
K-MEANS CLUSTERING ANALYSIS PADA PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI NUSA TENGGARA TIMUR BERDASARKAN INDIKATOR PENDIDIKAN

Oleh

Nony Lenama¹, Maria A. Kleden², Irvandi G. Pasangka³

^{1,2,3}Program Studi Matematika, Fakultas Sains dan Teknik, Universitas Nusa

Cendana

Email: nony.lenama30@gmail.com

Article History:

Received: 20-04-2023

Revised: 06-05-2023

Accepted: 18-05-2023

Keywords:

Indikator Pendidikan, K-Means Clustering Analysis

Abstract: Salah satu parameter kemajuan pendidikan dapat dilihat dari indikator pendidikan di suatu daerah. Di Provinsi Nusa Tenggara Timur memiliki indikator pendidikan yang berada di bawah angka nasional (95%). Tujuan dari penelitian ini adalah untuk mengelompokkan kabupaten/kota di Nusa Tenggara Timur berdasarkan indikator pendidikan dengan menggunakan K-means clustering analysis. Indikator yang digunakan dalam penelitian ini adalah indikator pendidikan yaitu angka partisipasi kasar, angka partisipasi murni dan angka partisipasi sekolah. Hasil penelitian ini menggambarkan 4 klaster yaitu klaster pertama terdiri dari 5 anggota yaitu Kabupaten Sumba Barat, Kupang, Sikka, Ende, dan Manggarai. Klaster kedua terdiri dari 7 anggota yaitu Kabupaten Alor, Rote Ndao, Manggarai Barat, Sumba Tengah, Sumba Barat Daya, Nagekeo, dan Manggarai Timur. Klaster ketiga terdiri dari 8 anggota yaitu Kabupaten Sumba Timur, Timor Tengah Selatan, Timor Tengah Utara, Belu, Lembata, Flores Timur, Ngada, dan Malaka. Sedangkan klaster keempat memiliki 2 anggota yaitu Kabupaten Sabu Raijua dan Kota Kupang.

PENDAHULUAN

Pendidikan merupakan hak asasi manusia setiap warga Negara. Pendidikan menjadi landasan kuat yang diperlukan untuk meraih kemajuan bangsa di masa depan, bahkan lebih penting lagi sebagai bekal dalam menghadapi era global dengan persaingan antara bangsa yang berlangsung sangat ketat. Pendidikan menjadi syarat mutlak yang harus dipenuhi karena merupakan faktor penentu bagi suatu bangsa untuk bisa bersaing dengan ruang lingkup internasional.

Salah satu parameter keberhasilan pendidikan dapat dilihat dari indikator pendidikan di suatu daerah. Indikator pendidikan di Indonesia memiliki komponen yang berbeda-beda. Salah satu indikator yang mendukung keberhasilan pendidikan adalah menuntaskan Angka Partisipasi Kasar (APK) dan Angka Partisipasi Murni (APM) mutu pendidikan hingga minimal mencapai 95%. Berdasarkan pernyataan tersebut pendidikan di Nusa Tenggara Timur masih belum maksimal. Apabila ditinjau dari jenjang pendidikan formal khususnya pada jenjang

pendidikan SMA/SMK/MA. Hal ini dapat dilihat dari data Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Timur 3 tahun terakhir yaitu tahun 2019-2021 untuk usia SMA berturut-turut persentase APK sebesar 84,63%, 84,70% dan 86,36%, dan untuk persentase APM sebesar 53,68%, 54,09% dan 54,29%. Data ini menunjukkan bahwa persentase APK dan APM usia SMA/SMK/MA di Nusa Tenggara Timur masih rendah karena belum mencapai 95% [1].

Berdasarkan uraian di atas, maka perlu dilakukan analisis untuk memperoleh gambaran pendidikan di Indonesia khususnya di Nusa Tenggara Timur (NTT) berdasarkan indikator pendidikan. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Nusa Tenggara Timur berdasarkan indikator pendidikan dengan menggunakan metode pengelompokan.

Pengelompokan adalah sebuah metode untuk mengumpulkan suatu objek kajian yang memiliki kemiripan karakteristik agar dapat menghasilkan sebuah kesimpulan bersifat nyata, karena data yang dikumpulkan harus berasal dari sumber yang bisa dipertanggungjawabkan. Untuk mengelompokkan objek-objek yang memiliki kesamaan karakteristik dapat digunakan analisis kluster (*clustering analysis*). Analisis kluster mirip seperti analisis faktor, analisis kluster juga meneliti seluruh hubungan interdependensi, tidak ada perbedaan variabel bebas dan variabel tak bebas dalam analisis kluster. Analisis kluster termasuk dalam analisis statistik multivariat metode interdependen, dan oleh karena itu tujuan analisis kluster tidak untuk menghubungkan atau membedakan dengan sampel atau variabel lain. Analisis kluster berguna untuk meringkas data dengan jalan mengelompokkan objek-objek berdasarkan kesamaan karakteristik tertentu di antara objek-objek yang akan diteliti.

K-Means clustering adalah metode analisis *cluster* yang paling sering digunakan dalam penelitian. Metode *K-Means clustering* berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, di mana data dalam satu kelompok mempunyai karakteristik yang sama, dan mempunyai karakteristik yang berbeda dengan data yang ada di kelompok lain di mana banyaknya kelompok (*cluster*) ditentukan terlebih dahulu. *K-Means clustering analysis* banyak digunakan dalam penelitian karena metode ini sangat sederhana, memiliki kecepatan yang lebih tinggi dibandingkan metode *hirarki* dan dapat menghasilkan *cluster* yang optimum [2-5].

LANDASAN TEORI

Analisis Multivariat

Analisis multivariat merupakan perluasan dari analisis univariat atau bivariat. Analisis multivariat adalah analisis multi variabel dalam satu atau lebih hubungan. Analisis ini berhubungan dengan semua teknik statistik yang secara simultan menganalisis sejumlah pengukuran pada individu atau objek [6-7]. Analisis multivariat berasal dari kata multi dan *variate*, dua kata ini dapat diartikan analisis lebih dari dua variabel. Analisis multivariat merupakan salah satu analisis statistik yang berkaitan dengan analisis banyak variabel, menurut Hair analisis multivariat mengacu pada semua teknik statistik yang secara bersamaan menganalisis beberapa pengukuran terhadap individu atau objek dalam suatu riset [8]. Jadi bisa dikatakan analisis multivariat adalah perluasan dari analisis univariate (seperti uji t) atau bivariat (seperti korelasi dan regresi linear sederhana).

Analisis Klaster

Analisis klaster adalah salah satu dari metode dalam analisis multivariat yang memiliki tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Analisis klaster mengelompokkan individu atau objek penelitian, sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam klaster yang sama. Klaster-klaster yang terbentuk dalam satu klaster mempunyai ciri yang relatif sama (homogen), sedangkan antar klaster mempunyai ciri yang berbeda (heterogen). Pengelompokan ini dilakukan berdasarkan variabel-variabel yang diamati [9].

Analisis klaster harus memenuhi dua asumsi yaitu sampel yang diambil harus benar-benar mewakili populasi dan multikolinearitas yaitu korelasi antar objek.

Ukuran Jarak

Dalam penelitian ini ukuran jarak digunakan untuk mengetahui seberapa mirip atau berbeda antara objek-objek yang diteliti menggunakan jarak *Euclidean*. Jarak *Euclidean* mengukur jumlah kuadrat perbedaan nilai antar masing-masing variabel. Prasetyo menyatakan bahwa jarak *euclidean* merupakan ukuran jarak yang dapat memberikan jarak terdekat atau terpendek antar dua data [10]. Adapun persamaan untuk menghitung jarak *Euclidean* adalah sebagai berikut

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}$$

K-Means Cluster

K-means merupakan metode *non-hirarki* yang paling banyak digunakan untuk pengelompokan karena kemudahannya dan kemampuannya untuk mengelompokkan dalam jumlah data yang besar dengan waktu komputasi yang cepat dan efisien.

K-means melihat jarak antar objek terhadap nilai *centroid cluster*. *K-means* mengelompokkan secara tegas data hanya pada satu klaster, dengan menggunakan ukuran jarak terkecil terhadap *centroid*.

Adapun langkah-langkah untuk *k-means clustering* adalah sebagai berikut :

1. Tentukan nilai k sebagai jumlah klaster yang diinginkan
2. Pilih k dari set data X sebagai *centroid*. Nilai *centroid* merupakan rata-rata objek dalam klaster tersebut.

$$C_{ij} = \frac{i}{n_i} \sum_{k=1}^{n_i} X_{jk}$$

3. Hitung jarak setiap objek ke masing-masing *centroid* dari masing-masing klaster. Hitung menggunakan jarak *Euclidean*
4. Alokasikan masing-masing objek ke dalam *centroid* yang paling dekat.
5. Hitung kembali *centroid* berdasarkan data yang mengikuti klaster masing-masing.
6. Ulangi langkah 3, 4 dan 5 hingga kondisi konvergen tercapai, yaitu tidak ada data yang berpindah klaster.

Pengertian Pendidikan

Pendidikan adalah suatu proses, teknik dan metode belajar mengajar dengan maksud mentransfer suatu pengetahuan dari seseorang kepada orang lain melalui prosedur yang

sistematis dan terorganisir yang berlangsung dalam jangka waktu yang relatif lama, pendidikan adalah proses mengubah sikap dan tata cara seseorang atau kelompok orang dalam usaha mendewasakan manusia melalui upaya pengajaran dan pelatihan [11-12]. Indikator Pendidikan merupakan besaran kuantitatif mengenai suatu konsep tertentu yang dapat digunakan untuk mengukur proses dan hasil pendidikan atau dampak dari suatu instrumen kebijakan pendidikan.

Indikator Pendidikan

Berdasarkan Profil Pendidikan pada Kementerian Pendidikan dan Kebudayaan tahun 2017, berikut dijelaskan mengenai indikator pendidikan yaitu :

1. Angka Partisipasi Kasar (APK)

APK adalah perbandingan antara banyak siswa pada jenjang pendidikan tertentu dengan banyak penduduk usia sekolah sesuai jenjang pendidikan tersebut yang dinyatakan dalam persentase.

Rumus :

$$APK_j = \frac{\text{siswa } j}{\text{penduduk usia sekolah } j} \times 100\%$$

2. Angka Partisipasi Murni (APM)

APM adalah perbandingan antara siswa usia sekolah tertentu pada jenjang pendidikan tertentu dengan penduduk usia yang sesuai pada jenjang pendidikan tersebut yang dinyatakan dalam persentase.

Rumus :

$$APM_j = \frac{\text{siswa usia } j}{\text{penduduk usia } j} \times 100\%$$

3. Angka Partisipasi Sekolah

APS adalah proporsi dari semua anak yang masih sekolah pada suatu kelompok umur tertentu terhadap penduduk dengan kelompok umur yang sesuai.

Rumus :

$$APS_j = \frac{\text{jumlah penduduk usia } j \text{ yang masih bersekolah}}{\text{jumlah penduduk usia } j} \times 100\%$$

METODE PENELITIAN

Sumber Data

Data yang digunakan dalam penelitian ini berupa data sekunder yang bersumber dari Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Timur berupa data Angka Partisipasi Kasar (APK), Angka Partisipasi Murni (APM) dan Angka Partisipasi Sekolah (APS).

Objek Penelitian

Obyek yang digunakan dalam penelitian ini diambil dari 22 kabupaten/kota, dengan periode waktu pengamatan 3 tahun yaitu tahun 2019-2021.

Variabel Penelitian

Variabel yang digunakan dalam penelitian ini adalah indikator-indikator pendidikan yaitu :

X_1 : Angka Partisipasi Kasar (APK)

X_2 : Angka Partisipasi Murni (APM)

X_3 : Angka Partisipasi Sekolah (APS)

Analisis Data

Sebelum dilakukan analisis *K-Means*, maka terlebih dahulu dilakukan analisis deskriptif untuk melihat karakteristik objek berdasarkan seluruh variabel yang digunakan. Analisis *K-Means* dilakukan dengan tahap-tahap sebagai berikut :

- 1 Menentukan banyaknya *cluster* yang akan dibentuk adalah 4 *cluster*.
- 2 Melakukan pengelompokan menggunakan metode *K-Means clustering* berdasarkan k *cluster* yang telah dibentuk pada langkah 1. Tahapan pengelompokan nya adalah sebagai berikut :
- 3 Pilih k dari set data X sebagai centroid. Nilai centroid merupakan rata-rata objek dalam klaster tersebut
- 4 Hitung jarak setiap objek ke masing-masing *centroid* dari masing-masing klaster. Hitung menggunakan jarak *Euclidean*

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

- 5 Alokasikan masing-masing objek ke dalam *centroid* yang paling dekat.
- 6 Hitung kembali *centroid* berdasarkan data yang mengikuti klaster masing-masing.
- 7 Ulangi langkah b hingga kondisi konvergen tercapai, yaitu tidak ada data yang berpindah klaster.
- 8 Pembentukan hasil klaster.

HASIL DAN PEMBAHASAN

Statistik deskriptif

Statistik deskriptif dilakukan untuk mengetahui gambaran secara umum karakteristik untuk masing-masing indikator atau variabel yang digunakan

Tabel 4.1 Statistik Deskriptif

Variabel	Minimum	Maksimum	Mean	Standar Deviasi
APK	64,34	111,57	83,43	13,29
APM	43,34	70,04	53,86	6,86
APS	63,31	90,77	77,92	6,48

Dari Tabel 4.1 dapat dilihat jumlah minimum, jumlah maksimum dan rata-rata ketiga variabel yaitu angka partisipasi kasar, angka partisipasi murni dan angka partisipasi sekolah di Provinsi Nusa Tenggara Timur pada tahun 2019-2021 secara berturut-turut adalah

Asumsi analisis klaster

Tabel 4. 2 Uji Kaiser-Meyer-Olkin

<i>KMO and Bartlett's Test</i>
<i>Kaiser-Meyer-Olkin Measure of 0,552</i>
<i>Sampling Adequacy.</i>

<i>Bartlett's Test of Sphericity</i>	<i>Approx. Square Df Sig.</i>	<i>Chi- 16,378 3 0,001</i>
--	---	------------------------------------

Berdasarkan Tabel 4.2 diperoleh nilai *Kaiser-Meyer-Olkin Measure of Sampling Adequacy* adalah 0,552. Nilai KMO 0,552 berkisar antara 0,5 sampai 1, sehingga dapat dikatakan bahwa sampel tersebut dapat mewakili populasi dan variabel-variabel tersebut dapat digunakan untuk di analisis lebih lanjut.

Tabel 4. 3 Uji Multikolinearitas

No	Variabel	Tolerance	VIF
1.	X_1	0,928	1,077
2.	X_2	0,693	1,444
3.	X_3	0,622	1,608

Berdasarkan Tabel 4.3 diperoleh nilai *tolerance* pada variabel X_1 adalah 0,928 lebih besar dari 0,1 dan nilai VIF 1,077 kurang dari 10 sehingga tidak terjadi multikolinearitas pada variabel X_1 , nilai *tolerance* pada variabel X_2 adalah 0,693 lebih besar dari 0,1 dan nilai VIF 1,444 kurang dari 10 sehingga tidak terjadi multikolinearitas pada variabel X_2 dan nilai *tolerance* pada variabel X_3 adalah 0,622 lebih besar dari 0,1 dan nilai VIF 1,608 kurang dari 10 sehingga tidak terjadi multikolinearitas pada variabel X_3 . Jadi dapat disimpulkan bahwa tidak terjadi multikolinearitas pada semua variabel, sehingga proses analisis kluster dengan metode *K-Means* bisa dilakukan.

Proses Pengelompokan Menggunakan K-Means Clustering Analysis

Dalam hasil analisis ini, terdapat 3 variabel untuk mengelompokkan Kabupaten/kota di Provinsi Nusa Tenggara Timur berdasarkan indikator pendidikan yaitu angka partisipasi kasar (X_1), angka partisipasi murni (X_2) dan angka partisipasi sekolah (X_3). Sebelum dilakukan uji kluster *K-Means*, data harus memenuhi asumsi klasik dalam *clustering* yaitu uji KMO untuk melihat syarat kecukupan sampel dan uji multikolinearitas untuk menguji ada atau tidaknya variabel independen yang mempunyai kemiripan antar variabel independen lainnya, karena data tersebut memenuhi kedua asumsi yang ada maka data dapat dilanjutkan dengan uji kluster *K-Means*.

Setelah data memenuhi kedua asumsi selanjutnya dilakukan standarisasi data. Standarisasi data dilakukan apabila terdapat perbedaan satuan yang signifikan antara variabel-variabel yang diteliti. Dapat dilihat bahwa variabel yang digunakan memiliki satuan yang bervariasi untuk itu penulis ingin melakukan standarisasi terhadap variabel-variabel yang akan digunakan terlebih dahulu sebelum melakukan pengolahan data.

Pada pengelompokan Kabupaten/Kota di Provinsi Nusa Tenggara Timur berdasarkan indikator pendidikan akan dibentuk 4 kluster. Diperoleh hasil pengelompokan 22 Kabupaten/Kota di Provinsi Nusa Tenggara Timur adalah kluster pertama terdiri dari 5 anggota yaitu Kabupaten Sumba Barat, Kupang, Sikka, Ende, dan Manggarai. Kluster kedua terdiri dari 7 anggota yaitu Kabupaten Alor, Rote Ndao, Manggarai Barat, Sumba Tengah, Sumba Barat Daya, Nagekeo, dan Manggarai Timur. Kluster ketiga terdiri dari 8 anggota yaitu

Kabupaten Sumba Timur, Timor Tengah Selatan, Timor Tengah Utara, Belu, Lembata, Flores Timur, Ngada, dan Malaka, sedangkan kluster keempat memiliki 2 anggota yaitu Kabupaten Sabu Raijua dan Kota Kupang.

Tingkat kebaikan kluster yang terbentuk dapat dilihat melalui jarak antar anggota kluster ke *centroid* dan jarak antar kluster. Semakin tinggi jarak antar kluster maka kluster yang terbentuk semakin baik. Selain itu, apabila jarak antar kluster lebih rendah dibandingkan dengan jarak antar anggota kluster dengan nilai *centroid* maka kluster yang terbentuk sudah baik.

Tabel 4.4 Jarak Antar Anggota Kluster dengan Centroid

Kabupaten/Kota	c_1	c_2	c_3	c_4
Sumba Barat	0,43155	2,26538	1,84380	2,10418
Sumba Timur	2,15655	1,06852	0,57967	3,69897
Kupang	0,40811	2,65103	2,10089	1,71058
Timor Tengah Selatan	2,14840	1,26816	0,37348	3,57515
Timor Tengah Utara	2,08984	1,06448	0,55061	3,55478
Belu	3,63153	2,34754	1,70488	4,74285
Alor	3,70443	1,56111	1,97241	5,28341
Lembata	1,87991	1,04656	0,64365	3,50087
Flores Timur	1,41541	2,02693	1,02864	2,45244
Sikka	1,18048	2,79941	2,37410	2,43764
Ende	0,84506	2,49370	1,59011	1,83289
Ngada	2,47732	2,50749	1,11884	3,13818
Manggarai	1,31382	2,73658	2,90817	2,81245
Rote Ndao	2,20966	0,63655	1,31950	3,92667
Manggarai Barat	2,48290	0,47201	1,63379	4,28680
Sumba Tengah	2,15830	1,00941	1,66787	4,05751
Sumba Barat Daya	2,25579	0,84348	1,93877	4,23850
Nagekeo	2,33703	0,82117	1,21386	3,93887
Manggarai Timur	2,64852	0,44655	1,85416	4,53601
Sabu Raijua	2,22845	4,15385	3,02670	0,79697
Malaka	1,57310	1,94498	1,47072	2,75611
Kota Kupang	2,11087	4,49173	3,78070	0,79696

Tabel 4.5 Jarak Antar Kluster

Kluster	Jarak Antar Kluster			
	1	2	3	4
1	0,000	2,430	2,012	2,019
2	2,430	0,000	1,418	4,252
3	2,012	1,418	0,000	3,330
4	2,019	4,252	3,330	0,000

Berdasarkan Tabel 4.5 dapat kita lihat jarak antar kluster yang paling tinggi adalah jarak kluster 2 dan kluster 4 yaitu 4,252 , kemudian jarak kluster 3 dan kluster 4 yaitu 3,330 , jarak antar kluster 1 dan kluster 4 adalah 2,019, jarak antar kluster 1 dan kluster 2 adalah 2,430, jarak antar kluster 1 dan kluster 3 adalah 2,012 dan terakhir jarak paling dekat adalah

jarak antar kluster 2 dan kluster 3 yaitu 1,418.

Tabel 4.6 Rata-rata Tiap Kluster

Variabel	Kluster			
	1	2	3	4
APK	96,61	72,64	78,58	107,62
APM	55,49	46,80	55,48	68,01
APS	83,39	76,75	73,76	84,94

Berdasarkan Tabel 4.6 diatas dapat disimpulkan bahwa pada kluster pertama angka partisipasi kasar, angka partisipasi murni dan angka partisipasi sekolah memiliki nilai rata-rata yang lebih tinggi dibandingkan dengan kluster kedua dan kluster ketiga sedangkan lebih rendah dari kluster keempat. Pada kluster kedua dapat disimpulkan bahwa angka partisipasi kasar dan angka partisipasi murni memiliki nilai rata-rata terendah dari keempat kluster yang ada sedangkan angka partisipasi sekolah memiliki nilai rata-rata yang lebih tinggi dibandingkan dengan kluster ketiga dan lebih rendah dari kluster pertama dan kluster keempat. Pada kluster ketiga dapat disimpulkan bahwa angka partisipasi kasar dan angka partisipasi murni memiliki nilai rata-rata yang lebih tinggi dibandingkan dengan kluster kedua dan lebih rendah dibandingkan dengan kluster pertama dan kluster keempat, sedangkan angka partisipasi sekolah memiliki nilai rata-rata terendah dari keempat kluster yang ada. Pada kluster keempat dapat disimpulkan bahwa angka partisipasi kasar, angka partisipasi murni dan angka partisipasi sekolah memiliki nilai rata-rata tertinggi dibandingkan dengan keempat kluster yang ada.

Untuk melihat apakah variabel-variabel yang telah membentuk kluster tersebut merupakan variabel pembeda dalam pengelompokan dapat dilihat pada Tabel 4.7 berikut

Tabel 4.7 ANOVA

	Kluster		Error		F	Sig
	Mean		Mean			
	Square	Df	Square	df		
APK	5,741	3	0,210	18	27,350	0,000
APM	5,548	3	0,242	18	22,927	0,000
APS	3,145	3	0,642	18	4,895	0,012

Validasi kluster yaitu uji parsial F dengan hipotesis :

H_0 : variabel i bukan variabel pembeda dalam pengelompokan,

H_1 : variabel i merupakan variabel pembeda dalam pengelompokan.

Kriteria Uji : Tolak H_0 jika $F > F_{\alpha, k-1, -k}$

Dipilih $\alpha = 0,05$ dan jumlah objek penelitian (n) = 22 sedangkan kluster yang dibentuk (k) adalah 4 sehingga dari tabel distribusi F diperoleh nilai $F_{\alpha, k-1, -k} = 3,16$

Berdasarkan Tabel 4.7 (Anova) dapat dilihat bahwa nilai F_{hitung} dari ketiga variabel yang digunakan lebih besar dari F_{tabel} sehingga dapat disimpulkan tolak H_0 atau terima H_1 , artinya ketiga variabel yang digunakan merupakan variabel pembeda dalam pengelompokan

sehingga kluster yang terbentuk valid.

Pembahasan

Dalam hasil analisis ini, terdapat 3 variabel untuk mengelompokkan Kabupaten/kota di Provinsi Nusa Tenggara Timur berdasarkan indikator pendidikan yaitu angka partisipasi kasar (X_1), angka partisipasi murni (X_2) dan angka partisipasi sekolah (X_3). Sebelum dilakukan uji kluster *K-Means*, data harus memenuhi asumsi klasik dalam *clustering* yaitu uji KMO untuk melihat syarat kecukupan sampel dan uji multikolinearitas untuk menguji ada atau tidaknya variabel independen yang mempunyai kemiripan antar variabel independen lainnya, karena data tersebut memenuhi kedua asumsi yang ada maka data dapat dilanjutkan dengan uji kluster *K-Means*.

Berdasarkan hasil pengelompokan objek menggunakan analisis kluster *k-means*. Dimana objek yang di kelompokkan adalah indikator pendidikan dengan mengambil nilai rata-rata pada tahun 2019-2021. Proses *clustering k-means* memproses semua objek penelitian secara sekaligus, dimana *k* adalah banyaknya kelompok yang telah ditentukan terlebih dahulu, proses penentuan jumlah *k* didasarkan pada kasus yang akan diteliti yaitu indikator pendidikan yang bertujuan untuk melihat tingkat pendidikan setiap daerah yang ada di Provinsi Nusa Tenggara Timur. Pada penelitian ini, peneliti membagi kluster dalam 4 kategori yaitu kategori rendah, sedang, cukup tinggi dan tinggi.

Pada kluster pertama berdasarkan tafsiran angka yang didapat dapat disimpulkan bahwa angka partisipasi kasar, angka partisipasi murni dan angka partisipasi sekolah memiliki nilai rata-rata yang lebih tinggi dibandingkan dengan kluster kedua dan kluster ketiga sedangkan lebih rendah dari kluster keempat. Maka dapat disimpulkan bahwa kluster pertama masuk dalam kategori pendidikan cukup tinggi dengan kabupaten yang masuk dalam kluster ini adalah Kabupaten Sumba Barat, Kupang, Sikka, Ende dan Manggarai.

Pada kluster kedua berdasarkan tafsiran angka yang didapat dapat disimpulkan bahwa angka partisipasi kasar dan angka partisipasi murni memiliki nilai rata-rata terendah dari keempat kluster yang ada sedangkan angka partisipasi sekolah memiliki nilai rata-rata yang lebih tinggi dibandingkan dengan kluster ketiga dan lebih rendah dari kluster pertama dan kluster keempat. Maka dapat disimpulkan bahwa kluster kedua masuk dalam kategori pendidikan rendah dengan kabupaten yang masuk dalam kluster ini adalah Kabupaten Alor, Rote Ndao, Manggarai Barat, Sumba Tengah, Sumba Barat Daya, Nagekeo dan Manggarai Timur.

Pada kluster ketiga berdasarkan tafsiran angka yang didapat dapat disimpulkan bahwa angka partisipasi kasar dan angka partisipasi murni memiliki nilai rata-rata yang lebih tinggi dibandingkan dengan kluster kedua dan lebih rendah dibandingkan dengan kluster pertama dan kluster keempat, sedangkan angka partisipasi sekolah memiliki nilai rata-rata terendah dari keempat kluster yang ada. Maka dapat disimpulkan bahwa kluster ketiga masuk dalam kategori pendidikan sedang dengan kabupaten yang masuk dalam kluster ini adalah Kabupaten Sumba Timur, Timor Tengah Selatan, Timor Tengah Utara, Belu, Lembata, Flores Timur, Ngada dan Malaka.

Pada kluster keempat berdasarkan tafsiran angka yang didapat dapat disimpulkan bahwa angka partisipasi kasar, angka partisipasi murni dan angka partisipasi sekolah memiliki nilai rata-rata tertinggi dibandingkan dengan keempat kluster yang ada. Maka dapat disimpulkan bahwa kluster keempat masuk dalam kategori pendidikan tinggi dengan kabupaten yang masuk dalam kluster ini adalah Kabupaten Sabu Raijua dan Kota Kupang.

Hasil kluster yang didapat berdasarkan uji ANOVA menunjukkan bahwa semua variabel yang digunakan mempunyai perbedaan yang signifikan pada keempat kluster yang terbentuk dan variabel yang memiliki perbedaan paling besar pada keempat kluster yang terbentuk adalah variabel angka partisipasi kasar dengan nilai F sebesar 27,350 dan signifikan 0,000. Karena variabel angka partisipasi kasar memiliki perbedaan paling besar terhadap kluster yang terbentuk, maka Kabupaten/kota yang memiliki jumlah angka partisipasi kasar terbesar masuk dalam kluster keempat.

KESIMPULAN

Berdasarkan hasil pengelompokan objek menggunakan analisis kluster *k-means*. Dimana objek yang di kelompokkan adalah indikator pendidikan angka partisipasi kasar, angka partisipasi murni dan angka partisipasi sekolah dengan mengambil nilai rata-rata pada tahun 2019-2021, maka dapat disimpulkan bahwa dari 4 kluster yang dibentuk dalam penelitian ini diperoleh kluster pertama terdiri dari terdiri dari 5 anggota yaitu Kabupaten Sumba Barat, Kupang, Sikka, Ende, dan Manggarai, di mana anggota pada kluster pertama merupakan kabupaten yang penanganan pendidikannya menjadi prioritas ketiga. Kluster kedua terdiri dari 7 anggota yaitu Kabupaten Alor, Rote Ndao, Manggarai Barat, Sumba Tengah, Sumba Barat Daya, Nagekeo, dan Manggarai Timur, di mana anggota pada kluster kedua merupakan kabupaten yang penanganan pendidikannya menjadi prioritas pertama. Kluster ketiga terdiri dari 8 anggota yaitu Kabupaten Sumba Timur, Timor Tengah Selatan, Timor Tengah Utara, Belu, Lembata, Flores Timur, Ngada, dan Malaka, di mana anggota pada kluster ketiga merupakan kabupaten yang penanganan pendidikannya menjadi prioritas kedua. Sedangkan kluster keempat memiliki 2 anggota yaitu Kabupaten Manggarai Timur dan Kota Kupang, anggota pada kluster keempat merupakan kabupaten yang penanganan pendidikannya menjadi prioritas keempat.

DAFTAR PUSTAKA

- [1] Badan Pusat Statistik. 2019. *Indikator Pendidikan 2019-2021*. Kupang : BPS NTT
- [2] Akramunnisa & Fajriani. 2017. "*K-Means Clustering Analysis pada Persebaran Tingkat Pengangguran Kabupaten/Kota di Sulawesi Selatan*". *Jurnal Varian* Vol. 3 No. 2 April 2020
- [3] Elok, F Rafikasari. 2016. "*Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Pendidikan Tahun 2013 Menggunakan Analisis Hierarchial Cluster*". *Jurnal Varian* Vol. 16 No. 2 September 2022
- [4] Hanna, S Karti. 2013. "*Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Pendidikan SMA/SMK/MA dengan Metode C-Means dan Fuzzy C-Means*". *Jurnal Varian* Vol. 2 No. 2 September 2022
- [5] Priskila, A Selly. 2021. "*K-Means Clustering Analysis pada Pengelompokan Kabupaten di Provinsi Nusa Tenggara Timur Berdasarkan Tingkat Pengangguran*". Skripsi. Diterbitkan. Jurusan Matematika Fakultas Sains dan Teknik. Universitas Nusa Cendana (UNDANA) : Kupang.
- [6] Santoso, S. 2002. *Statistik Multivariat Edisi Revisi konsep dan Aplikasi dengan SPSS*. Jakarta : PT Elex Media Komputindo.
- [7] Sigit, Nugroho. 2008. *Statistik Multivariat Terapan*. Bengkulu : UNIB Press.

- [8] Hair, J.R., Anderson, R.E., Tatham, R.L., Black W.C., 2006, *Multivariate Data Analysis with Readings*, 3th Edition, Macmillan Publishing Company, New York.
- [9] Tony, Wijaya dan Santi, Budiman. 2016. *Analisis Multivariate untuk Penelitian Manajemen*. Yogyakarta : Pohon Cahaya.
- [10] Prasetyo, E. 2012. *Data Mining Dan Aplikasi Menggunakan Matlab*. Yogyakarta : Andi.
- [11] Landasan Teori Pengertian Pendidikan di Indonesia. (Diakses pada 14 September (2022) dari <http://repository.uin-suska.ac.id>)
- [12] Pusat Bahasa Departemen Pendidikan Nasional . 2013. Pendidikan Diakses dari <http://www.sarjanaku.com/2012/12/pengertian-pendidikan-menurut-para-ahli.html> 11 Agustus 2022. Pukul 08.40 WIB

3376

JCI

Jurnal Cakrawala Ilmiah

Vol.2, No.9, Mei 2023

HALAMAN INI SENGAJA DIKOSONGKAN