

Visualisasi Data Pokok Pendidikan Provinsi Papua Semester Genap Tahun Ajaran 2021/2022 dengan Menggunakan Metode *K-Means*

Afnita Rahma Auliya Putri (221911091, 3SD1)

Dosen Mata Kuliah Visualisasi Data dan Informasi: Farid Ridho

Ringkasan—Kondisi pendidikan Provinsi Papua yang terbelakang di antara provinsi di Indonesia, serta belum adanya pemanfaatan data pokok pendidikan (DAPODIK) padahal data tersebut berpotensi besar sebagai pertimbangan pembangunan pendidikan di suatu wilayah, melatarbelakangi penelitian “Visualisasi Data Pokok Pendidikan Provinsi Papua Semester Genap Tahun Ajaran 2021/2022 dengan Menggunakan Metode *K-Means*” ini. Analisis menggunakan alat *Microsoft Excel 2019*, *Geoda*, *QGIS 3.16*, *R Studio*, *Tableau Desktop*, dan *Tableau Public*, menghasilkan dua klaster, dengan klaster yang perlu diperhatikan pendidikannya—karena terdapat ketimpangan rasio murid terhadap guru serta sekolah—adalah Kota Jayapura, Kab. Pegunungan Bintang, Kab. Yahukimo, Kab. Jayawijaya, Kab. Lanny Jaya, Kab. Tolikara, Kab. Puncak Jaya, Kab. Mimika, Kab. Deiyai, Kab. Paniai, Kab. Dogiyai, dan Kab. Mappi. Hasil klaster disajikan dalam peta tematik, didukung oleh *side-by-side bar chart* interaktif, yang dapat diakses pada laman https://public.tableau.com/views/UjianAkhirSemesterUASVisualisasiDataDanInformasi/Dashboard1?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link. Selain itu, semua hal terkait pengerjaan penelitian ini dapat diakses dalam tautan *GitHub* berikut <https://github.com/aafnitarahma/221911091-3SD1-VisDat>.
Kata Kunci—*k-means*, *bar chart*, guru, murid, sekolah.

I. LATAR BELAKANG

Pendidikan merupakan salah satu sektor penting dalam pembangunan di setiap negara. Berhasil tidaknya pendidikan yang dilaksanakan akan menentukan maju mundurnya negara tersebut (Taufik A. K., 2012). Hal ini sejalan dengan pernyataan Mendikbud, Muhadjir Effendy, bahwa dunia pendidikan sangat penting guna membingkai hadirnya sumber daya manusia yang berkualitas, demi terwujudnya Indonesia yang berkembang.

Ukuran keberhasilan pendidikan di suatu wilayah dapat tercermin dari nilai Angka Partisipasi Kasar (APK), Angka Partisipasi Murni (APM), dan/atau Angka Partisipasi Sekolah (APS) wilayah terkait (DPMPSTP, 2019). Angka Partisipasi Kasar (APK) adalah perbandingan antara jumlah penduduk yang masih bersekolah di jenjang pendidikan tertentu (tanpa memandang usia penduduk tersebut) dengan jumlah penduduk yang memenuhi syarat resmi penduduk usia sekolah di jenjang pendidikan yang sama, Angka Partisipasi Murni (APM) adalah proporsi dari penduduk kelompok usia sekolah tertentu yang sedang bersekolah tepat di jenjang pendidikan yang seharusnya (sesuai antara umur penduduk dengan ketentuan usia

bersekolah di jenjang tersebut) terhadap penduduk kelompok usia sekolah yang bersesuaian, sedangkan Angka Partisipasi Sekolah (APS) adalah proporsi dari penduduk kelompok usia sekolah tertentu yang sedang bersekolah (tanpa memandang jenjang pendidikan yang ditempuhi) terhadap penduduk kelompok usia sekolah yang bersesuaian. Sejak tahun 2007, Pendidikan Non Formal (Paket A, Paket B, dan Paket C) turut diperhitungkan dalam menghitung angka-angka tersebut (Sirusa BPS).

Berdasarkan data Susenas BPS, provinsi Papua adalah provinsi dengan APM terendah di semua tingkatan sekolah (secara berurutan dari SD/Sederajat sampai SMA/Sederajat adalah 2019: 79.19, 57.19, 44.32; 2020: 79.34, 57.95, 44.73; 2021: 80.38, 57.93, 44.41) dan dengan APS terendah di semua rentang umur (secara berurutan dari 7-12 sampai 16-18 adalah 2019: 82.67, 80.13, 63.50; 2020: 82.99, 80.48, 64.83; 2021: 83.43, 80.02, 63.98) selama tiga tahun terakhir—2019, 2020, dan 2021—berturut-turut. Untuk APK-nya, dalam tiga tahun terakhir, Papua terendah di tingkat pendidikan SD/ sederajat dan SMP/ sederajat, dan di tingkat SMA/ sederajat terendah di tahun 2021—tetapi secara umum, di tahun 2019 dan 2020 masih masuk tiga terbawah. Dengan demikian, dapat dikatakan bahwa pendidikan di Papua terbelakang dibandingkan provinsi lain di Indonesia.

Variabel rasio murid terhadap guru dan rasio murid terhadap sekolah berpengaruh signifikan terhadap APK SMA/ Sederajat (Rahmadina, R., P., *et al.*, 2021). Kedua variabel tersebut juga berpengaruh signifikan terhadap APS SMP/ Sederajat (Rondong, F.), serta terhadap APM SD/ Sederajat (Bappenas, 2009). Variabel penyusunnya (rasio murid terhadap guru dan rasio murid terhadap sekolah), yakni jumlah sekolah, jumlah guru, dan jumlah murid di setiap tingkatan pendidikan tersedia dalam web Data Pokok Pendidikan Kemendikbud (<https://dapo.kemdikbud.go.id/>). Data tersebut sejauh ini telah divisualisasikan dalam *dashboard* Kemendikbud berikut (<https://referensi.data.kemdikbud.go.id/dashboardgtk/telusur.php?id=30#>), tetapi hanya memuat data sebaran guru—tidak semua data DAPODIK divisualisasikan—serta untuk mengaksesnya diperlukan daya besar (jaringan internet sangat kuat dan *device* yang memadai). Untuk itu, peneliti merasa perlu untuk memvisualisasikan DAPODIK demi pemanfaatan lebih dan publikasi yang lebih luas, terutama yang berfokus pada Provinsi Papua yang kondisi pendidikannya terbelakang.

Dengan demikian, dengan variabel DAPODIK yang berpengaruh terhadap APK, APM, dan/atau APS, akan

dilakukan analisis kluster untuk mengelompokkan kabupaten dan kecamatan Provinsi Papua berdasarkan kondisi pendidikannya. Kemudian, hasil analisis tersebut akan disajikan dalam visualisasi peta supaya mempermudah penyampaian informasi kepada masyarakat. Hasil penelitian ini, dalam jangka panjang, diharapkan mampu menjadi acuan pemerintah dalam menyusun prioritas daerah di Papua yang perlu lebih diperhatikan pendidikannya serta menumbuhkan kesadaran masyarakat mengenai kondisi pendidikan di provinsi paling timur Indonesia.

II. TUJUAN PENELITIAN

Berdasarkan latar belakang yang telah diuraikan, tujuan penelitian ini adalah:

- 1) Mengetahui karakteristik variabel data pokok pendidikan lingkup kabupaten di Provinsi Papua pada setiap kluster yang terbentuk dari hasil pengelompokan dengan metode *K-Means*, serta
- 2) Menghasilkan peta hasil kluster lingkup kabupaten Provinsi Papua dan visualisasi yang sesuai untuk menyajikan data pokok pendidikan tingkat kecamatan di Provinsi Papua.

III. PENELITIAN TERKAIT

Penelitian ini terinspirasi oleh penelitian Wang, H. *et al.* pada 2021, yakni visualisasi hierarkis data area geografis dengan asosiasi atribut spasial. Untuk menjelajahi fitur global dan detail karakteristik pada saat yang sama dan mendapatkan wawasan mendalam tentang fitur potensial, mereka menggabungkan data area geografis ke dalam kelompok hierarkis berdasarkan kesamaan spasial yang diukur dengan jarak, kemudian hasilnya disajikan dalam peta. Dengan itu, peneliti ingin melakukan versi sederhana dari penelitian tersebut, yakni tanpa mempertimbangkan efek spasial, akan dilakukan pengelompokan non hierarki (lebih tepatnya *K-Means*) kemudian data akan disajikan dalam bentuk peta, sedangkan rincian untuk wilayah lebih kecil akan disajikan dalam bentuk *side-by-side bar chart*, seperti dalam penelitian Yang, Y. *et al.* (2021), yang melengkapi visualisasi data yang terkait fitur area dengan *bar chart*. Alasan peneliti memilih *side-by-side bar chart* dibandingkan visualisasi data hierarki lain, seperti *sunburst*, *treemap*, *icicle*, dsb. adalah karena banyaknya kategori yang ada (kategorinya adalah distrik dan sub-distrik, yang berarti akan kacau jika divisualisasikan dalam bentuk lingkaran), juga karena banyaknya nilai nol (0) yang mana tidak akan muncul jika digunakan visualisasi tersebut—padahal data bernilai 0 (nol) tersebut penting dan tetap harus terlihat di visualisasinya.

IV. METODE PENELITIAN

Suatu kesatuan individu atau subjek pada wilayah dan waktu dengan kualitas tertentu yang akan diamati atau diteliti disebut populasi. Berdasarkan ketercakupannya, populasi dapat dibedakan menjadi dua, yaitu populasi terbatas dan populasi tidak terbatas (Supardi, 2016). Seluruh kabupaten dan kecamatan di Provinsi Papua menjadi populasi dari penelitian ini, yang meliputi 29 kabupaten dan 576 kecamatan. Selanjutnya, bagian dari populasi yang dijadikan subjek

penelitian sebagai yang mewakili anggota populasi disebut sampel penelitian (Supardi, 2016). Seluruh anggota populasi dalam penelitian ini juga bertindak sebagai sampel karena kepentingan penelitian untuk mengelompokkan semua wilayah.

Penelitian ini menggunakan data sekunder, yakni yakni data yang pengumpulannya tidak diusahakan sendiri oleh peneliti, dengan mengambil data dari situs Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi (Kemendikbud). Data yang diambil adalah Data Pokok Pendidikan (DAPODIK) Semester Genap tahun ajaran 2021/2022, dengan pertimbangan data tersebut paling mewakili kondisi saat ini, berupa:

1. Data jumlah sekolah SD, SMP, SMA, dan SMK per kabupaten di Provinsi Papua (<https://dapo.kemdikbud.go.id/sp/1/250000>)
2. Data jumlah guru SD, SMP, SMA, dan SMK per kabupaten di Provinsi Papua (<https://dapo.kemdikbud.go.id/guru/1/250000>)
3. Data jumlah peserta didik SD, SMP, SMA, dan SMK per kabupaten di Provinsi Papua (<https://dapo.kemdikbud.go.id/guru/1/250000>)

yang selain terinci dalam kabupaten juga terinci dalam kecamatan, sehingga membentuk suatu data berhierarki. Selain data tersebut, diperlukan juga data geospasial Provinsi Papua, yang didapatkan dalam laman https://gadm.org/download_country.html (dalam bentuk satu/seluruh wilayah Indonesia). Data dalam tautan tersebut masih mengacu pada batas administrasi wilayah Indonesia per tahun 2015. Untuk data yang terkait dengan penelitian ini, data pada level kabupaten sesuai (yakni sebanyak 29 kabupaten), tetapi untuk level kecamatan tidak sesuai (hanya tersedia 385 kecamatan, sementara seharusnya terdapat 576 kecamatan di Provinsi Papua). Karena itu, analisis hanya akan dilakukan sampai batas kabupaten, sedangkan data level kecamatan akan disajikan dalam visualisasi yang sesuai.

Sebagai catatan, per tanggal 9 Juni 2022, yakni ketika peneliti mengumpulkan data, DAPODIK tersebut tersinkronisasi sebesar 83 persen—masih terdapat wilayah (sebanyak 17% dari total wilayah di Papua) yang belum melaporkan DAPODIK ke Kemendikbud. Dengan kata lain, terdapat observasi berupa *missing values*. Setelah diperiksa, data yang belum ada adalah data: Kecamatan Libarek, Pisugi, Popugoba, Silo Karno Doga, Trikora, Walaik, Welesi, dan Wita Waya di Kabupaten Jayawijaya; Kecamatan Pulau Kurudu, Pulau Yerui, Anotarei, dan Yawakukat di Kabupaten Kepulauan Yapen; Kecamatan Wege Muka, Wege Bino, Pugo Dagi, Muye, Nakama, Teluk Deya, Yagai, Youtadi, Baya Biru, Deiyai Miyu, Dogomo, Aweida, Topiyai, dan Fajar Timur di Kabupaten Paniai; Kecamatan Pagaleme, Gurage, Ilamburawi, Yambi, Lumo, Molanikme, Dokome, Kalome, Wanwi, Yamoneri, Waegi, Gubume, dan Taganombak di Kabupaten Puncak Jaya; Kecamatan Aswi, Der Koumur, Bectbamu, Kopay, Awyu, Joutu, Koroway Buluanop, Unir Sirau, Suru-Suru, dan Pulau Tiga di Kabupaten Asmat; Kecamatan Tagime, Danime, dan Yuko di Kabupaten Tolikara; Kecamatan Ismari, Bonggo Barat, Veen, Verkam, Muara Tor, Sungai Biri, Apawer Hilir, dan Apawer Tengah di Kabupaten Sarmi; Kecamatan Soyoi Mambai, Wonti, dan Walani di Kabupaten Waropen; Kecamatan Iwaso di Kabupaten Mamberamo Raya; Kecamatan

Wusi, Pasir Putih, Iniye, Mam, Alama, Yenggelo, Mbuwa Tengah, Pija, dan Moba di Kabupaten Nduga; Kecamatan Gollo, Wano Barat, Yugungwi, Nogi, Tiom Olo, Milimbo, Wiringgabut, Ayumnati, Kelulome, Lannyna, Yiginua, Niname, Nikogwe, Mokoni, Werek, Awina, Gupura, Bruwa, Karu, Melagai, Goa Balim, Yiluk, Kolawa, Balingga Barat, Gelok Beam, Guna, Muara, Kully Lanny, dan Buguk Gona di Kabupaten Lanny Jaya; Kecamatan Dervos, Beoga Timur, Ogamanin, Bina, Kembru, Pogoma, Mabugi, Gome Utara, Amungkalpia, dan Lambewi di Kabupaten Puncak; serta Kecamatan Tomosiga di Kabupaten Intan Jaya. Terlepas dari banyaknya data yang tidak diketahui, peneliti tetap menggunakan data ini dengan pertimbangan persentase sinkronisasi data terkait pada periode sebelumnya untuk Provinsi Papua-pun di sekitar nilai 80-an persen. Misalnya untuk Semester Ganjil 2021/2022 (85%), Semester Genap 2020/2021 (85%), dan Semester Ganjil 2020/2021 (85%).

Analisis yang digunakan dalam penelitian ini adalah analisis *clustering* dengan metode *K-Means*. Alat yang digunakan untuk *pre-processing* data, melakukan *K-Means clustering*, hingga membuat peta hasil kluster adalah *Microsoft Excel 2019*, *Geoda*, *QGIS 3.16*, *R Studio*, *Tableau Desktop*, dan *Tableau Public*.

Rincian tahapan yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Merumuskan masalah dan tujuan penelitian
2. Mengumpulkan data
3. Mengolah dan menganalisis data, termasuk/sekalius memvisualisasikan data
4. Menginterpretasi hasil
5. Menarik kesimpulan

Pre-processing data dilakukan dengan menyiapkan tabel untuk pengolahan, yakni dengan menyatukan semua DAPODIK kabupaten dalam satu *dataset* menggunakan *Microsoft Excel 2019*. Selain itu, dilakukan juga pemotongan *polygon (shp)* level kabupaten dan kecamatan di Indonesia untuk mengambil kabupaten dan kecamatan di Provinsi Papua saja—menggunakan *QGIS 3.16*, lalu penggabungan variabel DIPODIK dan data geospasial menggunakan *Geoda*.

Analisis *clustering* dilakukan dengan metode *K-Means* menggunakan *Tableau* dan *R Studio* secara simultan. *K-Means* merupakan salah satu algoritma dalam metode non hirarki yang paling sederhana. Algoritma ini mengelompokkan data menjadi *k cluster*, sehingga data yang memiliki kesamaan karakteristik dimasukkan ke dalam satu kelompok yang sama dan data dengan karakteristik berbeda dimasukkan ke dalam kelompok lainnya. Secara umum, proses dalam algoritma *K-Means Clustering* adalah sebagai berikut,

1. Menentukan jumlah *cluster (k)*.
2. Menginisialisasi secara acak nilai titik pusat *cluster* sejumlah *k*.
3. Menghitung jarak setiap data terhadap titik pusat *cluster* masing-masing dengan menggunakan jarak *Euclidean*. Jarak *Euclidean* antara objek *i* dan *j* dirumuskan dengan:

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

dengan

d_{ij} : jarak *Euclidean* dari data ke-*i* ke pusat *cluster j*

x_{ik} : nilai objek ke-*i* pada variabel ke-*k*

x_{jk} : nilai objek ke-*j* pada variabel ke-*k*

p : banyak variabel yang diamati

4. Mengelompokkan data berdasarkan jarak terdekat antara setiap data terhadap titik pusat *cluster*.
5. Menentukan titik pusat *cluster* yang baru dengan menghitung rata-rata seluruh data pada pusat *cluster* yang sama. Berikut persamaan titik pusat *cluster*:

$$C_{kj} = \frac{x_{1j} + x_{2j} + \dots + x_{aj}}{a}, j = 1, 2, \dots, p \quad (2)$$

dengan

C_{kj} : pusat *cluster* ke-*k* variabel ke-*j*

a : banyak data pada *cluster* ke-*k*

6. Mengulangi tahap ke-3 sampai ke-5 hingga titik pusat *cluster* tetap dan anggota *cluster* tidak berpindah (konvergen).

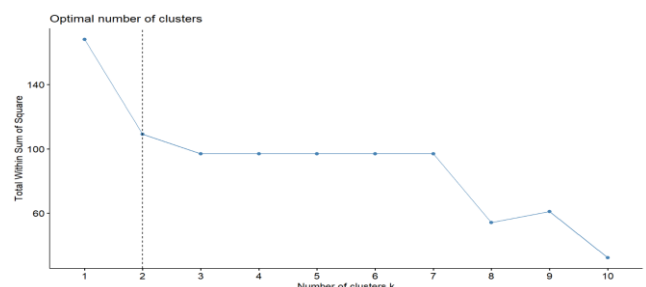
Untuk penentuan jumlah *cluster* optimal, terdapat beberapa metode. Dalam penelitian ini digunakan *silhouette method* dan *elbow method*.

Metode *silhouette* digunakan untuk menentukan jumlah *cluster* optimal dengan menggunakan data skala rasio. Algoritma ini akan menghitung jarak rata-rata dari suatu objek terhadap seluruh objek yang ada di dalam *cluster* yang sama dengan objek di *cluster* lainnya.

Metode *elbow* menentukan jumlah *cluster* optimal dengan menggunakan siku yang terbentuk pada suatu titik di grafik SSE dan didasarkan pada penurunan SSE yang besar. Metode ini menggunakan nilai *Sum of Square Error (SSE)* dari tiap jumlah *cluster*. Makin banyak *cluster* yang digunakan, maka nilai SSE akan makin mengecil, sehingga jumlah *cluster* terbaik adalah jumlah *cluster* dengan penurunan nilai SSE terbesar.

V. HASIL DAN PEMBAHASAN

Dalam analisis *clustering*, harus ditetapkan terlebih dahulu jumlah kluster (*k*). Terdapat beberapa cara untuk menentukan jumlah kluster optimal, tetapi dalam penelitian ini akan digunakan metode *elbow* dan metode *silhouette*. Metode *elbow*, yakni dengan mengukur nilai *within sum-square* dalam bentuk grafik ditampilkan pada Gambar 1.

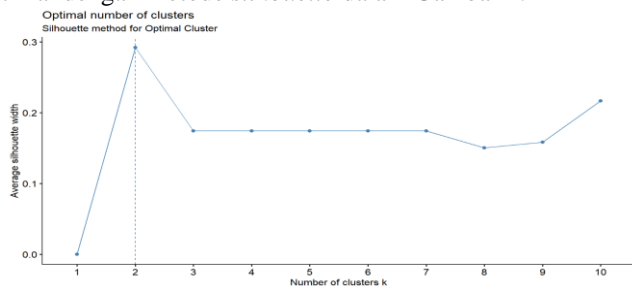


Gambar 1. Grafik metode *elbow* untuk menentukan jumlah kluster optimal

Dari grafik metode *elbow* di atas, terlihat pada jumlah *cluster* 2 terdapat lengan siku, yang berarti nilai *within sum-squares* mulai menunjukkan penurunan secara gradual. Dengan

pertimbangan tersebut, serta ketidak efisienan interpretasi jika jumlah kluster terlalu banyak, jumlah kluster optimal berdasarkan metode *elbow* adalah 2.

Hasil tersebut didukung oleh hasil pencarian jumlah kluster optimal dengan metode *silhouette* dalam Gambar 2.



Gambar 2. Grafik metode *silhouette* untuk menentukan jumlah kluster optimal

Pada metode *silhouette*, jumlah kluster optimal merupakan nilai estimasi dari jumlah kluster k yang mampu memaksimalkan nilai rata-rata indeks validitas *silhouette* $S_{(i)}$. Untuk mempermudah pengamatan perubahan indeks validitas *silhouette*, setiap perubahan jumlah kluster k dibentuk ke dalam plot Gambar 2.

Analisis *clustering* dalam penelitian ini dilakukan secara simultan menggunakan *Tableau* dan *R Studio*. Penentuan jumlah kluster optimal dalam analisis *clustering* menggunakan *R Studio* telah dijelaskan sebelumnya, sedangkan dengan *Tableau* diatur peneliti untuk ditentukan otomatis. Hasil kluster optimal yang terbentuk adalah sama, yakni sebanyak 2 kluster.

Tabel 1 berikut menunjukkan perbandingan hasil *clustering*—menggunakan *Tableau* dan *RStudio*—dengan 6 karakteristik (Rasio Murid terhadap Guru SD (x_1), Rasio Murid terhadap Guru SMP (x_2), Rasio Murid terhadap Guru SMA/SMK (x_3), Rasio Murid terhadap SD (x_4), Rasio Murid terhadap SMP (x_5), Rasio Murid terhadap Guru SMA/SMK (x_6)) sebagai atribut.

TABEL I

Perbandingan Keanggotaan Kluster berdasarkan Alat Pengolahan Datanya

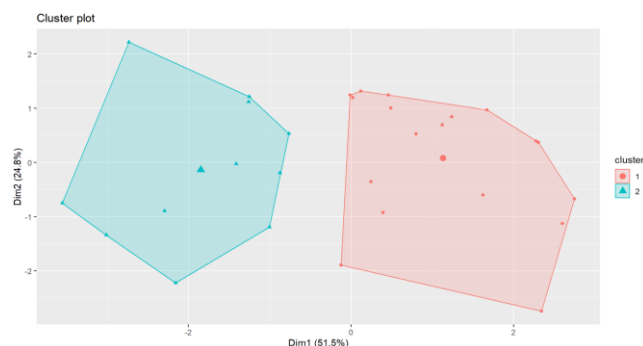
Alat Pengolahan	Anggota Kluster 1 ^a	Anggota Kluster 2
<i>Tableau</i>	Kota Jayapura, Kab. Pegunungan Bintang, Kab. Yahukimo, Kab. Jayawijaya, Kab. Lanny Jaya, Kab. Tolikara, Kab. Puncak Jaya, Kab. Mimika, Kab. Deiyai, Kab. Paniai, Kab. Dogiyai, Kab. Mappi	Kab. Nabire, Kab. Kepulauan Yapen, Kab. Biak Numfor, Kab. Supiori, Kab. Waropen, Kab. Mamberamo Raya, Kab. Puncak, Kab. Intan Jaya, Kab. Sarmi, Kab. Yalimo, Kab. Mamberamo Tengah, Kab. Keerom, Kab. Jayapura, Kab. Nduga, Kab. Asmat, Kab. Merauke, dan Kab. Boven Digoel
<i>R Studio</i>	Kota Jayapura, Kab. Yahukimo, Kab. Jayawijaya, Kab. Lanny Jaya, Kab. Tolikara, Kab. Puncak Jaya, Kab. Mimika, Kab. Deiyai,	Kab. Nabire, Kab. Kepulauan Yapen, Kab. Biak Numfor, Kab. Supiori, Kab. Waropen, Kab. Mamberamo Raya, Kab. Puncak, Kab. Intan Jaya, Kab. Sarmi, Kab.

	Kab. Paniai, Kab. Dogiyai, Kab. Mappi	Yalimo, Kab. Mamberamo Tengah, Kab. Keerom, Kab. Jayapura, Kab. Nduga, Kab. Asmat, Kab. Merauke, Kab. Pegunungan Bintang, dan Kab. Boven Digoel
--	---------------------------------------	---

^aPemberian nama kolom dalam Tabel I mengikuti kluster *Tableau*. Artinya, Kluster 1 *Tableu* sama dengan Kluster 2 *R Studio*.

Dari Tabel I, diketahui bahwa hasil perbedaan pengklasifikasian antara *Tableau* dan *R Studio* terletak pada Kabupaten Pegunungan Bintang. Selain itu, klasifikasi wilayah lain sama. Perbedaan hasil *cluster* tersebut kemungkinan besar dipengaruhi oleh pemilihan titik awal random yang berbeda. Karena pendekatan acak untuk memulai proses analisis kluster, sangat mungkin bahwa pusat segmen yang berbeda (rata-rata/sarana) akan terbentuk. Hal ini dapat terjadi bahkan ketika menggunakan kumpulan data yang sama dengan paket statistik yang sama dan hanya menjalankan kembali data ke dalam jumlah segmen yang sama.

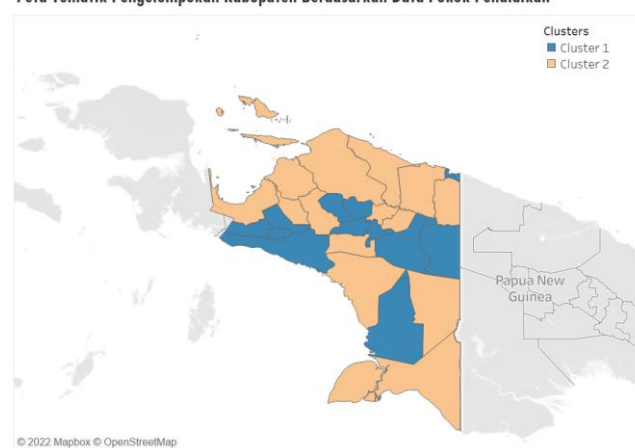
Gambar 3 menunjukkan visualisasi hasil *clustering* menggunakan *R Studio* dengan metode *K-Means* ($k = 2$). Dari plot terlihat tidak ada daerah yang saling tumpang tindih, artinya pengelompokan sudah baik.



Gambar 3. Cluster plot *K-Means* dengan jumlah kluster $k = 2$ (output *R Studio*)

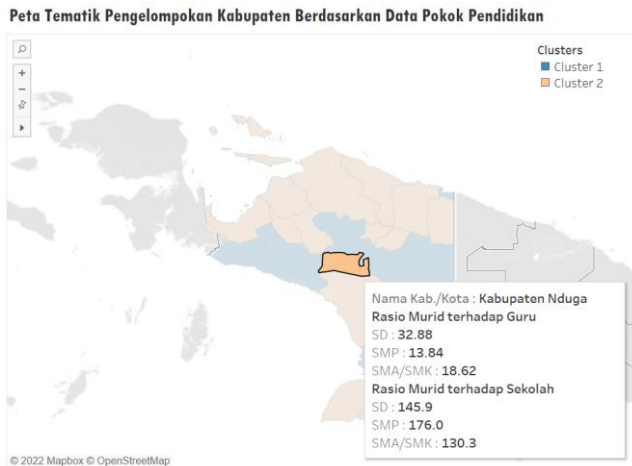
Hasil pengelompokan *Tableau* dengan metode *clustering* yang sama, disajikan dalam peta tematik pada Gambar 4 berikut.

Peta Tematik Pengelompokan Kabupaten Berdasarkan Data Pokok Pendidikan



Gambar 4. Peta tematik hasil *clustering* dengan metode *K-Means* jumlah klaster $k = 2$ (output Tableau)

Visualisasi Gambar 4 dapat dilihat dalam tautan berikut https://public.tableau.com/views/UjianAkhirSemesterUASVisualisasiDatadanInformasi/Dashboard1?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link. Peta tematik tersebut menggunakan palet warna yang ramah bagi penderita buta warna. Selain itu, peta tersebut juga dilengkapi fitur *highlight*, supaya pengguna dapat fokus ke wilayah yang diinginkannya, serta fitur *hover* yang menampilkan informasi berupa 6 karakteristik yang digunakan sebagai atribut pengklasteran. Gambar 5 berikut menunjukkan penggunaan *highlight* dan tampilan informasi yang muncul.



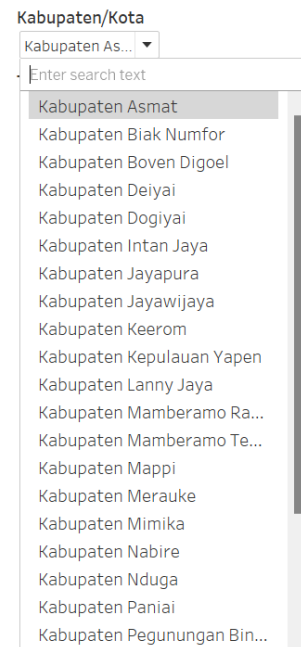
Gambar 4. Tampilan peta ketika kursor melayar di daerah terkait

Dengan tujuan merinci hasil klaster ini, yakni menunjukkan perbandingan data tiap kecamatan di suatu kabupaten tertentu, di bawah peta tersebut—masih di tautan yang sama, ditampilkan visualisasi interaktif berupa *side-by-side bar chart*. Dalam satu *bar chart* tersebut, akan ditampilkan data Rasio Murid terhadap Guru atau data Rasio Murid terhadap Sekolah (pengguna dapat memilih data yang ingin ditampilkan) per tingkatan pendidikan dan kecamatan dalam suatu kabupaten (yang juga ditentukan pengguna). Angka yang ditampilkan pada *bar chart* berupa angka maksimal dan minimal saja, menyesuaikan tujuan penelitian yang *urgent*-nya di sana, tetapi informasi yang tidak ditampilkan bisa muncul jika *bar-chart* terkait (yang ingin diketahui informasinya) tersentuh kursor. *Highlight* untuk data ini berupa tingkatan pendidikan, karena akan lebih berarti membandingkan data rasio dengan jenis yang sama untuk semua kecamatan di suatu kabupaten daripada melihat rasio data di semua tingkatan pendidikan untuk satu kecamatan saja. Sama seperti peta, pewarnaan pada *bar chart* tersebut menggunakan palet warna yang ramah bagi penderita buta warna, tetapi dibuat dengan pilihan warna yang cukup kontras dari peta karena warna tidak mewakili data yang sama. Gambar 5, 6, 7, dan 8 berikut memperjelas uraian tersebut.

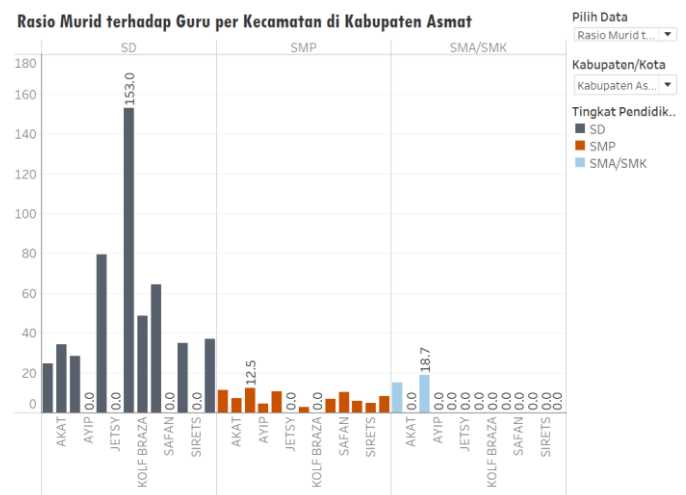
Pilih Data

Rasio Murid t...
Rasio Murid terhadap Guru
Rasio Murid terhadap Sekolah

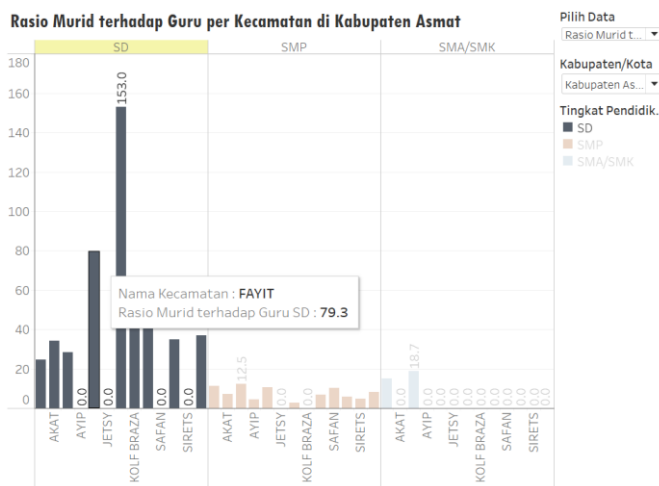
Gambar 5. Fitur berupa *drop down list* untuk memilih data yang ingin ditampilkan (antara Rasio Murid terhadap Guru atau Rasio Murid terhadap Sekolah)



Gambar 5. Fitur berupa *drop down list* untuk memilih data kabupaten yang ingin ditampilkan



Gambar 7. Paket visualisasi *side-by-side bar chart* secara keseluruhan.



Gambar 7. Highlight suatu bar tertentu

Tabel II dan Tabel III menunjukkan nilai rata-rata setiap variabel di tiap-tiap kluster. Kluster 1 (*Cluster 1 Tableau, Cluster 2 R Studio*) merupakan wilayah dengan Rasio Murid terhadap Guru dan Rasio Murid terhadap Sekolah di tiap tingkatan tinggi, sedangkan Kluster 2 (*Cluster 2 Tableau, Cluster 1 R Studio*) merupakan wilayah dengan Rasio Murid terhadap Guru dan Rasio Murid terhadap Sekolah di tiap tingkatan rendah.

Nilai rasio murid terhadap guru tinggi artinya satu guru menangani jumlah murid yang banyak. Makin tinggi nilai rasio ini berarti makin berkurang tingkat pengawasan dan perhatian guru terhadap murid, sehingga mutu pengajaran cenderung makin rendah (Sirusa BPS, Susenas 2011). Pasal 17 Peraturan Pemerintah Nomor 74 Tahun 2008 tentang Guru, menyebutkan bahwa pada jenjang SD, SMP, dan SMA idealnya satu guru bertanggung jawab terhadap 20 murid, sedangkan pada jenjang SMK idealnya satu guru bertanggung jawab pada 15 murid. Dari Tabel II dan Tabel III terlihat bahwa rata-rata rasio murid terhadap guru di tingkat SD dan SMP pada Kluster 1 melebihi nilai ideal 20, atau juga dapat dikatakan timpang.

Di sisi lain, nilai rasio murid terhadap sekolah pun demikian. Para peneliti telah mencapai konsensus umum bahwasanya ukuran kelas yang lebih kecil dapat menghasilkan keterampilan sosial, emosional, dan akademis yang positif. Hal ini juga ditemukan dalam berbagai penelitian, termasuk penelitian ukuran kelas yang dilakukan oleh Glass dan Smith pada tahun 1978 dan *Project STAR Tennessee* yang dilakukan pada pertengahan 1980-an. Kedua penelitian tersebut menyimpulkan bahwa ukuran kelas yang lebih kecil menghasilkan nilai ujian yang lebih tinggi, menghasilkan lebih sedikit murid yang putus sekolah, dan menyamakan kedudukan bagi minoritas dan anak-anak yang hidup dalam kemiskinan. Dari Tabel II dan Tabel III terlihat bahwa rata-rata rasio murid terhadap sekolah di tingkat SD, SMP, dan SMA/SMK bernilai besar pada Kluster 1.

TABEL II

Nilai Rata-Rata Hasil Pengelompokan Kluster dengan *Tableau*

Clusters	Number of Items	x_1	x_2	x_3	x_4	x_5	x_6
Cluster 1	12	41.193	21.268	15.665	224.2	220.72	252.54

Cluster 2	17	23.676	13.054	12.957	120.98	148.44	205.41
Not Clustered	0						

TABEL III

Nilai Rata-Rata Hasil Pengelompokan Kluster dengan *R Studio*

Clusters	Number of Items	x_1	x_2	x_3	x_4	x_5	x_6
Cluster 1 ^a	11	39.623	21.355	15.668	233.22	230.52	258.44
Cluster 2	18	25.609	13.457	13.106	121.20	146.47	204.42

^aKeluaran asli dalam *R Studi* berupa Kluster 2, begitupun sebaliknya. Di tabel ini akan dinamai menurut *output Tableau*.

VI. KESIMPULAN DAN SARAN

Dari penjabaran hasil analisis, dapat disimpulkan bahwa wilayah dalam Kluster 1 merupakan wilayah dengan ketimpangan rasio murid terhadap guru serta rasio murid terhadap sekolah. Kabupaten di Provinsi Papua yang termasuk dalam kluster ini adalah Kota Jayapura, Kab. Pegunungan Bintang, Kab. Yahukimo, Kab. Jayawijaya, Kab. Lanny Jaya, Kab. Tolikara, Kab. Puncak Jaya, Kab. Mimika, Kab. Deiyai, Kab. Paniai, Kab. Dogiyai, dan Kab. Mappi. Visualisasi dalam bentuk peta hasil kluster serta rincian data untuk setiap kecamatannya, dapat dilihat dalam tautan https://public.tableau.com/views/UjianAkhirSemesterUASVisualisasiDataDanInformasi/Dashboard1?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link.

Saran peneliti untuk penelitian selanjutnya adalah mengembangkan visualisasi yang otomatis memperbarui ketika terdapat DAPODIK baru disinkronkan dalam web Kemendikbud. Selain itu, dengan adanya data geospasial yang memuat lengkap informasi kecamatan hingga desa di Provinsi Papua, dapat dilakukan analisis *clustering* lingkup kecamatan, bahkan hingga lingkup desa. Untuk mewujudkan ini, perlu bantuan pemerintah Indonesia untuk melonggarkan akses terhadap data geospasial yang lengkap dan terbaru. Terkait hasil kluster, peneliti menyarankan pemerintah Indonesia untuk lebih memperhatikan mutu pendidikan wilayah dalam Kluster 1. Ketimpangan rasio murid terhadap guru serta murid terhadap sekolah dapat diatasi guru dan sekolah dengan menggunakan teknologi untuk mengisi kesenjangan, meningkatkan komunikasi lintas kurikulum dengan ujian berstandar, serta menawarkan insentif bagi siswa. Solusi tersebut tentu sulit diwujudkan jika tidak ada bantuan dari pemerintah. Selain itu, pemerintah juga dapat membuka lowongan untuk menjadi guru di Provinsi Papua atau juga dengan membangun/ merenovasi sekolah yang ada di sana.

DAFTAR PUSTAKA

- [1] Acer for Education (2021, 12) Berapa Rasio Guru-Murid dan Jumlah Siswa dalam Kelas yang Ideal? Available: <https://acerforeducation.id/blog-kami/edukasi/rasio-guru-dan-murid-yang-ideal/>

- [2] Clusteranalysis4marketing (2022) Why does cluster analysis sometimes gives different results? Available: <https://www.clusteranalysis4marketing.com/technical-aspects-cluster-analysis/why-does-cluster-analysis-sometimes-give-different-results/>.
- [3] Databoks. (2021, 11) Belum Ideal, Rasio Murid-Guru di SMK Timbang [Online]. Available: <https://databoks.katadata.co.id/datapublish/2021/11/29/belum-ideal-rasio-murid-guru-di-smk-timpang>
- [4] DPMPTSP (2019) Sosial. Available: <https://pmperizinan.jogjakota.go.id/web/konten/72/sosial>.
- [5] Elfarabi, M., F., 2018. “Analisis Faktor-Faktor yang Mempengaruhi Angka Partisipasi Sekolah di Indonesia”. Skripsi Elektronik, Yogyakarta: Universitas Islam Indonesia, pp. 47. Diakses pada 18 Juni 2022, dari <https://dspace.uui.ac.id/bitstream/handle/123456789/6233/SKRIPSI.pdf?sequence=1>.
- [6] Kurniawan T. A., 2012. “Penerapan Model Pembelajaran Kooperatif Tipe CIRC (*Cooperative Reading and Composition*) untuk Meningkatkan Hasil Belajar Bahasa Indonesia Siswa Kelas IV SD Negeri Pajang III Surakarta Tahun 2011/ 2012”. Skripsi Elektronik. Surakarta: Universitas Muhammadiyah. Diakses pada 20 Juni 2022, dari <http://eprints.ums.ac.id/17488/>.
- [7] Lestari, N., A., Analisis Faktor-Faktor yang Mempengaruhi Angka Partisipasi Sekolah serta Angka Putus Sekolah Tingkat Sekolah Dasar dan Sekolah Menengah Pertama: Data Panel 33 Provinsi di Indonesia Tahun 2006 hingga 2011. Tesis, Yogyakarta: Universitas Gadjah Madha, abstrak. Diakses pada 18 Juni 2022, dari <http://etd.repository.ugm.ac.id/penelitian/detail/76441>.
- [8] Rahmatin, U., Z., dan Prof. Dr. H. Ady Soejoto, “Pengaruh Tingkat Kemiskinan dan Jumlah Sekolah terhadap Angka Partisipasi Sekolah di Kota Surabaya,” *Jurnal Pendidikan Ekonomi Manajemen dan Keuangan*, vol. 1, no. 2, pp. 127-140, November 2017, <https://journal.unesa.ac.id/index.php/jpeka/article/download/1950/1335>.
- [9] Rondong, F., Faktor-Faktor yang Mempengaruhi Partisipasi Sekolah pada Jenjang Pendidikan Dasar: SD/MI dan SMP/MTs di Indonesia. Tesis, Jakarta: Universitas Indonesia, abstrak. Diakses pada 18 Juni 2022, dari <https://lib.ui.ac.id/file?file=pdf/abstrak-133149.pdf>.
- [10] Sirusa, Badan Pusat Statistik (BPS).
- [11] Wang, H., et. AL., “Hierarchical visualization of geographical areal data with spatial attribute association,” *Visual Informatics*, vol. 5, no. 3, pp. 82–91, September 2021, <https://doi.org/10.1016/j.visinf.2021.09.001>.
- [12] Y. Yang, T. Dwyer, K. Marriott, B. Jenny and S. Goodwin, "Tilt Map: Interactive Transitions Between Choropleth Map, Prism Map and Bar Chart in Immersive Environments," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 12, pp. 4507-4519, 1 Dec. 2021, doi: 10.1109/TVCG.2020.3004137.