

**SEGMENTASI WILAYAH BERDASARKAN KETERSEDIAAN FASILITAS PUBLIK
(STUDI KASUS: KABUPATEN/KOTA DI PROVINSI NUSA TENGGARA TIMUR)**

Oleh :

Anne Mudya Yolanda¹⁾, Agung Satrio Wicaksono²⁾

annemudyayolanda@lecturer.unri.ac.id

Prodi Statistika, Universitas Riau¹⁾

Prodi Administrasi Publik, Universitas Sultan Ageng Tirtayasa²⁾

ABSTRAK

Ketersediaan fasilitas publik dapat menjadi salah satu indikator dalam pembangunan daerah. Penelitian ini dilakukan untuk mendeskripsikan segmentasi ketersediaan fasilitas publik di Provinsi Nusa Tenggara Timur menurut kabupaten/kota. Analisis K-Means Clustering diterapkan pada data untuk mengelompokkan kabupaten/kota di Provinsi Nusa Tenggara Timur menjadi tiga kelompok. Variabel yang digunakan dalam analisis K-Means Clustering ini berdasarkan Statistik Potensi Desa di provinsi Nusa Tenggara Timur yaitu Ruang Publik, Pos Bindu, Sumber Penerangan Jalan Utama, Ketersediaan Angkutan Umum, dan Keberadaan Sarana Komunikasi. Berdasarkan hasil dan pembahasan, tiga segmen ketersediaan fasilitas publik menurut kabupaten atau kota di Provinsi Nusa Tenggara Timur masing-masing terdiri dari 8, 5, dan 9 kabupaten/kota untuk kelompok pertama, kedua, dan ketiga. Persentase keragaman yang diperoleh sebesar 61,4%. Segmen dengan ketersediaan fasilitas publik terendah adalah Kabupaten Sumba Barat, Kabupaten Belu, Kabupaten Sumba Tengah, Kabupaten Sabu Raijua, dan Kota Kupang. Segmen menengah terdiri atas Kabupaten Sumba Timur, Kabupaten Alor, Kabupaten Sikka, Kabupaten Ngada, Kabupaten Rote Ndao, Kabupaten Sumba Barat Daya, Kabupaten Nagekeo, dan Kabupaten Malaka. Daerah dengan ketersediaan fasilitas tertinggi adalah Kabupaten Kupang, Kabupaten Timor Tengah Selatan, Kabupaten Timor Tengah Utara, Kabupaten Lembata, Kabupaten Flores Timur, Kabupaten Ende, Kabupaten Manggarai, Kabupaten Manggarai Barat, dan Manggarai Timur.

Kata kunci : fasilitas publik, k-means, clustering

ABSTRACT

The availability of public facilities could be an indicator of regional development. The purpose of this study was to describe the segmentation of the availability of public facilities in Nusa Tenggara Timur province by regency/city. The K-Means Clustering analysis was used to divide the regencies/cities in Nusa Tenggara Timur province into three groups. The variables used in this K-Means Clustering analysis based on Nusa Tenggara Timur Village Potential Statistics are Public Spaces, Integrated Health Counseling Post, The Source of Main Street Illumination, Availability of Public Transportation, and Availability of Communication Facility. According to the findings and discussions, the first, second, and third clusters of the availability of public facilities by regencies/cities in Nusa Tenggara Timur Province consist of 8, 5, and 9 regencies/cities, respectively. Percentage of variability for this data is about 61.4 %. The segments with the lowest availability of public facilities are Sumba Barat regency, Belu regency, Sumba Tengah regency, Sabu Raijua regency, and Kupang City. The medium segment consists of Sumba Timur regency, Alor regency, Sikka regency, Ngada regency, Rote Ndao regency, Sumba Barat Daya regency, Nagekeo regency, and Malacca regency. The areas with the highest availability of facilities are Kupang regency, Timor Tengah Selatan regency, Timor Tengah Utara Regency, Lembata regency, Flores Timur regency, Ende regency, Manggarai regency, Manggarai Barat regency, and Manggarai Timur regency.

Key Words : publik facilities, k-means, clustering, segmentation

PENDAHULUAN

Pendahuluan Kemiskinan merupakan pekerjaan rumah yang masih harus diselesaikan di provinsi Nusa Tenggara Timur (NTT). Statistik menunjukkan pada September 2020 Persentase penduduk miskin meningkat 0.31 persen poin terhadap Maret 2020 dan meningkat 0.59 persen poin terhadap September 2019, yaitu mencapai 21.21 persen atau sebesar 1173.53 ribu orang. Angka ini meningkat 19.77 ribu orang terhadap Maret 2020 dan meningkat 44.07 ribu orang terhadap September 2019 (Badan Pusat Statistik Provinsi Nusa Tenggara Timur, 2021a). Angka ini jauh berada di atas dari Persentase penduduk miskin secara nasional pada September 2020 sebesar 10.19 persen (Badan Pusat Statistik,

2021b).

Ada banyak faktor yang mempengaruhi kemiskinan yang terjadi di berbagai provinsi di Indonesia. Salah satunya dapat dilihat dari Indeks Pembangunan Manusia atau IPM (Yolanda, Yunitaningtyas and Indahwati, 2019; Fahrika, Salam and Buhasyim, 2020). IPM sendiri tersusun atas tiga dimensi dasar, yaitu Umur panjang dan hidup sehat, Pengetahuan, dan Standar hidup layak (Badan Pusat Statistik, 2021a).

Pada era otonomi daerah seperti sekarang ini IPM dapat menjadi proksi yang ideal untuk mengukur kinerja pemerintah daerah pada tiga bidang utama yaitu pendidikan, kesehatan masyarakat, dan ekonomi. Peningkatan IPM sekaligus mencerminkan

peningkatan kinerja pelayanan publik, mengingat yang pada akhirnya dapat meningkatkan kesejahteraan masyarakat (Niswaty, Mano and Akib, 2015). Nusa Tenggara Timur sebagai daerah yang masih berjuang memberantas kemiskinan dan meningkatkan IPM, tentunya harus mau dan mampu meningkatkan kinerja pelayanan publik. Hal ini beriringan pula pada ketersediaan fasilitas publik di provinsi Nusa Tenggara Timur untuk memudahkan masyarakat dalam beraktivitas dan mengakses kebutuhan yang dapat meningkatkan IPM. Tersedianya fasilitas publik yang memadai menjadi salah satu indikator berjalannya pembangunan di suatu daerah.

Ketersediaan fasilitas publik dapat menjadi faktor penunjang dalam peningkatan kesejahteraan masyarakat. Kesejahteraan masyarakat memiliki beberapa indikator, yaitu Kependudukan, Kesehatan dan Gizi, Pendidikan, Ketenagakerjaan, Taraf dan Pola Konsumsi, Perumahan dan Lingkungan, Kemiskinan, serta Sosial Lainnya yang menjadi acuan dalam upaya peningkatan kualitas hidup (Badan Pusat Statistik, 2020). Adanya kebijakan dan pembangunan fasilitas publik sebagai sarana dan prasarana pendukung indikator tersebut merupakan bagian dari usaha meningkatkan potensi suatu daerah.

Pada tingkat daerah, ketersediaan fasilitas publik merupakan bagian dari pengembangan potensi desa. Bahkan, potensi desa menjadi salah satu kegiatan wajib yang dilakukan sebelum Pemerintah melakukan sensus. Pendataan potensi desa terakhir dilakukan tahun 2018 dengan tujuan sebagai berikut: 1) menyediakan data

yang mendukung perencanaan kegiatan sensus penduduk 2020, (2) sebagai sarana untuk pemutakhiran Master File Desa (MFD), (3) menyediakan data tentang keberadaan dan perkembangan potensi yang dimiliki desa/kelurahan yang meliputi: sosial, ekonomi, sarana, dan prasarana wilayah, (4) menyediakan data bagi keperluan pemutakhiran klasifikasi/tipologi desa, misalnya perkotaan-perdesaan, pesisir non pesisir, dan sebagainya, (5) sebagai sumber pemutakhiran peta wilayah kerja statistik, (6) menyediakan data pokok bagi penyusunan statistik wilayah kecil (7) menyediakan data bagi penyusunan berbagai analisis seperti identifikasi dan penentuan desa tertinggal, variabel konteks dalam PMT, identifikasi desa rawan bencana, dan identifikasi desa yang mempunyai kesulitan geografis, (8) menyediakan data bagi penghitungan indikator-indikator pembangunan/kemajuan desa (Badan Pusat Statistik Provinsi Nusa Tenggara Timur, 2018).

Statistik Potensi Desa berisi penjelasan mengenai wilayah administrasi pemerintahan setingkat desa menurut ketersediaan infrastruktur dan potensi yang dimiliki oleh setiap wilayah. Atau dengan kata lain berkaitan dengan kondisi fasilitas publik terkini di suatu wilayah. Selain menjadi acuan dalam pembangunan, ketersediaan fasilitas publik menjadi salah satu daya tarik bagi penduduk luar wilayah untuk berkunjung, terutama untuk daerah-daerah dengan potensi wisata, termasuk provinsi Nusa Tenggara Timur.

Ada 1305 daya Tarik Wisata di provinsi Nusa Tenggara Timur pada tahun 2020 yang berhasil dicatat Dinas Pariwisata dan Ekonomi Kreatif Provinsi Nusa Tenggara Timur (Badan Pusat

Statistik Provinsi Nusa Tenggara Timur, 2020a). Keadaan ini menjadikan provinsi Nusa Tenggara Timur sebagai salah satu tujuan populer bagi wisatawan untuk berkunjung, baik wisatawan domestik, maupun mancanegara. Pada tahun 2019 (sebelum memasuki masa pandemi Covid-19), Tingkat Penghunian Kamar (TPK) tahun 2019 untuk Hotel bintang adalah 52.17 dan non bintang sebesar 23.52 (Badan Pusat Statistik Provinsi Nusa Tenggara Timur, 2021b), bahkan Jumlah Wisatawan Mancanegara yang Masuk ke Indonesia Melalui Pintu Masuk Atambua 2019 pernah mencapai 12.561 pada bulan Desember 2019 (Badan Pusat Statistik Provinsi Nusa Tenggara Timur, 2020b). Angka ini, begitu pula jumlah kunjungan wisatawan tentunya dapat terus ditingkatkan dengan mengeluarkan berbagai kebijakan, diantaranya menambah ketersediaan fasilitas publik, seperti sarana transportasi, sarana komunikasi, ruang publik, sumber penerangan, dan fasilitas kesehatan.

Berdasarkan uraian di atas peneliti tertarik untuk melihat ketersediaan fasilitas publik yang ada di provinsi Nusa Tenggara Timur. Kajian ini akan diperluas dengan membagi kabupaten atau kota yang ada di provinsi Nusa Tenggara Timur menjadi beberapa kelompok berdasarkan ketersediaan fasilitas publik. Hasil pengelompokan dapat menjadi gambaran bagi pemangku kepentingan sekaligus menjadi acuan dalam merumuskan kebijakan agar lebih tersegmentasi.

Pengelompokan dilakukan dengan bantuan analisis *K-Means Clustering*. Secara umum, Analisis *Clustering* adalah teknik yang digunakan untuk mengelompokkan objek yang serupa. Data yang akan dikelompokkan tidak

harus memiliki label atau kategori tertentu. Algoritma *Clustering* akan membaca struktur data mendeskripsikan objek yang menarik dan menentukan cara terbaik untuk mengelompokkan objek (EMC Education Services, 2015).

Pada *K-Means Clustering*, data sebanyak n yang terdiri atas beberapa variabel akan dibagi menjadi beberapa kelompok (*cluster*). Setiap kelompok atau *cluster* tersebut memiliki beberapa anggota atau kumpulan objek serupa. Pengelompokan menggunakan "mean vektor" sebagai *centroid* (Templ, 2016). Algoritma tersebut akan diterapkan pada data untuk mencapai tujuan dari penelitian ini yaitu membuat segmentasi pada data terkait fasilitas publik di provinsi Nusa Tenggara Timur dengan menerapkan analisis *K-Means Clustering*.

TINJAUAN PUSTAKA

Analisis *clustering* adalah proses pengelompokan sekumpulan obyek yang memiliki sifat yang mirip (paling dekat kesamaannya) sehingga mengelompok ke dalam satu *cluster* (kelompok) yang sama (Han and Kamber, 2001). Analisis *clustering* disebut juga segmentasi data. Pada analisis ini dilakukan pembagian data yang besar ke dalam beberapa kelompok berdasarkan kesamaan yang dimiliki objek dalam kumpulan data (Han, Kamber and Pei, 2012).

Cara pengklasifikasian data dapat dibagi menjadi dua, yaitu teknik *Unsupervised* dan teknik *Supervised*. Pada teknik pengelompokan *unsupervised*, amatan pada data yang digunakan tidak memiliki label atau tidak ada informasi mengenai kelompok atau *cluster*. Teknik klasifikasi ini dilakukan

dengan tujuan untuk membagi amatan data yang diolah menjadi beberapa kelompok. Analisis ini sering juga dikenal sebagai analisis gerombol (*clustering analysis*).

Pada teknik klasifikasi *supervised*, amatan telah mempunyai label atau memiliki informasi mengenai kelompok atau grup. Teknik diterapkan untuk menentukan aturan pembeda antar kelompok amatan. Aturan pembeda tersebut dapat dimanfaatkan untuk menentukan keanggotaan dari amatan lain yang saat ini belum ada dalam data, dengan kata lain dapat digunakan untuk pendugaan data atau amatan baru.

Penentuan banyak *cluster* atau kelompok salah satunya dapat ditentukan dengan metode *elbow*. Metode *Elbow* dipakai untuk mendapatkan informasi dalam mencari banyaknya kelompok terbaik yang harus dibuat pada data. Metode ini menghitung persentase hasil perbandingan antara banyaknya kelompok (*cluster*) yang akan membentuk *elbow* (jika dibuat grafik akan membentuk pola seperti siku) pada suatu titik (Madhulata, 2012; Merliana, Ernawari and Santoso, 2015). Algoritma dari metode ini memberikan cara perhitungan untuk menentukan banyak *cluster* yang akan dibuat pada suatu kelompok data, diantaranya dengan melihat nilai *totwithins* atau total within. Nilai total within merupakan nilai jarak antar amatan yang berada dalam satu kluster yang sama.

Algoritma yang dijalankan menggunakan nilai Within *cluster sum of squares by cluster* (*wss*) atau jumlah kuadrat dalam *cluster* (kelompok) digunakan untuk melihat kedekatan antar data dan menjadikannya dalam

kelompok yang sama. Within *cluster sum of squares by cluster*, terkait dengan kesamaan anggota dalam satu kelompok. Nilai ini dapat menjadi salah satu indikator untuk melihat kebaikan dari banyaknya kelompok.

Persamaan 1:

$$wss = \sum_{i=1}^n \sum_{j=1}^p (x_{ij} - \bar{x}_{kj})^2 \quad (1)$$

dengan $i=1, 2, \dots, n, j=1, 2, \dots, p$ dan \bar{x}_j adalah rata-rata nilai variabel bebas ke- j dalam kelompok atau *cluster k* tertentu.

Jumlah Kuadrat Antar *Cluster* atau dikenal dengan *The between-cluster sum of square (bss)* merupakan jarak antar *cluster*. Semakin besar nilai ini maka semakin berbeda satu *cluster* dengan *cluster* lainnya.

Persamaan 2:

$$bss = \sum_{k=1}^c \sum_{j=1}^p (\bar{x}_{kj} - \bar{x}_j)^2 \quad (2)$$

dengan $i = 1, 2, \dots, n, j=1, 2, \dots, p, k = 1, 2, \dots, c$ (banyaknya *cluster* atau kelompok), \bar{x}_{kj} adalah rata-rata nilai variabel bebas ke- j dalam gerombol- k , \bar{x}_j adalah rata-rata variabel bebas ke- j .

Jarak yang digunakan dalam analisis adalah Jarak *Euclidian*. Jika amatan sebut saja $\{X_{ij}\}$, dengan $i = 1, 2, \dots, n, j=1, 2, \dots, p$, terdiri dari p variabel bebas. Andaikan $d_{ii'}$ merupakan jarak antara amatan ke i dan i' maka nilai yang paling banyak di terapkan yaitu jarak kuadrat *Euclidian*.

Persamaan 3: Jarak *Euclidian*

$$d_{ii'} = \sum_{j=1}^p (x_{ij} - x_{ij'})^2 \quad (3)$$

Penentuan *cluster* atau kelompok

menggunakan kriteria *elbow* merasiokan nilai RMSSTD dan RS statistics. RMSSTD (*Root Means Square Standard Deviation*) digunakan sebagai ukuran kemiripan (*homogeneity*) amatan yang terdapat di dalam *cluster* yang ditemukan (*within clusters*). Semakin rendah nilai *within cluster* yang dihasilkan, maka semakin mirip pula data atau amatan yang ada dalam suatu *cluster*. Adapun *R Squared* atau RS merupakan suatu angka yang menggambarkan ukuran tingkat kesamaan antara *cluster* atau kelompok (*between clusters*). Semakin rendah nilai RMSSTD artinya kemiripan data di dalam *cluster* yang ditemukan semakin tinggi. Sedangkan RS, memiliki nilai antara nol sampai dengan satu. Nilai 0 untuk kelompok yang sama dan nilai satu untuk kelompok yang benar-benar berbeda.

Saat menerapkan metode *elbow* dalam analisis, banyaknya *cluster* yang akan membentuk siku pada suatu titik merupakan banyaknya *cluster* atau kelompok terbaik. Hal ini dapat dilihat dengan membuat tampilan visual secara grafik. Pada grafik yang dihasilkan, akan terlihat penurunan hasil setiap nilai banyaknya *cluster* (*k*). apabila grafik menunjukkan penurunan yang drastis maka diasumsikan nilai *k* tersebut adalah banyaknya *cluster* terbaik, sampai setelahnya *k* akan menurun secara perlahan dan stabil. Pada sisi lain, apabila telah memutuskan banyaknya *cluster* atau kelompok yang diinginkan, maka analisis *clustering* yang dilakukan dapat langsung berlanjut pada menentukan isi dalam setiap *cluster* menggunakan algoritma yang dipilih.

Ada banyak algoritma yang diusulkan untuk dapat melakukan proses *clustering*. Pada penelitian ini, penentuan jumlah *cluster* terbaik

diperoleh dengan menggunakan *K-Means Clustering*. Analisis *K-Means Clustering* merupakan salah satu metode pengelompokan data (*clustering*) yang mengelompokkan data berdasarkan kemiripan karakteristik ke dalam sebuah *cluster* dan karakteristik lain dikelompokkan ke *cluster* yang berbeda. Algoritma yang dijalankan oleh metode ini berusaha untuk meminimalkan variansi atau ragam antar data yang ada di dalam suatu *cluster* dan memaksimalkan variansi dengan data yang ada di *cluster* lainnya (Hung *et al.*, 2005). Hasil pengelompokan yang baik terjadi ketika objek dalam kelompok atau *cluster* yang sama memiliki keragaman yang rendah sedangkan objek antar kelompok memiliki tingkat keragaman yang tinggi (Serban & Grigoretta 2006).

Metode *K-Means Clustering* ini termasuk dalam teknik pengelompokan dengan membuat penyekatan (*partition*) yang membagi atau memisahkan objek ke *k* daerah bagian yang terpisah (Johnson and Wichern, 2007). Setiap kelompok dalam metode ini diwakili oleh masing-masing centroid (Wu, 2012). Adapun Algoritma *K-Means Clustering* yang akan dijalankan dalam analisis adalah sebagai berikut.

- 1) Pilih nilai *k* sebagai banyaknya *cluster* atau kelompok yang ingin dibuat.
- 2) Bangkitkan secara acak *k* centroid sebagai titik pusat awal dari *cluster* atau kelompok secara acak.
- 3) Hitung jarak setiap amatan dalam data ke masing-masing centroid
- 4) Pengelompokan dilakukan pada setiap amatan berdasarkan jarak terdekat ke amatan dengan

centroid.

- 5) Cari pusat atau centroid baru dengan menghitung nilai rata-rata dari amatan yang ada pada centroid yang sama.
- 6) Ulangi langkah 3 apabila posisi centroid baru tidak sama dengan centroid lama. Lakukan tahapan ini sampai konvergen.

Analisis Clustering telah banyak dilakukan untuk mengelompokkan data, seperti penelitian dengan judul Pengelompokan Kabupaten/Kota di Provinsi Nusa Tenggara Timur Berdasarkan Indikator Indeks Pembangunan Manusia. Penelitian tersebut membagi Indeks Pembangunan Manusia di provinsi Nusa Tenggara Timur menjadi tiga kelompok berdasarkan K-Medoids. Berdasarkan hasil analisis diperoleh rekomendasi bahwa hasil segmentasi dapat menjadi acuan dalam mengevaluasi kinerja pemerintah untuk pemerataan pembangunan dengan kebijakan berdasarkan karakteristik wilayah sesuai hasil analisis clustering (Wicaksono and Yolanda, 2021).

Pada penelitian lain, analisis *K-Means Clustering* diterapkan pada Segmentasi Provinsi Berdasarkan Sarana dan Perlengkapan Fasilitas Kesehatan Keluarga Berencana (Faskes KB) Tahun 2021. Berdasarkan hasil analisis, sarana dan perlengkapan Faskes KB di Indonesia bergerombol menjadi empat kelompok dengan *within cluster sum of squares by cluster* untuk kelompok 1, 2, 3, dan 4 masing-masing adalah 9.376602, 1.029898, 3.866469, dan 16.833687. Segmentasi yang terbentuk yaitu kelompok pertama sebagai provinsi dengan sarana dan perlengkapan faskes

KB yang bisa dipakai tahun 2021 sangat memadai, kelompok tiga dikategorikan sebagai kelompok memadai, kelompok empat untuk kelompok cukup memadai, dan kelompok dua dalam kategori kurang memadai. (Yolanda and Yunitaningtyas, 2021).

METODE PENELITIAN

Sumber data pada penelitian ini adalah data yang berkaitan dengan ketersediaan fasilitas publik di provinsi Nusa Tenggara Timur. Data yang digunakan merupakan data sekunder dari Pendataan Potensi Desa (Podes) menurut Kabupaten atau kota di provinsi Nusa Tenggara Timur tahun 2018. Pendataan Podes merupakan kegiatan rutin yang dilakukan sejak tahun 1980 dan dikumpulkan sebanyak tiga kali dalam kurun waktu sepuluh tahun dan dilaksanakan dua tahun Sebelum pelaksanaan sensus sebagai bagian dari siklus sepuluh tahunan kegiatan sensus (Badan Pusat Statistik Provinsi Nusa Tenggara Timur, 2018). Pada tingkat provinsi, data potensi desa disajikan dalam bentuk agregasi data tingkat kabupaten/kota.

Data potensi desa dapat menjadi acuan dalam menggambarkan Kondisi suatu wilayah dan mengambil berbagai kebijakan. Statistik ini merupakan suatu nilai yang mewakili ketersediaan fasilitas publik pada desa atau kelurahan dalam bentuk agregasi. Adapun variabel dari Statistik Potensi Desa yang digunakan dalam analisis *K-Means Clustering* sebagai berikut.

1. Banyaknya Desa/Kelurahan Menurut Keberadaan Sarana Komunikasi Ruang Publik

Ruang Publik atau pada data penelitian ini adalah ruang publik terbuka adalah lahan umum yang utamanya diperuntukkan sebagai tempat berkumpul warga seperti untuk bersantai, bermain tanpa perlu membayar. Ruang publik terbuka dapat berupa lapangan terbuka atau alun-alun, taman, tempat bermain, dan sebagainya.

2. Banyaknya Desa/Kelurahan Menurut Kegiatan Pos Pembinaan Terpadu

Pos pembinaan terpadu atau Posbindu dalam penelitian mengikuti konsep yang saat ini dikenal sebagai Posbindu PTM (penyakit tidak menular). Posbindu merupakan peran serta masyarakat dalam melakukan kegiatan deteksi dini dan pemantauan faktor risiko PTM Utama yang dilaksanakan secara terpadu, rutin, dan periodik.

3. Banyaknya Desa/Kelurahan Menurut Sumber Penerangan Jalan Utama

Sumber Penerangan Jalan Utama yang digunakan dalam penelitian ini adalah jenis penerangan dan sumber pembiayaan penerangan yang ada di jalan utama desa. Data yang digunakan dalam proses segmentasi wilayah merupakan total dari penerangan dengan listrik diusahakan oleh pemerintah, listrik non-pemerintah, dan non-listrik.

4. Banyaknya Desa/Kelurahan Menurut Ketersediaan angkutan umum

Angkutan Umum yang dimaksud dalam penelitian ini adalah sarana angkutan pemindahan orang dan atau barang dari suatu tempat ke tempat lain dengan menggunakan kendaraan bermotor yang disediakan untuk dipergunakan untuk umum dengan dipungut bayaran. Data ketersediaan angkutan umum yang akan dijadikan karakteristik dalam proses segmentasi adalah total banyaknya desa atau kelurahan menurut angkutan umum yang tersedia, baik dengan trayek tetap maupun tidak tetap.

5. Banyaknya Desa/Kelurahan Menurut Keberadaan Sarana Komunikasi

Kegiatan Komunikasi yang dimaksud dalam tulisan ini adalah kegiatan usaha telekomunikasi, penyajian penerbitan, pos, dan giro. Data keberadaan sarana komunikasi dalam analisis *K-Means Clustering* adalah jumlah dari sarana komunikasi berupa warnet, Kantor Pos/Pos Pembantu/Rumah Pos, pos keliling, dan Perusahaan Jasa Ekspedisi Swasta yang tersedia di desa atau kelurahan menurut kabupaten atau kota di provinsi Nusa Tenggara Timur.

Analisis K-Means Clustering dapat dijalankan dengan bantuan berbagai *software* statistika. Pada penelitian ini, *K-Means Clustering* akan dikerjakan dengan bantuan program R Studi. Adapun sintak yang digunakan menggunakan algoritma Hartigan and Wong dari library

stats yang tersedia di program R (The R Core Team, 2021).

PEMBAHASAN

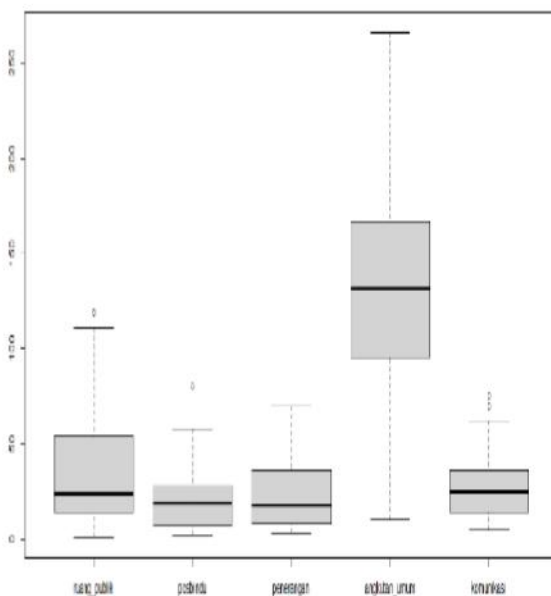
Statistik Deskriptif

Statistik potensi desa terdiri atas beberapa kategori, diantaranya ruang publik merupakan kategori hiburan dan olahraga, posbindu sebagai kategori pendidikan dan kesehatan, sumber penerangan jalan utama pada kategori perumahan dan lingkungan hidup, ketersediaan angkutan umum dan keberadaan sarana komunikasi sebagai kategori Angkutan, Komunikasi dan Informasi. Kategori ini dapat menjadi ukuran ketersediaan fasilitas publik di daerah. Satuan data dalam statistik potensi desa adalah banyaknya desa atau kelurahan.

Sumber: Statistik Potensi Desa provinsi Nusa Tenggara Timur 2018 (diolah)

Gambar 1. Box plot Ketersediaan Fasilitas provinsi Nusa Tenggara Timur tahun 2018

Berdasarkan data pada Gambar 1 diketahui bahwa terdapat pencilan pada data dilihat dari adanya titik-titik amatan yang berada di luar diagram kotak garis. Variabel dengan pencilan yaitu variabel ruang publik, posbindu, dan keberadaan sarana komunikasi. Eksplorasi ini menunjukkan bahwa ada wilayah kabupaten atau kota di provinsi Nusa Tenggara Timur yang memiliki jumlah ruang publik, posbindu, dan keberadaan sarana komunikasi cukup jauh berbeda dibandingkan dengan kabupaten atau kota lainnya. Keadaan ini turut mengindikasikan adanya daerah dengan ketersediaan fasilitas publik belum memadai.



Tabel 1. Ringkasan Ketersediaan Fasilitas Umum di provinsi Nusa Tenggara Timur tahun 2018

Statistik	A	B	C	D	E
Min	1.00	2.00	3.00	10.00	5.00
Q 1	14.25	7.50	8.50	97.25	14.00
Median	23.50	19.00	18.00	131.50	25.00
Mean	39.36	22.20	24.80	131.50	25.00
Q 3	54.00	27.00	35.50	164.00	35.70

Statistik	A	B	C	D	E
Maks	119.0	80.0	70.0	266.0	75.0
	0	0	0	0	0

Sumber: Statistik Potensi Desa provinsi Nusa Tenggara Timur 2018 (diolah)

Keterangan:

A = Ruang Publik

B = Pos Pembinaan Terpadu

C = Sumber Penerangan Jalan Utama

D = Ketersediaan angkutan umum

E = Keberadaan Sarana Komunikasi

Jika merujuk pada nilai rata-rata (mean) yang disajikan pada Tabel 1 tampak bahwa rata-rata ruang publik adalah $39.36 \approx 40$ desa. Wilayah dengan ruang publik paling sedikit adalah kabupaten Sikka (hanya satu desa). Namun, ada pula kabupaten yang punya sangat banyak mencapai 119 kelurahan atau desa dengan ruang publik, yakni kabupaten Lembata. Jika diurutkan dari terkecil ke terbesar, sebanyak 25 % data berada pada rentang 1 – 14.25 (Kuartil 1 atau Q 1). Artinya terdapat cukup banyak kabupaten/kota yang desa atau kelurahannya baru memiliki sedikit ruang publik. 75% data berada pada batas 54 (kuartil 3 atau Q 3). Nilai ini masih berada cukup jauh dari nilai maksimal (lebih dari dua kali lipat nilai kuartil 3), sehingga dapat dikatakan cukup sedikit kabupaten atau kota di provinsi Nusa Tenggara Timur yang memiliki ruang publik di atas rata-rata.

Pada ketersediaan posbindu di provinsi Nusa Tenggara Timur dapat dilihat dari kabupaten Sumba Barata memiliki 2 kelurahan atau desa dengan posbindu. Wilayah dengan posbindu terbanyak dimiliki oleh kabupaten Flores Timur terdapat 80 kelurahan atau desa.

Kuartil tiga dari data sebesar 27, juga berada sangat jauh dari nilai maksimal. Berdasarkan Kondisi ini dapat dinyatakan bahwa sebagian besar kabupaten atau kota hanya memiliki sedikit posbindu.

Nilai tengah (median) dari variabel ruang publik dan posbindu berada di bawah nilai rata-rata. Hal ini menunjukkan ada cukup banyak kabupaten atau kota yang ketersediaan fasilitasnya berada di bawah ketersediaan rata-rata. Jangkauan (*range*) dari kedua variabel juga cukup lebar.

Pada variabel keberadaan sarana komunikasi diperoleh rata-rata sebesar 25 desa. Kabupaten Sabu Raijua merupakan wilayah dengan total sarana komunikasi paling sedikit (5 kelurahan atau desa). Kota Kupang sebagai ibu kota provinsi adalah wilayah dengan total sarana komunikasi terbanyak (75 kelurahan atau desa). Mean dan median pada variabel ini sama, sehingga dapat dikatakan bahwa keberadaan sarana komunikasi sebagian besar totalnya berada di sekitar nilai rata-rata dan median atau sebagian besar wilayah memiliki cukup banyak sarana komunikasi.

Kuartil 1 (Q 1) keberadaan sarana komunikasi sebesar 14, menunjukkan bahwa sekitar 25 % kabupaten atau kota memiliki banyak desa atau kelurahan dengan sarana komunikasi relatif kecil. Kuartil 3 yang mewakili posisi 75 % data memiliki batas nilai sebesar 35.75. Angka ini juga masih relatif kecil dan memperkuat dugaan ada daerah dengan ketersediaan fasilitas publik belum memadai.

Pada Sumber Penerangan Jalan Utama, wilayah dengan total

penerangan paling sedikit adalah kabupaten Sabu Raijua. Begitu pula untuk ketersediaan angkutan umum. Kelurahan atau desa di Kabupaten Flores Timur memiliki sumber penerangan paling banyak, sedangkan secara kelurahan atau desa di kabupaten Ende memiliki total ketersediaan angkutan umum paling banyak.

Pada variabel ini kuartil 1 dan 3 masing-masing didistribusikan sebanyak 7.5 dan 35.50. Angka ini juga masih dapat terus ditingkatkan agar lebih banyak lagi desa atau kelurahan yang memiliki sumber penerangan utama. Pola dari variabel ini serupa dengan variabel keberadaan sarana komunikasi.

Analisis K-Means Clustering

Sebelum melakukan analisis *K-Means Clustering*, langkah awal yang terlebih dahulu harus dilakukan adalah menentukan banyaknya *cluster* (*k*) yang akan digunakan peneliti untuk mengelompokkan data. Penentuan jumlah *cluster* yang optimum bisa menggunakan kriteria elbow. Namun, pada data dalam penelitian ini, peneliti tidak melakukan langkah tersebut, karena data akan langsung dikelompokkan menjadi tiga segmen.

Ketersediaan fasilitas publik menurut kabupaten atau kota di provinsi Nusa Tenggara Timur akan dikelompokkan dengan Analisis *K-Means Clustering* dan membentuk tiga segmentasi ketersediaan fasilitas publik dengan label rendah, sedang, dan tinggi. Luaran dari Analisis *K-Means Clustering* yang diterapkan disajikan pada Gambar 2.

Gambar 2. Hasil Analisis *K-Means Clustering* terhadap Ketersediaan Fasilitas Umum di provinsi Nusa Tenggara Timur tahun 2018

Semakin kecil jumlah kuadrat dalam kelompok menjelaskan bahwa semakin dekat jarak antara amatan dengan pusat kelompok atau centroid, maka semakin mirip karakteristik dalam kelompok. Artinya, Semakin kecil nilai yang dihasilkan maka anggota dalam *cluster* atau kelompok akan semakin homogen. Nilai *Within cluster sum of squares by cluster (withinss)* yang dihasilkan untuk adalah 9033.375, 7809.6, dan 35157.333. Hasil dari ketiga kelompok yang terbentuk menunjukkan bahwa paling tinggi ditunjukkan oleh kelompok ketiga sebesar 35157.333 dan terendah sebesar 7809.6 oleh kelompok dua.

Jarak antar *cluster* yang terbentuk sebesar 82629.1 (*betweenss*) yang mana semakin besar nilainya maka hasil pengelompokan semakin baik. Analisis *cluster* dengan metode K-means menuntut keragaman di dalam *cluster* sekecil-kecilnya, sedangkan antar *cluster* diharapkan sebesar-besarnya. Persentase keragaman yang dihasilkan sebesar 61.4 % ($between_ss = 82629.1 / total_ss = 134629.4$). Hal ini berarti antar *cluster* mampu menjelaskan keberagaman sebanyak 61.4 %. Hal ini menunjukkan, *cluster* atau kelompok

k-means clustering with 3 clusters of sizes 8, 5, 9

Cluster means:			
	Ruang_Publik	Pos_Pembinaan_Terpada_Posbindu	Sumber_Penerangan_Jalan_Utama
1	21.00000	22.00000	23.50000
2	14.60000	6.40000	19.40000
3	69.44444	31.33333	29.11111
	ketersediaan_angkutan_umum	Keberadaan_Sarana_Komunikasi	
1	120.7500	27.62500	
2	46.6000	29.80000	
3	168.6667	29.44444	

yang dihasilkan cukup baik.

Pada analisis ini data dikelompokkan dengan metode k-means dengan tujuan untuk membagi data menjadi tiga kelompok sehingga jumlah kuadrat dari titik ke pusat kelompok yang ditetapkan dapat diminimalkan. Terdapat Centres atau centroid yang mana dalam analisis *K-Means Clustering* dibentuk dari *cluster means*. Centroid merupakan pusat dari *cluster* atau kelompok (disajikan pada Tabel 2). Berdasarkan nilai centroid untuk masing-masing *cluster* dan variabel dihasilkan dari program di atas (Gambar 2) dapat ditentukan karakteristik untuk setiap segmen *cluster* yang terbentuk.

Tabel 2. *Cluster Means* dari analisis *K-Means Clustering*

<i>Cluster</i>	A	B	C	D	E
1	21.0	22.0	23.5	120.75	27.625
2	14.6	6.4	19.4	46.60	29.80
3	69.4	31.3	29.1	188.67	29.44

Keterangan:

A = Ruang publik

B = Pos Pembinaan Terpadu

C = Sumber Penerangan Jalan Utama

D = Ketersediaan angkutan umum

E = Keberadaan Sarana Komunikasi

Analisis *K-Means Clustering* yang dilakukan membagi amatan menjadi tiga kelompok. Kelompok 1, 2, dan 3 masing-masing secara berurutan berisi 8, 5, dan 9 wilayah kabupaten atau kota. Adapun wilayah yang tercakup beserta karakteristik yang terbentuk dari hasil kemiripan data pada algoritma

pengelompokan tiga segmen tersebut adalah sebagai berikut.

1. Kelompok 1

Kelompok 1 terdiri atas 8 kabupaten, yaitu kabupaten Sumba Timur, kabupaten Alor, kabupaten Sikka, kabupaten Ngada, kabupaten Rote Ndao, kabupaten Sumba Barat Daya, kabupaten Nagekeo, dan kabupaten Malaka. Kabupaten pada kelompok 1 memiliki ketersediaan fasilitas publik cukup sedang untuk variabel ruang publik, posbindu, Sumber Penerangan Jalan Utama, dan ketersediaan angkutan umum dengan keberadaan sarana komunikasi paling sedikit. Berdasarkan hasil tersebut, wilayah pada kelompok 1 disegmentasikan sebagai kelompok sedang.

2. Kelompok 2

Kelompok 2 terdiri atas 5 wilayah, yaitu kabupaten Sumba Barat, kabupaten Belu, kabupaten Sumba Tengah, kabupaten Sabu Raijua, dan Kota Kupang. Kedekatan antar wilayah pada kelompok ini terbentuk karena setiap wilayah memiliki kemiripan pola, yaitu hanya memiliki sedikit fasilitas publik dari sektor ruang publik, posbindu, Sumber Penerangan Jalan Utama, dan ketersediaan angkutan umum. Namun, memiliki desa atau kelurahan dengan keberadaan sarana komunikasi paling banyak. Oleh karenanya, daerah pada kelompok ini digolongkan sebagai segmen dengan ketersediaan fasilitas rendah.

3. Kelompok 3

Kelompok 3 terdiri atas 9 wilayah, yaitu kabupaten Kupang, kabupaten Timor Tengah Selatan, kabupaten Timor Tengah Utara, kabupaten Lembata, kabupaten Flores Timur, kabupaten Ende, kabupaten Manggarai, kabupaten Manggarai Barat, dan kabupaten Manggarai Timur. Variabel ruang publik, posbindu, Sumber Penerangan Jalan Utama, dan ketersediaan angkutan umum pada kelompok ini ketersediaannya paling banyak di antara dua kelompok yang lain. Mengacu pada hasil ini, kabupaten pada kelompok 3 adalah daerah dengan ketersediaan fasilitas publik paling tinggi.

Pada hasil *K-Means Clustering*, kelompok rendah yaitu kabupaten Sumba Barat, kabupaten Belu, kabupaten Sumba Tengah, kabupaten Sabu Raijua, dan Kota Kupang memiliki total fasilitas publik paling sedikit di antara dua kelompok lainnya. Namun, jika dilihat lebih jauh kelompok ini sudah memiliki keberadaan sarana komunikasi paling banyak. Kondisi ini menunjukkan bahwa pada daerah-daerah dalam segmen ini diasumsikan penambahan sarana komunikasi bukanlah prioritas. Penambahan fasilitas untuk meningkatkan kesejahteraan masyarakat dapat berfokus pada sektor ruang publik, posbindu, Sumber Penerangan Jalan Utama, dan ketersediaan angkutan umum di berbadai desa atau kelurahan. Hal ini tentunya harus juga diselaraskan dengan jumlah penduduk dan kondisi geografis setempat.

Pada segmen sedang terdapat kabupaten Sumba Timur, kabupaten Alor, kabupaten Sikka, kabupaten Ngada, kabupaten Rote Ndao, kabupaten Sumba

Barat Daya, kabupaten Nagekeo, dan kabupaten Malaka yang mana tercatat secara umum memiliki total keberadaan sarana komunikasi lebih sedikit dibandingkan dua kelompok lain. Kondisi ini menunjukkan, apabila pemangku kepentingan setempat ingin menambahkan fasilitas atau memiliki alokasi dana untuk memperbanyak sarana komunikasi maka kabupaten-kabupaten pada kelompok ini berada pada daftar teratas sebagai wilayah tujuan alokasi.

Pada kabupaten Kupang, kabupaten Timor Tengah Selatan, kabupaten Timor Tengah Utara, kabupaten Lembata, kabupaten Flores Timur, kabupaten Ende, kabupaten Manggarai, kabupaten Manggarai Barat, dan kabupaten Manggarai Timur memiliki banyak fasilitas dibandingkan daerah lainnya. Oleh karenanya, wilayah ini belum menjadi urutan prioritas mengacu pada hasil analisis *K-Means Clustering* yang dilakukan. Namun, hal ini tentunya perlu dicocokkan dengan kondisi terkini, seperti penambahan jumlah penduduk, kepadatan penduduk, aktivitas penduduk terutama kegiatan ekonomi, dan kondisi geografis setiap kabupaten atau kota, serta aktivitas masyarakat setempat. Jika daerah-daerah ini merupakan daerah tujuan wisata atau berada pada jalur utama lalu lintas warga, maka tentunya penambahan beberapa fasilitas penunjang tetap diperlukan, seperti penambahan Sumber Penerangan Jalan Utama dan ketersediaan angkutan umum.

Hasil analisis di atas dapat dijadikan acuan bagi pemangku kebijakan di provinsi Nusa Tenggara Timur untuk membuat kebijakan yang tersegmentasi. Pengelompokan yang dilakukan

memberikan informasi wilayah, dalam hal ini kabupaten atau kota sesuai kondisi ketersediaan fasilitas publik saat ini. Jika pemerintah ingin menambah fasilitas, maka alokasinya menjadi lebih terarah, sebab sudah diketahui kabupaten atau kota mana yang jumlah fasilitasnya masih butuh tambahan. Pada Akhirnya, pemangku kebijakan dapat menggunakan hasil klasifikasi dari metode *K-Means Clustering* agar arah pembangunan dapat lebih efektif.

Analisis *K-Means Clustering* dapat dilakukan lebih luas yaitu pemetaan sampai tingkat yang lebih kecil. Analisis pada tingkat wilayah desa (jika data tersedia) dapat menggambarkan ketersediaan fasilitas publik menurut desa atau kelurahan. Hasil penelitian yang diperoleh dapat menjadi semakin rinci, karena diketahui wilayah pada tingkat desa atau kelurahan yang sangat membutuhkan perhatian khusus dari Pemerintah.

SIMPULAN

Analisis *K-Means Clustering* yang dilakukan, dapat disimpulkan bahwa banyaknya desa atau kelurahan menurut keberadaan fasilitas publik di provinsi Nusa Tenggara Timur membagi kabupaten atau kota menjadi tiga segmen atau kelompok. Berdasarkan hasil dan pembahasan yang dilakukan dapat disimpulkan bahwa kelompok 1, 2, dan 3 masing-masing terdiri atas 8, 5, dan 9 kabupaten atau kota. Persentase keragaman yang dihasilkan dari hasil segmentasi sebesar 61.4 %.

Segmen dengan ketersediaan fasilitas publik paling rendah adalah kelompok 2 yang terdiri atas kabupaten Sumba Barat, kabupaten Belu, kabupaten

Sumba Tengah, kabupaten Sabu Raijua, dan Kota Kupang. Segmen dengan ketersediaan fasilitas publik kategori sedang yaitu kelompok 1 terdiri atas kabupaten Sumba Timur, kabupaten Alor, kabupaten Sikka, kabupaten Ngada, kabupaten Rote Ndao, kabupaten Sumba Barat Daya, kabupaten Nagekeo, dan kabupaten Malaka. Wilayah dengan banyaknya desa atau kelurahan dengan ketersediaan fasilitas publik paling tinggi adalah kelompok 3, yakni kabupaten Kupang, kabupaten Timor Tengah Selatan, kabupaten Timor Tengah Utara, kabupaten Lembata, kabupaten Flores Timur, kabupaten Ende, kabupaten Manggarai, kabupaten Manggarai Barat, dan kabupaten Manggarai Timur. Hasil segmentasi membantu pemangku kepentingan dan pembuat kebijakan dalam merumuskan program terkait ketersediaan fasilitas sesuai hasil segmentasi. Pemetaan dan alokasi seperti penambahan fasilitas dapat mengacu pada karakteristik tiap kelompok sehingga sesuai dengan kebutuhan setiap desa dan kelurahan di provinsi Nusa Tenggara Timur.

Analisis lanjutan dari penelitian ini dapat dilakukan dengan menambahkan variabel yang menjadi acuan dalam pengelompokan data, misalnya memasukkan semua variabel dari Statistik Potensi Desa. Lebih lanjut, hasil yang diperoleh akan lebih mendalam jika dapat dilakukan pengelompokan menurut desa atau kelurahan. Penerapan teknik lain juga dapat diperluas menggunakan algoritma *K-Medoids Clustering*. Analisis spasial dan analisis regresi juga dapat diterapkan untuk melihat pengaruh yang diberikan oleh variabel-variabel dalam Statistik Potensi Desa terhadap beberapa indikator penting dalam melihat

performa daerah baik tingkat kabupaten atau kota maupun provinsi, seperti Produk Domestik Regional Bruto (PDRB), Produk Domestik Regional Bruto (PDRB) pariwisata, angka kemiskinan, Pendapatan Asli Daerah (PAD), jumlah kunjungan wisatawan lokal dan mancanegara ke kabupaten atau kota, dan lainnya.

DAFTAR PUSTAKA

- Badan Pusat Statistik (2020) *Indikator Kesejahteraan Rakyat 2020*. Edited by Sub-direktorat Indikator Statistik. Jakarta: Badan Pusat Statistik Republik Indonesia.
- Badan Pusat Statistik (2021a) *Indeks Pembangunan Manusia 2020*. Jakarta. Available at: <https://www.bps.go.id/publication/2021/04/30/8e777ce2d7570ced44197a37/indeks-pembangunan-manusia-2020.html> (Accessed: 18 May 2021).
- Badan Pusat Statistik (2021b) *Profil Kemiskinan Indonesia, Berita Resmi Statistik*. Available at: https://www.bps.go.id/website/materi_ind/materiBrsInd-20210215114827.pdf.
- Badan Pusat Statistik Provinsi Nusa Tenggara Timur (2018) *Statistik Potensi Desa Provinsi Nusa Tenggara Timur 2018*. Edited by Bidang Statistik Sosial. Nusa Tenggara Timur. Available at: https://ntt.bps.go.id/backend/materi_ind/materiBrsInd-20210215113503.pdf.
- Badan Pusat Statistik Provinsi Nusa Tenggara Timur (2020a) *Jumlah Daya Tarik Wisata Menurut Kabupaten/Kota 2018-2020, Pariwisata*. Available at: <https://ntt.bps.go.id/indicator/16/1374/1/jumlah-daya-tarik-wisata-menurut-kabupaten-kota.html> (Accessed: 21 May 2021).
- Badan Pusat Statistik Provinsi Nusa Tenggara Timur (2020b) *Jumlah Wisatawan Mancanegara yang Masuk ke Indonesia Melalui Pintu Masuk Atambua 2019, Pariwisata*. Available at: <https://ntt.bps.go.id/indicator/16/1029/3/jumlah-wisatawan-mancanegara-yang-masuk-ke-indonesia-melalui-pintu-masuk-atambua.html> (Accessed: 21 May 2021).
- Badan Pusat Statistik Provinsi Nusa Tenggara Timur (2021a) *Profil Kemiskinan di NTT, Berita Resmi Statistik*. Available at: <https://ntt.bps.go.id/pressrelease/2021/02/15/1019/persentase-penduduk-miskin-september-2020-naik-menjadi-21-21-persen.html>.
- Badan Pusat Statistik Provinsi Nusa Tenggara Timur (2021b) *Tingkat Penghunian Kamar (TPK) Menurut Jenis Hotel dan Bulan 2019-2021, Pariwisata*. Available at: <https://ntt.bps.go.id/indicator/16/68/1/tingkat-penghunian-kamar-tpk-menurut-jenis-hotel-dan-bulan.html> (Accessed: 21 May 2021).
- EMC Education Services (2015) *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. Indianapolis, Indiana: John Wiley & Sons, Inc.
- Fahrika, A. I., Salam, H. and Buhasyim, M. A. (2020) 'Effect of Human Development Index (HDI), Unemployment, and Investment Realization toward Poverty in South

- Sulawesi- Indonesia', *The International Journal of Social Sciences World provide*, 2(02), pp. 110–116. doi: doi.org/10.5281/zenodo.4080749.
- Han, J. and Kamber, M. (2001) *Data Mining: Concepts and Techniques*. Second Edi. San Francisco: Morgan Kauffman Publishers.
- Han, J., Kamber, M. and Pei, J. (2012) *Data Mining: Concepts and Techniques*. Third Edit. USA: Morgan Kaufmann Publishers.
- Hung, C. . *et al.* (2005) 'An Efficient k-Means Clustering Algorithm Using Simple Partitioning', *Journal Of Information Science And Engineering*, XXI(1), pp. 1157–1177.
- Johnson, R. A. and Wichern, D. W. (2007) *Applied Multivariate Statistical Analysis*. 6th editio. New Jersey (US): Pearson Prentice Hall.
- Madhulata, T. S. (2012) 'An Overview On Clustering Methods', *IOSR Journal of Engineering*, 02(04), pp. 719–725. doi: <https://dx.doi.org/10.9790/3021-0204719725>.
- Merliana, N. P. E., Ernawari and Santoso, A. J. (2015) 'Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means Clustering', in *Prosing Seminar Nasional Multi Disiplin Ilmu & Call for Papers UNISBANK (SENDI_U)*. UNISBANK, pp. 978–979. Available at: <https://www.unisbank.ac.id/ojs/index.php/sendu/article/view/3333>.
- Niswaty, R., Mano, J. and Akib, H. (2015) 'An Analysis of The Public Service Performance based on Human Development Index in Makassar City , Indonesia', *IJABER*, 13(6), pp. 4395–4403. Available at: www.researchgate.net/publication/301680972_An_analysis_of_the_public_service_performance_based_on_human_development_index_in_makassar_city_Indonesia.
- Templ, M. (2016) *Simulation for Data Science with R*. Birmingham, United Kingdom: PACKT Publishing.
- The R Core Team (2021) 'R : A Language and Environment for Statistical Computing'. R Foundation for Statistical Computing, p. 1468. Available at: <https://cran.r-project.org/doc/manuals/r-devel/fullrefman.pdf>.
- Wicaksono, A. S. and Yolanda, A. M. (2021) 'Pengelompokan Kabupaten / Kota di Provinsi Nusa Tenggara Timur Berdasarkan Indikator Indeks Pembangunan Manusia Menggunakan K-Medoids Clustering', *JSTAR: Jurnal Statistika Terapan*, 1(1), pp. 79–90. doi: 10.5300/jstar.v1i1.7.
- Wu, J. (2012) *Advances in K-Means Clustering A Data Mining Thinking*. New York (US): Springer.
- Yolanda, A. M. and Yunitaningtyas, K. (2021) 'Segmentasi Provinsi Berdasarkan Sarana dan Perlengkapan Fasilitas Kesehatan Keluarga Berencana Tahun 2021', *Jurnal Keluarga Berencana*, 6(01), pp. 20–30. doi: <https://doi.org/10.37306/kkb.v6i1>.
- Yolanda, A. M., Yunitaningtyas, K. and Indahwati (2019) 'Spatial Data Panel Analysis for Poverty in East Java Province 2012- 2017', *Journal of Physics: Conf. Series*, 1265(012027). doi: 10.1088/1742-6596/1265/1/012027.