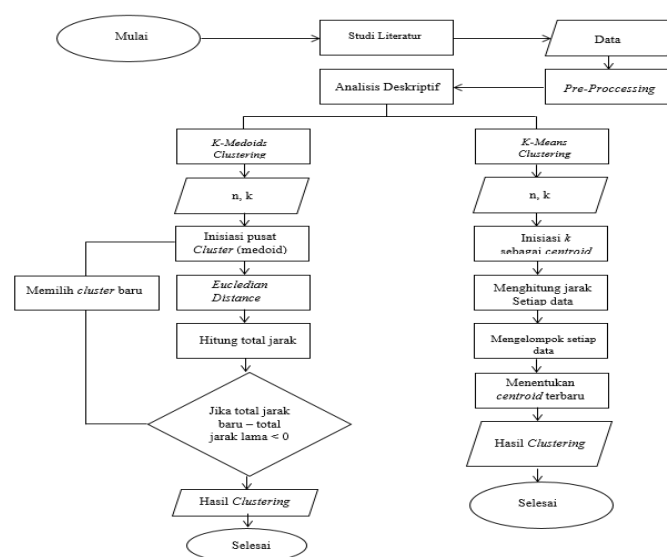
Metode Terbaik dengan Simpangan Baku

Sebuah metode pengelompokan jika mempunyai nilai simpangan baku dalam kelompok (S_w) yang minimum dan nilai simpangan baku antar kelompok (S_b) yang maksimum (Jannah, 2009). Pengelompokan yang baik akan memiliki nilai s_w yang minimum dan s_b maksimum atau metode terbaik dapat menghasilkan nilai rasio simpangan baku minimum s_w terhadap s_b dengan rumus seperti berikut (W. J. Bunkers, 1996):

$$S = \frac{s_w}{s_b} \times 100\% \quad (4)$$

Flow chart

Gambar 1 Flowchart Penelitian



Langkah Kerja

1. Studi Literatur: mengumpulkan berbagai sumber acuan atau referensi yang relevan.
2. *Pre-Processing*: pembersihan data (proses menghilangkan noise dan data yang tidak konsisten), seleksi data (memilih atribut atau variabel yang sesuai untuk dianalisis), dan transformasi data (proses pengubahan atau penggabungan data ke dalam format yang sesuai).
3. Analisis Deskriptif: untuk menggambarkan karakteristik atau deskripsi dari setiap variabel penelitian.
4. *Clustering K-Means dan K-Medoids*: dijelaskan pada subbab Metode *K-Mean dan K-Medoids*.
5. Hasil *clustering*.

Data dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu data aspek pendidikan pemuda menurut Provinsi di Indonesia tahun 2019 diperoleh dari website resmi Badan Pusat Statistik Indonesia.

Tabel 1 Definisi Operasional Variabel Penelitian

Variabel	Satuan	Definisi
X_1 = Tamatan SD	%	Jumlah pemuda dengan jenjang pendidikan terakhir SD
X_2 = Sekolah Menengah Pertama (SMP)	%	Jumlah pemuda dengan jenjang pendidikan terakhir SMP
X_3 = Sekolah Menengah Atas (SMA)	%	Jumlah pemuda dengan jenjang pendidikan terakhir SMA
X_4 = Perguruan Tinggi	%	Jumlah pemuda yang sedang menempuh pendidikan perguruan tinggi/Universitas
X_5 = Menggunakan Komputer	%	Penggunaan komputer pada pemuda berdasarkan jenis kelamin
X_6 =Angka Buta huruf (ABH)	%	Jumlah pemuda yang tidak dapat membaca dan atau menulis huruf latin atau huruf lainnya yang disebut angka buta huruf(ABH)

sumber : <http://bps.go.id/>

HASIL DAN PEMBAHASAN

Analisis deskriptif

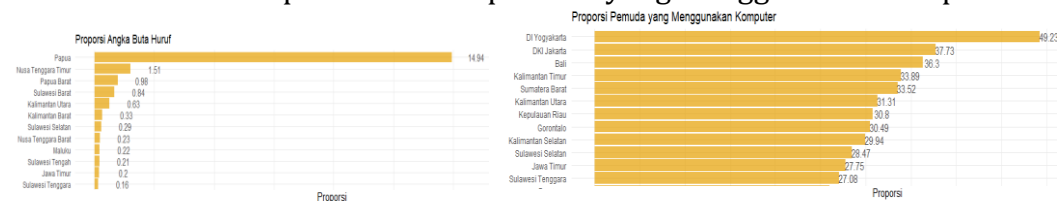
Tabel 2 Hasil Analisis Deskriptif

Variabel	Ukuran Pemusatan		
	Maksimum	Rata-rata	Minimum
Tamatan SD	13,76	0,65	2,99
Tamatan SMP	39,08	30,12	21,89
Tamatan SMA	55,43	42,57	28,52
Tamatan Perguruan Tinggi	19,84	14,52	9,95

Berdasarkan "Tabel 2", diketahui bahwa presentase rata-rata tamatan Sekolah Dasar (SD) di Indonesia sebesar 0,65% dengan presentase tamatan SD

paling tinggi berada pada Provinsi Kep. Bangka Belitung yaitu sebesar 13,76% dan presentase paling rendah berada pada Provinsi DI Yogyakarta yaitu sebesar 2,99%. Presentase rata-rata tamatan Sekolah Menengah Pertama (SMP) di Indonesia adalah sebesar 30,12% dengan presentase tamatan SMP paling tinggi berada pada Provinsi Sulawesi Barat yaitu sebesar 39,08% dan presentase paling rendah yaitu sebesar 21,89% terdapat pada Provinsi DI Yogyakarta. Selanjutnya untuk Sekolah Menengah Atas (SMA), presentase rata-rata tamatannya di Indonesia ada sebanyak 42,57%, dengan presentase tamatan tertinggi yaitu sebesar 55,43% berada pada Provinsi DI Yogyakarta dan presentase terendah yaitu sebesar 28,52% terdapat pada Provinsi Sulawesi Barat. Kemudian untuk tamatan perguruan Tinggi di Indonesia, presentase rata-rata tamatannya berada sebesar 14,52%, dengan presentase tertinggi yaitu sebesar 19,84% berada pada Provinsi Aceh dan presentase terendah yaitu sebesar 9,95% berada pada Provinsi Jawa Barat.

Gambar 2 Proporsi ABH dan pemuda yang Menggunakan Komputer



Berdasarkan “Gambar 2”, presentase rata-rata Angka Buta huruf di Indonesia adalah sebesar 0,65%. Kemudian, presentase Angka Buta Huruf tertinggi berada pada Provinsi Papua yaitu sebesar 14,84% dan presentase terendah berada pada Provinsi DI Yogyakarta yaitu sebesar 0%, berdasarkan data ini, dapat dikatakan bahwa tidak terdapat pemuda yang buta huruf di Provinsi DI Yogyakarta. Namun, dilihat dari beberapa sumber, diketahui bahwa sebenarnya masih terdapat pemuda yang buta huruf, tetapi dengan proporsi yang sangat kecil.

Kemudian dapat dilihat rata-rata presentase pemuda yang menggunakan komputer dalam kehidupan sehari-hari di Indonesia ada sebesar 26,11%, dengan presentase tertinggi berada pada Provinsi DI Yogyakarta yaitu sebesar 49,23% dan presentase terendah berada pada Provinsi Papua yaitu sebesar 11,71%. Terlihat bahwa ada perbedaan yang signifikan dari kedua provinsi tersebut, hal ini bisa saja terjadi karena adanya keterbatasan fasilitas, listrik, dan infrastruktur TIK di Provinsi Papua.

Pengujian Asumsi Multikolinieritas

Pengujian multikolinieritas dapat dilihat dari nilai *tolerance* dan lawannya *variance inflation factor* (VIF). Berikut merupakan pengujian hipotesis *no* multikolinieritas dengan menggunakan nilai VIF sebagai berikut:

i) Hipotesis:

H_0 : Tidak terjadi korelasi atau hubungan antar variabel independen (*No* Multikolinieritas)

H_1 : Terjadi korelasi atau hubungan antar variabel independen (Multikolinieritas)

- ii) Titik kritis: $VIF \geq 10$ maka tolak H_0
 iii) Keputusan:

Tabel 3 Hasil Keputusan Uji Multikolinieritas

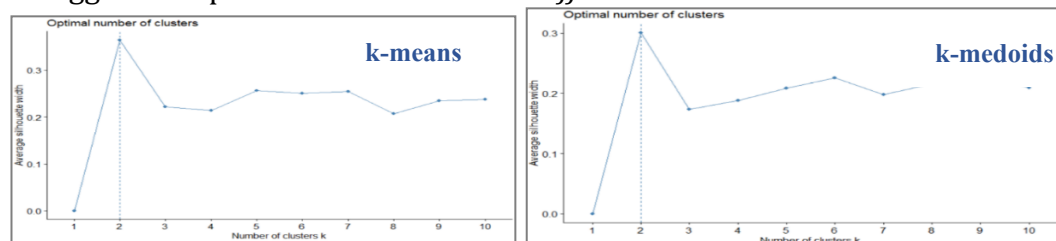
Variabel	Nilai VIF	tanda	Cutoff VIF	Keputusan
ABH	1,37	<	10	Gagal Tolak H_0 atau Tidak terdapat Multikolinearitas
SD	3,15	<		
SMP	3,31	<		
SM	4,50	<		
PT	3,01	<		
Menggunakan Komputer	1,33	<		

- iv) Kesimpulan:

Dengan menggunakan *cutoff* nilai $VIF \geq 10$, data yang ada gagal menolak hipotesis H_0 yang artinya tidak terjadi multikolinieritas.

Penentuan Jumlah Cluster

Peneliti akan menggunakan metode penentuan k-optimum menggunakan pendekatan *silhouette coefficient*.

**Gambar 4.** K-Optimum *Silhouette Coefficient* Metode *K-Means* dan *K-Medoids*

Berdasarkan plot k-optimum “Gambar 4” dapat dilihat bahwa dari nilai *average silhouette* metode *K-Medoid* dan *K-Means* bahwa jumlah *cluster* yang baik dibentuk adalah sebanyak 2. Maka pada penelitian ini menggunakan sebanyak 2 *cluster* yang akan menggambarkan kualitas pendidikan pemuda di Indonesia baik dan kurang baik.

Pembentukan dan Profilisasi Cluster *K-Mean*

Tabel 4 Hasil Keanggotaan Cluster Menggunakan *K-Means*

Cluster	Jumlah Anggota	Provinsi
1	19	Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Kepulauan Riau, DKI Jakarta, DI Yogyakarta, Bali, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Selatan, Sulawesi Tenggara, Maluku, Maluku Utara
2	14	Lampung, Kep. Bangka Belitung, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Sulawesi Tengah, Gorontalo, Sulawesi Barat, Papua Barat

Tabel 5 Profilisasi *cluster* Metode K-Means

Variabel	Cluster 1	Cluster 2
ABH	0,1157	0,3421
SD	6,4010	9,5650
SMP	28,0263	32,6407
SM	45,2436	39,3528
PT	16,2573	12,3571
Menggunakan Komputer	28,6547	23,67

Dari tabel profilisasi *cluster* di “Tabel 5”, terdapat dua warna yaitu warna kuning menunjukkan persentase tinggi dan hijau menunjukkan persentase sedang. Maka diperoleh karakteristik masing-masing *cluster* sebagai berikut.

- 1) *Cluster 1* menggambarkan kualitas pendidikan pemuda sudah baik. Karakteristik pendidikan pemuda pada kluster 2 yaitu merupakan wilayah yang memiliki persentase tinggi pada variabel pendidikan tertinggi yang ditamatkan SMA/SMK, Perguruan tinggi dan kemampuan menggunakan komputer.
- 2) *Cluster 2* menggambarkan kualitas pendidikan pemuda kurang baik. Karakteristik pendidikan pemuda pada kluster 1 yaitu merupakan wilayah yang memiliki persentase tinggi pada variabel Angka Buta Huruf (ABH), pendidikan tertinggi yang ditamatkan SD dan SMP.

Menentukan Keباikan metode *clustering* dengan Simpangan Baku

Hasil *cluster* terbaik adalah yang memiliki nilai S_w lebih kecil daripada nilai S_B . Kemudian, untuk melihat metode terbaik, dapat dilihat dari nilai ratio.

Tabel 6 Nilai Rata-rata Simpangan Baku

Metode	Rata-rata Simpangan Baku	
	<i>Average Within</i> (S_w)	<i>Average Between</i> (S_B)
<i>K-Means</i>	12,16392	23.04042
<i>K-Medoid</i>	21,93839	12,3134

Dapat dilihat hasil pada “Tabel 6”, menunjukkan bahwa kedua metode memiliki nilai rata-rata S_w yang lebih kecil daripada nilai S_B . Sehingga dapat dikatakan bahwa *cluster* yang dihasilkan cukup baik dan homogen terhadap anggota dalam *cluster*-nya.

Tabel 7 Nilai *Ratio*

Metode	Nilai <i>Ratio</i>
<i>K-Means</i>	0,527941
<i>K-Medoid</i>	0,5612719

Hasil pada “Tabel 7” menunjukkan bahwa metode *K-Means* memiliki performa yang lebih baik dibandingkan dengan meto *K-Medoid*.hal ini dapat dilihat dari

nilai *ratio* metode *K-Means* yang lebih kecil yaitu 0,527941, dibandingkan nilai *ratio* metode *K-Medoid* yaitu sebesar 0,5612719

SIMPULAN DAN SARAN

Kesimpulan

1. Rata-rata angka buta huruf pemuda di Indonesia sebesar 0,65%. Dilihat dari persentase pemuda menggunakan komputer, persentase tertinggi ada pada provinsi DI Yogyakarta sebesar 49,23% dan persentase terendah pada Provinsi Papua sebesar 11,71%. Persentase paling banyak dari pendidikan terakhir yang ditamatkan yaitu sebesar 19,84% terdapat pada Provinsi Aceh dan persentase paling sedikit yaitu sebesar 9,95% pada Provinsi Jawa Barat.
2. Hasil pengelompokan menggunakan algoritma *K-Means* dengan jarak *euclidean* membentuk 2 cluster yaitu cluster 1 terdiri dari 19 Provinsi, cluster 2 terdiri dari 14 Provinsi. Sedangkan hasil pengelompokan menggunakan algoritma *K-Medoid* dengan jarak *euclidean* membentuk 2 cluster yaitu cluster 1 terdiri dari 22 Provinsi, cluster 2 terdiri dari 11 Provinsi.
3. Dari perbandingan hasil pengelompokan menggunakan nilai *ratio*, metode *K-Means* memiliki nilai rasio sebesar 0,5279 lebih kecil dibandingkan metode *K-Medoid* yang memiliki nilai rasio sebesar 0,5612719. Sehingga diketahui metode *K-Means* memiliki performa yang lebih baik, khususnya pada pengelompokan provinsi-provinsi di Indonesia berdasarkan variabel penunjang kualitas pemuda di Indonesia tahun 2019.

Saran

1. Untuk cluster 1 metode *K-Means*. Dilihat dari karakteristik pendidikan pemuda pada cluster ini memiliki persentase tinggi pada variabel pendidikan tertinggi yang ditamatkan SMA/SMK, Perguruan tinggi dan kemampuan menggunakan komputer. Pada wilayah ini menggambarkan kualitas pendidikan pemuda sudah baik. Namun pemerintah juga tetap harus memperhatikan pendidikan pemuda di wilayah ini dan terus meningkatkan program-program pendidikan yang lebih baik.
2. Untuk cluster 2 metode *K-Means*. Dilihat dari karakteristik pendidikan pemuda pada cluster ini memiliki persentase tinggi pada variabel Angka Buta Huruf (ABH), pendidikan tertinggi yang ditamatkan SD dan SMP. Menggambarkan pada wilayah ini kualitas pendidikan pemuda masih kurang baik. Sehingga bagi pemerintah selaku pengambil kebijakan agar bisa melakukan tindakan seperti meningkatkan program-program pendidikan atau memberikan sosialisasi agar pemuda lebih memperhatikan pendidikannya.

3. Saran untuk penelitian selanjutnya dapat menggunakan metode cluster yang berbeda dengan studi kasus yang sama.

DAFTAR PUSTAKA

- D. A. S. Simamora, M. T. (2017). Clustering Data Kejadian Tsunami Yang Disebabkan Oleh Gempa Bumi Dengan Menggunakan Algoritma K-Medoids. *International Journal of Computer Applications in Technology*, 1(8), 635-640.
- D. F. Pramesti, L. M. (2017). Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer(J-PTIIK)*, 1(9), 723-732.
- D. Marlina, N. L. (2018). Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak. *Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, 4(2), 64.
- Fuady, M. B., & Nugraha, J. (2017). IMPLEMENTASI METODE K-MEANS DAN K-MEDOIDS UNTUK MENGELOMPOKKAN 82 KOTA DI INDONESIA BERDASARKAN INDEKS HARGA KONSUMEN. (pp. 327-337). dspace uii.
- I. Kamila, U. K. (2019). Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau. *jurnal Ilmu Rekayasa dan Manajemen Sistem informasi*, 5(1), 119.
- Jannah, U. (2009). *Perbandingan Jarak Euclid dengan Jarak Mahalanobis pada Analisis Cluster Hirarki*. Universitas Islam Negeri Maulana Malik Ibrahim.
- N. K. Kaur, U. K. (2018). K-Medoid clustering algorithm-a review. *International Journal of Computer Applications in Technology*, 1(1), 2349-1841.
- Santosa, T. A. (2012). Analisis Perbandingan Metode Hierarchical Clustering, K-Means dan Gabungan Keduanya dalam membentuk Cluster Data (Studi Kasus: Problem Kerja Praktek Jurusan Teknik Industri ITS. *Jurnal Teknik ITS*, 1(1), 1-5.
- Statistik, B. P. (2019). *Statistik Pemuda Indonesia 2019*.
- W. J. Bunkers, J. R. (1996). Definition of Climate Regions in the Northern Plains Using an Objective Cluster Modification Technique. *Journal Climate*, 9, 130-146.
- Wichern, R. A. (2002). *Applied Multivariate Statistical Analysis* (5th ed.). Prentice-Hall.

UCAPAN TERIMAKASIH

Terimakasih pada semua pihak yang ikut serta membantu menyelesaikan penelitian ini, terutama pak Edi Widodo selaku dosen pendamping.