

BAB II

TINJAUAN PUSTAKA

Pada Bab II akan dibahas mengenai landasan-landasan teori yang mendukung penelitian ini, yaitu analisis kelompok, metode pengelompokan hierarki (*single linkage* dan *complete linkage*), indeks *Davies Bouldin*, dan indikator pendidikan.

A. Analisis Kelompok

Cluster analysis (analisis kelompok) adalah salah satu teknik statistik yang bertujuan untuk mengklasifikasikan objek-objek ke dalam kelompok-kelompok sedemikian sehingga kelompok-kelompok yang terbentuk memiliki homogenitas tinggi di antara anggotanya dan heterogenitas tinggi antarobjek dalam kelompok yang berbeda (Sharma, 1996:185). Analisis kelompok adalah salah satu teknik multivariat yang memiliki tujuan utama untuk mengklasifikasi objek-objek berdasarkan karakteristik yang dimiliki. Analisis kelompok mengklasifikasi objek-objek ke dalam kelompok-kelompok sehingga objek-objek yang berada dalam satu kelompok serupa sedang objek yang berada dalam kelompok yang lain berbeda. Analisis tersebut bertujuan untuk memaksimalkan homogenitas objek dalam kelompok sementara itu juga memaksimalkan heterogenitas antarkelompok (Hair, et. al, 2009:478).

Analisis kelompok termasuk ke dalam analisis statistik multivariat metode interdependen, sehingga tujuannya tidak menghubungkan atau membedakan sampel/variabel lain (Ary, 2015). Analisis kelompok berguna untuk meringkas data dengan jalan mengelompokkan objek-objek berdasarkan

kesamaan karakteristik tertentu diantara objek-objek yang akan diteliti. Analisis kelompok terbagi menjadi dua metode, yaitu metode hierarki dan metode nonhierarki. Analisis kelompok adalah suatu alat untuk mengelompokkan sejumlah n objek berdasarkan p variabel yang secara relatif mempunyai kesamaan karakteristik diantara objek-objek tertentu, sehingga keragaman dalam suatu kelompok tersebut lebih kecil dibandingkan keragaman antarkelompok. Objeknya dapat berupa barang, jasa, hewan, manusia (responden, konsumen, dan lain-lain). Objek-objek tersebut akan diklasifikasikan dalam satu atau lebih kelompok, sehingga objek-objek yang berada dalam satu kelompok akan mempunyai kemiripan karakter. Jika terdapat n objek dan p variabel, maka observasi x_{ij} dengan $i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, p$ dan dapat digambarkan seperti pada Tabel 1. Kelompok yang baik mempunyai ciri-ciri sebagai berikut:

- memiliki homogenitas yang tinggi antar anggota dalam satu kelompok (*within cluster*)
- memiliki heterogenitas yang tinggi antarkelompok yang satu dengan kelompok yang lain (*between cluster*).

Dari dua hal di atas dapat disimpulkan bahwa sebuah kelompok yang baik adalah kelompok yang mempunyai anggota-anggota yang semirip mungkin satu dengan yang lain, namun sangat tidak mirip dengan anggota-anggota kelompok yang lain (Santoso, 2010:113).

Tabel 1. Data Observasi x_{ij}

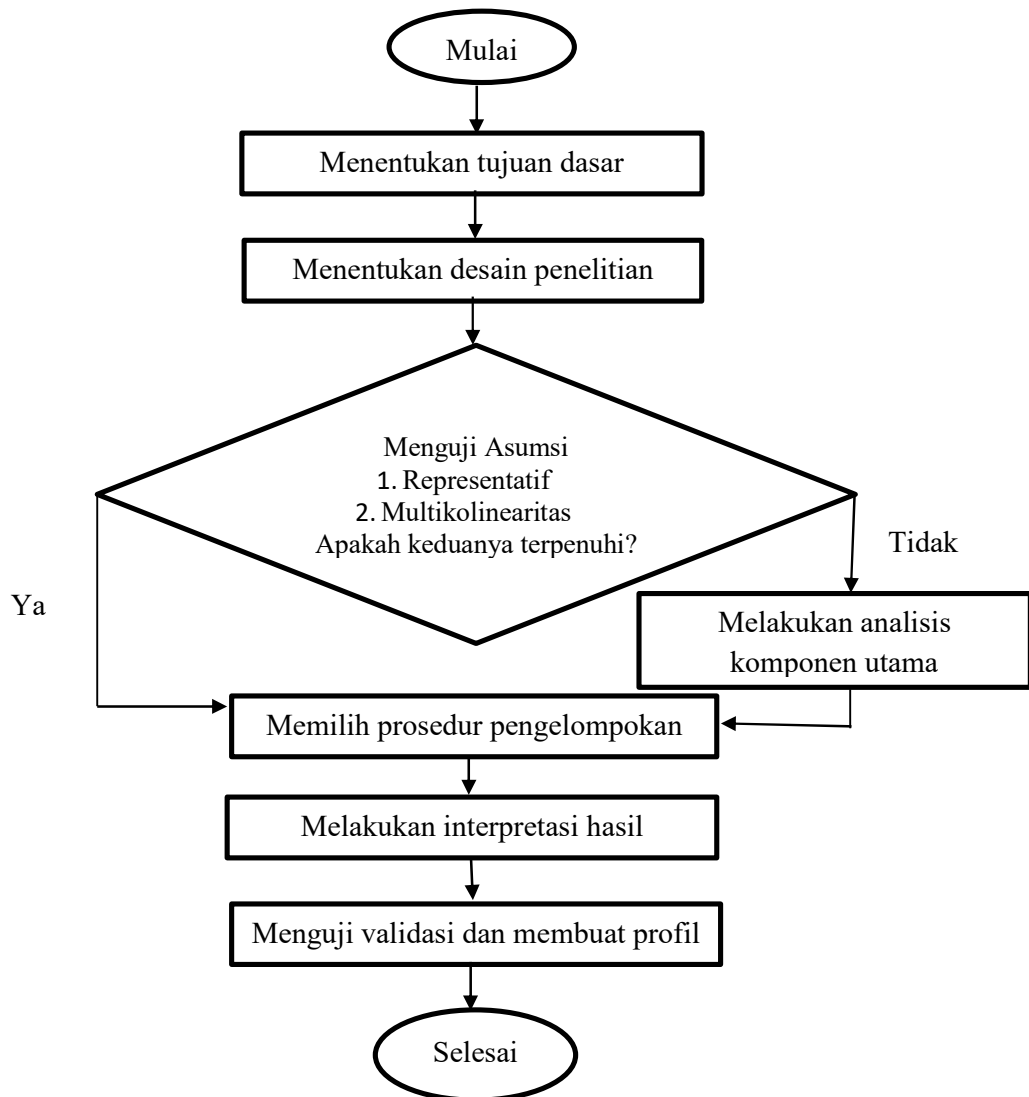
	Var 1	Var 2	...	Var j	...	Var p
Objek 1	x_{11}	x_{12}	...	x_{1j}	...	x_{1p}
Objek 2	x_{21}	x_{22}	...	x_{2j}	...	x_{2p}
.
.
Objek i	x_{i1}	x_{i2}	.	x_{ij}	.	x_{ip}
.
.
Objek n	x_{n1}	x_{n2}	.	x_{nj}	.	x_{np}

Dalam proses analisis kelompok terdapat 6 tahapan yang harus dilakukan, yaitu menentukan tujuan analisis kelompok, menentukan desain penelitian analisis kelompok, menentukan asumsi analisis kelompok, memilih prosedur pengelompokan, menginterpretasi hasil analisis kelompok, dan mengukur tingkat validasi hasil analisis kelompok (Simamora, 2011:206). Tahapan-tahapan dalam proses analisis kelompok dapat direpresentasikan seperti pada Gambar 1.

1. Tujuan Dasar

Tujuan utama analisis kelompok adalah untuk menempatkan sekumpulan objek ke dalam dua atau lebih kelompok berdasarkan kesamaan-kesamaan objek atas dasar berbagai karakteristik. Terdapat tiga tujuan analisis

kelompok yang tersedia bagi peneliti (Hair, et.al, 2009:489), yaitu sebagai berikut.



Gambar 1. Diagram Alir Analisis Kelompok

a. Deskripsi taksonomi (*taxonomy description*)

Deskripsi taksonomi merupakan tujuan tradisional analisis kelompok. Dengan taksonomi, dapat dikelompokkan sekumpulan objek secara empiris. Memang, analisis kelompok telah dipakai untuk keperluan-

keperluan pengelompokkan, namun kemampuannya tidak hanya sebatas itu. Analisis kelompok dapat pula dipakai untuk memunculkan hipotesis tentang struktur objek-objek yang diteliti. Jadi, sekalipun banyak dipandang sebagai alat analisis untuk keperluan eksploratori, analisis kelompok dapat pula dipakai untuk tujuan konfirmatori.

b. Simplikasi data (*data simplification*)

Dari struktur yang diperoleh, analisis kelompok menghasilkan perspektif sederhana mengenai observasi atau data dengan mengelompokkannya. Selanjutnya, analisis kelompok menetapkan struktur dari observasi atau data bukan variabel.

c. Identifikasi hubungan (*relationship identification*)

Setelah kelompok-kelompok terbentuk dan struktur data yang mendasarinya diperlihatkan dalam kelompok tersebut, peneliti mendapatkan informasi tentang hubungan antar observasi yang tidak mungkin diperoleh dengan menganalisis observasi secara individu. Analisis kelompok sering kali memperlihatkan kesamaan-kesamaan maupun perbedaan-perbedaan yang tidak terdeteksi metode-metode lain.

Setiap pengelompokan terhadap objek apa pun harus memiliki dasar untuk mengelompokkan. Pengelompokan dengan hanya memakai satu variabel umumnya tidak memuaskan karena tidak menggambarkan profil kelompok secara jelas. Dalam analisis kelompok pemilihan variabel-variabel yang dilibatkan tidak boleh sembarangan. Terdapat empat aturan dalam pemilihan variabel. Keempat aturan tersebut adalah sebagai berikut.

- a. Variabel yang digunakan dapat membentuk kelompok-kelompok yang bermanfaat dari segi ekonomi maupun pengetahuan.
- b. Variabel yang dipilih dapat memberikan gambaran tentang persamaan dalam kelompok dan perbedaan antarkelompok.
- c. Tidak menggunakan variabel yang memiliki level semua objek sama (*extremely homogenous*) dan variabel yang selalu berbeda dari objek yang satu dengan objek yang lainnya (*extremely heterogenous*).
- d. Memilih variabel-variabel yang secara konseptual diterima.

2. Desain penelitian dalam Analisis Kelompok

Sebelum melakukan proses pemilahan objek ke dalam beberapa kelompok, terlebih dahulu peneliti harus melakukan deteksi *outliers*, menentukan ukuran kesamaan, dan menentukan kebutuhan standarisasi data.

a. Mendeteksi *Outliers*

Outliers adalah data sampel yang memiliki karakteristik berbeda daripada mayoritas data sampel. Misalnya nilainya terlalu tinggi atau terlalu rendah dibandingkan sebagian besar data sampel yang lainnya. Dalam analisis kelompok, kehadiran *outliers* perlu dideteksi, sebab akan mengganggu analisis. Bila data telah tersedia, untuk mendeteksi keberadaan *outliers* dapat dilakukan dengan menggunakan dua cara berikut ini (Gudono, 2011:149):

1) Menggunakan *z-score*

Deteksi terhadap univariat *outlier* dapat dilakukan dengan menentukan nilai batas yang akan dikategorikan sebagai data *outlier*,

yaitu dengan cara mengkonversi nilai data ke dalam skor standardized atau yang biasa disebut *z-score*, yang memiliki nilai *means* (rata-rata) nol dan standar deviasi satu. Jika sebuah data memiliki *z-score* lebih besar dari 3 maka dipertimbangkan sebagai *outliers*. Jika data diketahui terdapat satu atau lebih data *outliers*, pada *outliers* tersebut bisa dilakukan beberapa penanganan. Data dihilangkan, jika data *outliers* tersebut didapat karena kesalahan pengambilan data, kesalahan pemasukan data pada komputer, dan sebagainya. Data *outlier* tetap dipertahankan dan tidak perlu dihilangkan, jika tidak terdapat kesalahan pada proses sampling maupun pemasukan data. Nilai *z-score* diperoleh menggunakan rumus persamaan berikut.

$$z = \frac{x_i - \bar{x}}{s} \quad (2.1)$$

dimana :

z = nilai standar/*z-score*

x_i = nilai observasi/data

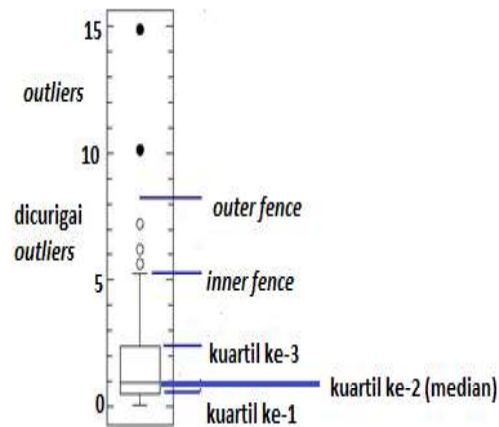
\bar{x} = mean data

s = deviasi standar sampel

2) Membuat *box plot*

Box plot atau *boxplot* (juga dikenal sebagai diagram box-and-whisker) merupakan suatu *box* (kotak berbentuk bujur sangkar). *Box plot* adalah cara standar untuk menampilkan distribusi data berdasarkan lima rangkuman (minimum, kuartil pertama, median, kuartil ketiga, dan maksimum). *Box plot* dapat diilustrasikan seperti Gambar 2.

Dalam penggambarannya, *box plot* dapat diilustrasikan secara horizontal maupun vertikal.



Sumber : <http://www.physics.csbsju.edu>

Gambar 2. Box plot

Dalam *box plot* yang paling sederhana, persegi panjang tengah membentang kuartil pertama ke kuartil ketiga (*interquartil range* atau IQR). Nilai IQR dapat diperoleh menggunakan rumus sebagai berikut (Walpole, 1992:63).

$$IQR = Q_3 - Q_1 \quad (2.2)$$

dengan

IQR = jarak antar kuartil

Q_3 = kuartil ke-3

Q_1 = kuartil ke-1

Rumus untuk Q_1 adalah:

$$Q_1 = \text{data ke } -\frac{n}{4} \quad (2.3)$$

Rumus untuk Q_3 adalah:

$$Q_3 = \text{data ke } -\frac{3n}{4} \quad (2.4)$$

dengan

n = banyaknya data

Segmen di dalam persegi panjang menunjukkan median, dan "kumis" di atas dan di bawah persegi panjang menunjukkan lokasi minimum dan maksimum. Data yang terletak antara pagar dalam (*inner fence*) dan pagar luar (*outer fence*) pantas dicurigai sebagai *outliers*. Data yang berada di luar *outer fence* harus sangat dicurigai sebagai *outliers*. *Outliers* adalah $3 \times \text{IQR}$ atau lebih di atas kuartil ketiga atau $3 \times \text{IQR}$ atau lebih di bawah kuartil pertama. Dugaan *outliers* adalah $1,5 \times \text{IQR}$ atau lebih di atas kuartil ketiga atau $1,5 \times \text{IQR}$ atau lebih di bawah kuartil pertama. Jika salah satu jenis *outlier* hadir, kumis di sisi yang sesuai diambil sampai $1,5 \times \text{IQR}$ dari kuartil (pagar dalam) dari titik data maksimum atau minimum, dan titik data individual ditampilkan sebagai lingkaran yang tidak terisi (untuk dugaan *outliers*) atau lingkaran yang diisi (untuk outlier). Pagar luar adalah $3 \times \text{IQR}$ dari kuartil (Hoffmann, 1991).

b. Menentukan ukuran kesamaan

Jarak yang dipakai sebagai ukuran kesamaan menjelaskan seberapa mirip atau seberapa berbeda objek-objek yang dikelompokkan. Misalkan terdapat p variabel. Pengukuran variabel V_1, V_2, \dots, V_p dilakukan terhadap setiap individu anggota sejumlah n dengan a_1, a_2, \dots, a_n sebagai

individu anggota. Jarak antar individu anggota dapat dinyatakan dalam bentuk matriks seperti pada Persamaan (2.5).

$$D_{n \times n} = \begin{bmatrix} D_{11} & D_{12} & \cdots & D_{1j} & \cdots & D_{1n} \\ D_{21} & D_{22} & \cdots & D_{2j} & \cdots & D_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ D_{i1} & D_{i2} & \cdots & D_{ij} & \cdots & D_{in} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ D_{n1} & D_{n2} & \cdots & D_{nj} & \cdots & D_{nn} \end{bmatrix} \quad (2.5)$$

Pada Persamaan (2.5), matriks D adalah matriks jarak yang terdiri atas kumpulan jarak antar objek sejumlah n dengan D_{ij} merupakan jarak antara individu ke- i dan ke- j (untuk setiap $i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, n$). Semakin kecil nilai D , maka semakin besar kemiripan antara kedua pengamatan tersebut. Sebaliknya bila D besar, semakin besar ketidakmiripan dari pengamatan tersebut (Yulianto & Hidayatullah, 2014). Terdapat beberapa ukuran jarak yang dapat digunakan untuk melakukan pengelompokan, beberapa di antaranya adalah sebagai berikut (Johnson & Wichern, 2007:673-674):

1) *euclidean distance* (jarak euclid)

$$D_{ij} = \sqrt{\sum_{k=1}^p (a_{ik} - a_{jk})^2} \quad (2.6)$$

dengan

D_{ij} = jarak antara objek ke- i dan ke- j

a_{ik} = nilai pengamatan objek ke- i variabel ke- k

a_{jk} = nilai pengamatan objek ke- j variabel ke- k

2) *squared euclidean distance*

$$D_{ij}^2 = \sum_{k=1}^p (a_{ik} - a_{jk})^2 \quad (2.7)$$

dengan

D_{ij} = jarak antara objek ke- i dan ke- j

a_{ik} = nilai pengamatan objek ke- i variabel ke- k

a_{jk} = nilai pengamatan objek ke- j variabel ke- k

3) *mahalanobis distance*

$$d(a_i, a_j) = \sqrt{(a_i - a_j)' S^{-1} (a_i - a_j)} \quad (2.8)$$

dengan

$d(a_{ik}, a_{jk})$ = jarak objek pengamatan x dan y

a_{ik} = nilai pengamatan objek ke- i

a_{jk} = nilai pengamatan objek ke- j

S = matriks varian kovarian

c. Standarisasi data

Sebelum melakukan standarisasi data, beberapa masalah harus diperhatikan, misalnya, jarak nilai dari masing-masing variabel karena perbedaan skala. Secara umum, variabel dengan penyebaran nilai yang tinggi mempunyai dampak yang lebih pada hasil akhir. Karena itu, pengukuran dari variabel-variabel diharapkan diketahui secara lengkap. Standarisasi data dilakukan jika ditemukan adanya perbedaan skala yang besar di antara variabel-variabel yang digunakan dalam penelitian. Untuk

melakukan standarisasi data menggunakan rumus seperti pada Persamaan (2.1).

3. Asumsi-asumsi dalam Analisis Kelompok

Analisis kelompok pada hakikatnya adalah teknik algoritme, bukan alat inferensi statistik. Oleh sebab itu persyaratan seperti distribusi data yang harus normal (*normality*) ataupun hubungan linier antar variabel (*linearity*) tidak menjadi syarat dalam analisis kelompok. Dalam analisis kelompok, masalah seberapa besar sampel mewakili populasi (representatif) dan adanya multikolinearitas harus lebih diperhatikan (Santoso, 2010:116). Kedua asumsi tersebut harus dipenuhi terlebih dahulu sebelum melakukan analisis lebih lanjut.

a. Representatif

Data yang diolah dalam analisis kelompok biasanya hanya sebagian kecil dari populasi sehingga agar hasilnya bisa digeneralisasi, data yang diolah sebaiknya mencerminkan gambaran umum atau bersifat representatif.

b. Multikolinieritas

Pengujian multikolinieritas digunakan untuk mengetahui ada tidaknya variabel independen yang memiliki kemiripan dengan variabel independen lainnya. Jenis uji parametrik untuk pengujian hubungan, yaitu korelasi *product moment pearson*. Pengujian ini digunakan untuk menguji dua variabel apakah ada hubungan atau tidak, dengan jenis data keduanya sama yaitu rasio atau interval. Rumus sebagai berikut:

$$r_{xy} = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{\{(n \sum x^2) - (\sum x)^2\} \{(n \sum y^2) - (\sum y)^2\}}} \quad (2.9)$$

dengan

r_{xy} = koefisien korelasi r pearson

n = ukuran sampel/observasi

x = variabel pertama

y = variabel kedua

Nilai koefisien korelasi merupakan nilai yang digunakan untuk mengukur kekuatan suatu hubungan antarvariabel. Koefisien korelasi memiliki nilai antara -1 hingga +1. Sifat nilai koefisien korelasi antara plus (+) atau minus (-). Makna sifat korelasi:

- 1) korelasi positif (+) berarti bahwa jika variabel x_1 mengalami kenaikan maka variabel x_2 juga akan mengalami kenaikan, begitu sebaliknya.
- 2) korelasi negatif (-) berarti bahwa jika variabel x_1 mengalami penurunan maka variabel x_2 akan mengalami kenaikan, begitu sebaliknya.

Sifat korelasi akan menentukan arah korelasi. Keeratan korelasi dapat dikelompokkan sebagai berikut:

- 1) 0,00 sampai 0,20 berarti korelasi memiliki keeratan sangat lemah
- 2) 0,21 sampai 0,40 berarti korelasi memiliki keeratan lemah
- 3) 0,41 sampai 0,70 berarti korelasi memiliki keeratan kuat
- 4) 0,71 sampai 0,90 berarti korelasi memiliki keeratan sangat kuat
- 5) 0,91 sampai 0,99 berarti korelasi memiliki keeratan kuat sekali

6) 1 berarti korelasi sempurna

Jika ada nilai korelasi antarvariabel bebas lebih besar dari 0,8 maka dapat diindikasikan adanya multikolinearitas (korelasi sangat kuat). Data yang mengandung multikolinearitas maka dilakukan proses analisis komponen utama.

Analisis Komponen Utama (*Principal Component Analysis*) adalah analisis multivariat yang mentransformasi variabel-variabel asal yang saling berkorelasi menjadi variabel-variabel baru yang tidak saling berkorelasi lagi (Rufaidah & Effindi, 2017). Variabel baru tersebut disebut sebagai komponen utama (*principal component*). Semua variabel dalam analisis komponen utama merupakan variabel dengan skala interval/rasio. Analisis tersebut mengurangi banyaknya dimensi variabel yang saling berkorelasi menjadi komponen utama yang tidak saling berkorelasi dengan mempertahankan sebanyak mungkin keragaman dalam himpunan data tersebut, artinya dimensi yang lebih kecil diharapkan lebih mudah melakukan penafsiran atau interpretasi tanpa kehilangan banyak informasi tentang data. Komponen utama (KU) yang terbentuk diharapkan seminimal mungkin, akan tetapi mampu menerangkan keragaman total yang maksimal. Banyaknya KU yang dipilih dapat ditentukan dengan beberapa prosedur, yaitu berdasarkan *eigenvalues*, *scree plot*, *percentage of variance accounted for*, *split-half reliability*, dan *significance test*.

Secara aljabar linier, KU adalah kombinasi-kombinasi linier tertentu dari p variabel acak V_1, V_2, \dots, V_p dimana koefisiennya adalah vektor

ciri (*eigen vector*). KU tergantung pada matriks kovarian (Σ) dan matriks korelasi (ρ) dari V_1, V_2, \dots, V_p dikarenakan vektor ciri dihasilkan dari akar ciri (*eigen value*) matriks kovarian atau dapat juga dihasilkan dari matriks korelasi.

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix} \text{ dan } \rho = \begin{bmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \dots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \dots & \rho_{pp} \end{bmatrix}$$

Melalui persamaan karakteristik matriks kovarian diperoleh akar ciri-akar cirinya, yaitu $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ dan vektor-vektor cirinya, yaitu e_1, e_2, \dots, e_p . Menyusutkan dimensi variabel asal V menjadi KU Y , karena itu sering disebut vektor pembobot.

KU pertama adalah kombinasi linier terbobot dari variabel asal yang mempunyai varian terbesar. KU kedua juga merupakan kombinasi linier terbobot dari variabel asal dengan varian terbesar kedua dan antara kedua KU tersebut tidak saling berkorelasi demikian seterusnya. Secara umum, KU ke- i dapat dituliskan sebagai,

$$Y_i = e_{1i}V_1 + e_{2i}V_2 + e_{3i}V_3 + \dots + e_{pi}V_p = e_i'V; i = 1, 2, 3, \dots, p \quad (2.10)$$

dimana e_i adalah vektor ciri (*eigen vector*) ke- i yang berpadanan dengan akar cirinya. Melalui persamaan di atas akan didapat $Var(Y_i) = \lambda_i$ dan $Cov(Y_{i-1}, Y_i) = 0$. Hal tersebut menunjukkan bahwa komponen utama tidak saling berkorelasi dan komponen utama ke- i memiliki keragaman sama dengan akar ciri ke- i . Oleh karena itu, keragaman total mampu diterangkan

oleh setiap komponen utama adalah proporsi antara akar ciri komponen utama tersebut dengan jumlah akar ciri matriks kovarian.

Apabila KU yang diambil adalah q buah, dimana $q < p$, maka proporsi keragaman data yang bisa diterangkan adalah:

$$\frac{\lambda_1 + \lambda_2 + \dots + \lambda_p}{p} \times 100\% \quad (2.11)$$

atau

$$\frac{\sum_{i=1}^p \lambda_i}{p} \times 100\% \quad (2.12)$$

Matriks kovarian Σ digunakan jika variabel yang diamati ukurannya pada skala dengan perbedaan tidak besar atau jika satuan ukurannya sama. Bila variabel yang diamati ukurannya pada skala dengan perbedaan sangat besar, atau ukurannya tidak sama, maka variabel tersebut perlu dibakukan sehingga KU ditentukan dari variabel baku. Variabel baku (Z) diperoleh dari transformasi terhadap variabel asal dalam matriks sebagai berikut:

$$Z = (V^{\frac{1}{2}})^{-1}(V - \mu) \quad (2.13)$$

dengan $V^{\frac{1}{2}}$ adalah matriks simpangan baku dengan unsur diagonal utama $(\sigma_{ii})^{1/2}$ dengan unsur lainnya adalah 0. Nilai harapan $E(Z) = 0$ dan keragamannya $Cov(Z) = \left(V^{\frac{1}{2}}\right)^{-1} S \left(V^{\frac{1}{2}}\right)^{-1} = r$. Dengan demikian KU dari Z dapat ditentukan dari vektor ciri yang diperoleh melalui matriks korelasi variabel asal r .

Untuk mencari akar ciri dan menentukan vektor pembobotnya sama seperti pada matriks S . Akar ciri yang didapat dari matriks korelasi akan

memiliki *trace* (jumlah seluruh akar ciri) sama dengan jumlah peubah yang dipakai (Yulianto, 2013:30-39).

$$\sum_{i=1}^p \lambda_i = \lambda_1 + \lambda_2 + \dots + \lambda_p = p \quad (2.14)$$

4. Prosedur Pengelompokan

Clustering algorithm (algoritme pengelompokan) yang akan digunakan untuk membentuk kelompok harus ditentukan pada tahap pertama. Selanjutnya diputuskan jumlah kelompok yang akan dibentuk. Dua hal ini mempunyai implikasi yang substansial tidak hanya pada hasil yang akan diperoleh tetapi juga pada interpretasi hasil tersebut. Secara umum algoritme pengelompokan digolongkan dalam dua kategori, yaitu metode hierarki dan metode nonhierarki.

a. Metode Hierarki

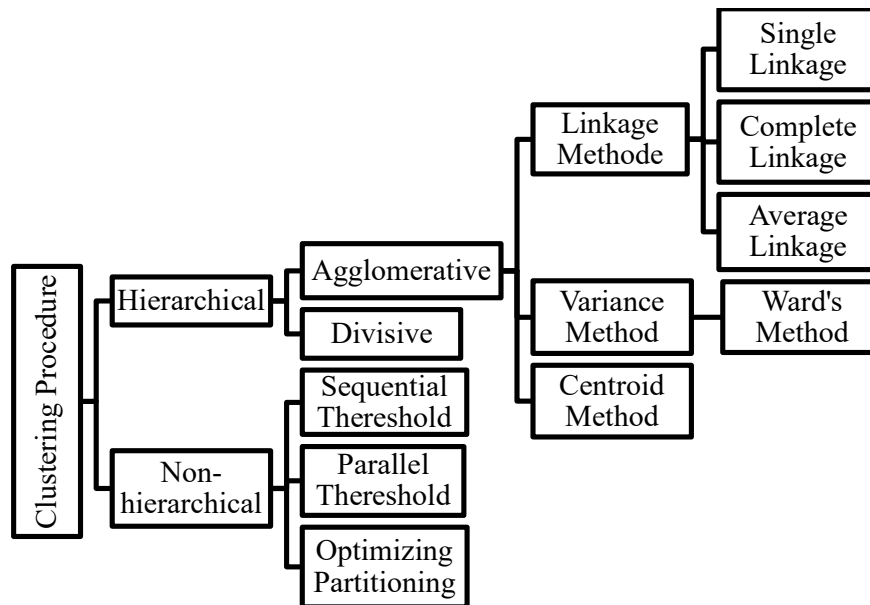
Dalam metode hierarki terdapat dua tipe dasar pengelompokan, yaitu *agglomerative* (pemusatan) dan *divisive* (penyebaran). Dalam metode pemusatan setiap objek atau observasi dianggap sebagai sebuah kelompok tersendiri. Kemudian dua kelompok yang mempunyai kemiripan digabungkan menjadi sebuah kelompok baru demikian seterusnya. Proses ini akan berjalan sehingga pada tahap terakhir, semua objek menjadi satu kelompok. Teknik ini terbagi menjadi *linkage method*, *variance method*, dan *centroid method*. Lalu, *linkage method* terbagi atas *single linkage method*, *complete linkage method*, dan *average linkage method*. *Variance method* terdiri atas *ward method*. Sebaliknya, dalam metode *divisive* dimulai dengan menempatkan semua objek sebagai satu kelompok. Selanjutnya, objek atau

observasi yang paling tinggi nilai ketidakmiripannya dipisahkan demikian seterusnya sampai semua objek menjadi kelompok-kelompok sendiri-sendiri.

b. Metode Nonhierarki

Kebalikan dari metode hierarki, metode nonhierarki tidak meliputi proses “*treelike construction*“. Justru metode nonhierarki menempatkan objek-objek ke dalam kelompok sekaligus sehingga terbentuk sejumlah kelompok tertentu. Langkah pertama adalah memilih sebuah kelompok sebagai kelompok pusat mula-mula, dan semua objek dalam jarak tertentu ditempatkan pada kelompok yang terbentuk. Kemudian memilih kelompok selanjutnya dan penempatan dilanjutkan sampai semua objek ditempatkan. Objek-objek bisa ditempatkan lagi jika jaraknya lebih dekat pada kelompok lain daripada kelompok asalnya. Metode nonhierarki berkaitan dengan *K-means clustering*, dan ada tiga pendekatan yang digunakan untuk menempatkan masing-masing observasi pada satu kelompok (*Sequential Threshold*, *Parallel Threshold*, dan *Optimization partitioning*).

Klasifikasi prosedur pengelompokan selengkapnya dapat dilihat pada Gambar 3.



Sumber : Analisis Multivariat Pemasaran (Simamora, 2005:256)

Gambar 3. Klasifikasi Prosedur Pengelompokan

Banyaknya kelompok yang akan dibentuk tidak dapat ditentukan secara sembarang. Penentuan banyaknya kelompok yang akan dibentuk harus berdasarkan suatu teori yang akan mempengaruhi jumlah kelompok secara langsung atau alami.

5. Interpretasi Kelompok

Pada tahap interpretasi dilakukan pengujian terhadap masing-masing kelompok untuk menamai atau memberikan keterangan secara tepat sebagai profil sifat dari kelompok. Ketika memulai proses interpretasi, ada satu ukuran yang sering digunakan yaitu *cluster centroid*. Jika prosedur pengelompokan dilakukan terhadap data asli, maka proses interpretasi akan memberikan gambaran yang logis. Tetapi jika data telah distandarisasi, peneliti harus

mengembalikan skor asli untuk variabel asal dan menghitung rata-rata profil menggunakan data tersebut.

6. Validasi dan Profil Kelompok

Tahap terakhir adalah melakukan uji validitas terhadap hasil pengelompokan yang dilakukan. Pengujian ini dilakukan untuk mengetahui jumlah kelompok optimal yang terbentuk. Terdapat beberapa indeks validitas yang dapat digunakan untuk menentukan jumlah kelompok optimal, diantaranya yaitu *Root Mean Squared Standard Deviation (RMSSTD)*, *R-Squared (RS)*, *Modified Hubert F Statistic (F)*, *Calinski-Harabasz index (CH)*, *I index (I)*, *Dunn's indices (D)*, *Silhouette index (S)*, *Davies Bouldin index (DB)*, *Xie-Beni index (XB)*, *SD validity index (SD)*, dan *S_Dbw validity index (S_Dbw)* (Liu, et al, 2010). Setelah diketahui kelompok optimal, maka tahap yang terakhir yang perlu dilakukan adalah mengidentifikasi profil dengan penggambaran karakteristik masing-masing kelompok untuk menjelaskan perbedaan masing-masing kelompok.

Secara umum langkah-langkah dalam analisis kelompok adalah sebagai berikut.

1. Menentukan variabel yang akan digunakan sebagai data penelitian.
2. Mendeteksi adanya outliers.
3. Memilih ukuran kesamaan (jarak) yang digunakan.
4. Melakukan standarisasi data jika diperlukan.
5. Melakukan pengujian terhadap asumsi-asumsi dalam analisis kelompok.
6. Memilih prosedur pengelompokan.

7. Menentukan jumlah kelompok yang akan dibentuk
8. Melakukan interpretasi terhadap kelompok yang terbentuk.
9. Melakukan uji validasi terhadap hasil pengelompokan untuk menentukan jumlah kelompok optimal.
10. membuat profil kelompok yang terbentuk

B. Metode Hierarki : *Single Linkage* dan *Complete Linkage*

Pengelompokan hierarki adalah metode analisis kelompok yang berusaha untuk membangun sebuah hierarki kelompok. Strategi untuk pengelompokan hierarki umumnya jatuh kepada dua jenis : aglomeratif dan divisif.

Pengelompokan hierarki aglomeratif (*agglomerative hierarchical clustering/AHC*) merupakan metode pengelompokan hierarki dengan pendekatan bawah-atas (*bottom up*). Proses pengelompokan dimulai dari masing-masing data sebagai satu buah kelompok, kemudian secara rekursif mencari kelompok terdekat sebagai pasangan untuk bergabung sebagai satu kelompok yang lebih besar. Proses tersebut diulang terus sehingga tampak bergerak ke atas membentuk jenjang (hierarki). Cara ini membutuhkan parameter kedekatan kelompok (*cluster proximity*) (Prasetyo, 2012:213).

Pengelompokan hierarki sering ditampilkan dalam bentuk grafis menggunakan diagram yang mirip pohon (*tree*) yang disebut dengan dendogram. Cabang dalam mewakili kelompok-kelompok. Cabang-cabang yang bergabung pada suatu node sepanjang sumbu jarak menunjukkan bahwa pada tingkat tersebut terjadi penggabungan. Oleh karena itu, sumbu vertikal menunjukkan jarak dan sumbu horizontal menunjukkan nomer identitas data (Jiang, et al, 2006).

Algoritme pengelompokan hierarki aglomeratif dijabarkan dalam algoritme berikut.

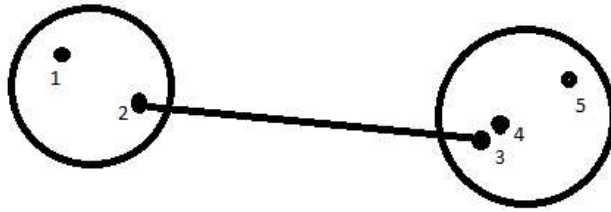
1. Hitung matriks kedekatan berdasarkan jenis jarak yang digunakan.
2. Ulangi langkah 3 sampai 4, hingga hanya satu klompok yang tersisa.
3. Gabung dua kelompok terdekat berdasarkan parameter kedekatan yang ditentukan.
4. Perbarui matriks kedekatan untuk merefleksikan kedekatan di antara kelompok baru dan kelompok asli yang sudah digabung.

Dalam pengelompokan hierarki terdapat beberapa metode yang digunakan untuk melakukan pengelompokan. Dua diantaranya adalah *single linkage* dan *complete linkage*.

a. *Single Linkage*

Single linkage (jarak terdekat) atau tautan tunggal merupakan metode hierarki yang mengelompokkan data berdasarkan tetangga terdekat (*nearest neighbour*) atau similaritas maksimum antara dua data (Amponsah, et al, 2013). Pada metode tautan tunggal (MIN), kedekatan di antara dua kelompok ditentukan dari jarak terdekat (terkecil) di antara pasangan di antara dua data dari dua kelompok yang berbeda (satu dari kelompok yang pertama dan satu dari kelompok yang lain) atau disebut juga nilai kemiripan termaksimal. Maka, dengan cara tersebut dimulai dari masing-masing data sebagai kelompok, kemudian mencari tetangga terdekat (*nearest neighbours*) dan menggunakan tautan tunggal untuk menggabungkan dua kelompok berikutnya hingga semuanya bergabung menjadi satu kelompok.

Dalam metode *single linkage* untuk menentukan jarak antar kelompok perlu melihat semua jarak antardua kelompok yang ada dan kemudian memilih yang terkecil (“terdekat” jaraknya). Secara teoritis hal tersebut dapat digambarkan seperti dalam Gambar 4.



Gambar 4. Jarak antar Kelompok dalam *Single Linkage Method*

Berikut rumus yang digunakan untuk menentukan jarak antar kelompok.

$$D_{(uv)w} = \min(D_{uw}, D_{vw}) \quad (2.15)$$

Dimana, D_{uw}, D_{vw} adalah jarak antara tetangga terdekat dari kelompok (U dan W) juga kelompok V dan W , begitu sebaliknya (Maylana, 2014).

Sesuai dengan Gambar 4, jarak antara item data 2 dan 3 adalah yang terpendek (terkecil) sehingga pasangan item data tersebut yang dipilih.

Sebagai contoh, terdapat data seperti pada Tabel 2. Terdapat enam item data (S1, S2, S3, S4, S5, dan S6). Jarak antara S1 dan S2 (D_{12}) dapat ditentukan berdasarkan rumus jarak euclid kuadrat (Persamaan 2.7) berikut.

$$D_{ij}^2 = \sum_{k=1}^p (a_{ik} - a_{jk})^2$$

Tabel 2. Data Nilai X_1 dan X_2

No.	X_1	X_2
1	17	5
2	17	4
3	8	11
4	9	13
5	20	18
6	18	20

Perhitungannya sebagai berikut.

$$D_{12}^2 = (17 - 17)^2 + (5 - 4)^2 = 1$$

Selanjutnya perhitungan jarak antara S1 dan S3 adalah 117; S2 dan S3 adalah 130 dan seterusnya dengan menggunakan Persamaan (2.7). Selengkapnya hasil perhitungan antar subjek dapat dilihat pada Tabel 3.

Tabel 3. Matriks Kesamaan (*Proximity*)

Item/ subjek	1	2	3	4	5	6
1	0	1	117	128	178	226
2	1	0	130	145	205	257
3	117	130	0	5	193	181
4	128	145	5	0	146	130
5	178	205	193	146	0	8
6	226	257	181	130	8	0

Jika prinsip *single linkage* tersebut diterapkan untuk data dalam matriks kesamaan seperti pada Tabel 3, maka akan terlihat bahwa jarak item data antara S1 dan S2 adalah yang terkecil ($D=1$), sehingga pasangan data tersebut dapat ditetapkan sebagai kelompok I ($=S1\&S2$). Selanjutnya setelah

S1 dan S2 bergabung menjadi kelompok I, terdapat 5 kelompok, yaitu S1&S2, S3, S4, S5, dan S6. Selanjutnya untuk menentukan item data mana yang akan dipilih sebagai pasangan kelompok berikutnya akan ditentukan lagi jarak antar data dan dikembangkan matriks kesamaan berikutnya. Dalam hal tersebut, misalnya, jarak antara kelompok I (yang beranggotakan S1 dan S2) dengan S3, dihitung dengan memilih mana yang paling kecil antara jarak S1 dan S3 (D_{13}) dan jarak antara S2 dan S3 (D_{23}). Perhitungannya adalah sebagai berikut.

$$D_{13}^2 = 117 \text{ dan } D_{23}^2 = 130 \text{ dipilih yang } 117$$

Untuk menentukan jarak euclid kuadrat dari S1 ke S3 ($= D_{13}^2$) tidak perlu lagi menghitung ulang. Cukup melihat data yang ada di Tabel 3. Begitupun dengan jarak item data yang lain. Setelah jarak euclid kuadrat semua item data ditentukan, diperoleh matriks kesamaan yang baru seperti pada Tabel 4.

Tabel 4. Matriks Kesamaan Tahap II

Item/ Subjek	1&2	3	4	5	6
1&2	0	117	128	178	226
3	117	0	5	193	181
4	128	5	0	146	130
5	178	193	146	0	8
6	226	181	130	8	0

Dari matriks di Tabel 4 diketahui bahwa kelompok berikutnya beranggotakan S3 dan S4, karena skor jaraknya terendah ($D=5$) dibandingkan terendah berikutnya ($D=8$). Kemudian dikembangkan lagi matriks kesamaan

(lihat Tabel 4). Jarak (S3&S4) dengan kelompok item 5 dihitung dengan membandingkan jarak S3 ke S5 dengan S4 ke S5 sebagai berikut.

$$D_{35}^2 = 193 \text{ dan } D_{45} = 146 \text{ dipilih yang } 146$$

Jika teknik tersebut diulang untuk item-item yang lainnya maka akan diperoleh Tabel Matriks Kesamaan (Tahap III) tampak pada Tabel 5 dan matriks kesamaan (tahap IV) tampak pada Tabel 6.

Tabel 5. Matriks Kesamaan (Tahap III)

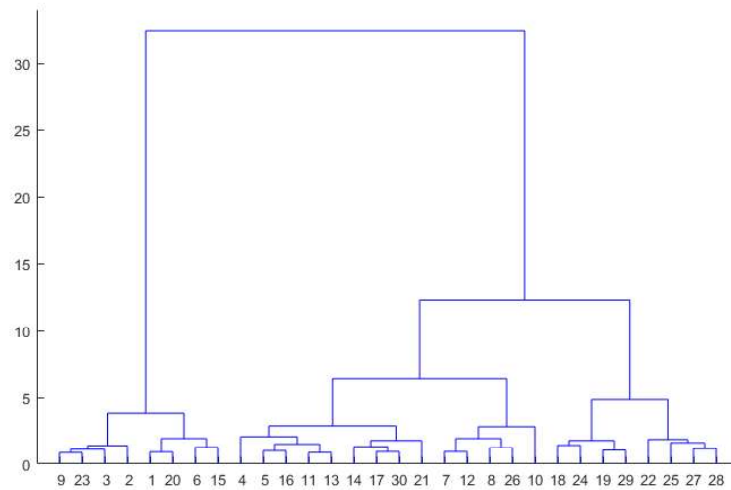
Item/ Subjek	1&2	3&4	5	6
1&2	0	117	178	226
3&4	117	0	146	130
5	178	146	0	8
6	226	130	8	0

Tabel 6. Matriks Kesamaan (Tahap IV)

Item/ Subjek	1&2	3&4	5&6
1&2	0	117	178
3&4	117	0	130
5	178	130	0

Jika dikehendaki dapat dilakukan pengelompokan lebih lanjut. Dengan memperhatikan skor jarak euclid kuadrat di Tabel 6, dapat disimpulkan selanjutnya kelompok (S1&S2) & (S3&S4) dapat dibentuk dan pada tahap selanjutnya semua item data bisa digabungkan menjadi sebuah kelompok beranggotakan semua item data. Tampak jelas bahwa metode hierarki secara bertahap (hierarkis) membentuk kelompok-kelompok sehingga

akhirnya semua item data tercakup menjadi satu bagian kelompok tertentu. Secara visual hal ini sering digambarkan dalam sebuah diagram pohon (dendrogram) seperti yang terlihat pada Gambar 5 (Gudono, 2011:262-265).



Sumber : <https://www.statistikian.com>

Gambar 5. Dendrogram

b. *Complete Linkage*

Untuk penjelasan mengenai langkah-langkah pembentukan kelompok menggunakan metode *complete linkage* secara umum sama seperti langkah-langkah pembentukan kelompok pada metode *single linkage*. Metode ini mengelompokkan data berdasarkan jarak terjauh dengan menggunakan Persamaan (2.16).

$$D_{(uv)w} = \max(D_{uw}, D_{vw}) \quad (2.16)$$

Dimana, D_{uw}, D_{vw} adalah jarak antara tetangga terjauh dari kelompok U dan W juga kelompok V dan W , begitu sebaliknya.



Gambar 6. Jarak antar Kelompok dalam *Complete Linkage Method*

Sesuai dengan Gambar 6, jarak antara item data 1 dan 5 adalah yang terpanjang (terbesar) sehingga pasangan item data tersebut yang dipilih.

C. *Davies Bouldin Index (Indeks Davies Bouldin)*

Indeks *Davies Bouldin* adalah pengukuran yang memaksimalkan jarak inter-kelompok antara kelompok C_k dan C_l dan pada waktu yang sama mencoba untuk meminimalkan jarak antar titik dalam sebuah kelompok. Andai $Sc(k)$ adalah jarak intra-kelompok dalam kelompok k . Rumus untuk menghitung $Sc(k)$ ialah (Sinurat, 2014):

$$Sc(k) = \frac{\sum_i^{N_k} |X_i - C_k|}{N_k} \quad (2.17)$$

dengan

N_k = banyak titik dalam kelompok k

C_k = centroid dari kelompok k

X_i = data ke- i dalam kelompok k

Jarak inter-kelompok didefinisikan:

$$d_{kl} = |C_k - C_l| \quad (2.18)$$

dengan

d_{kl} = jarak antara kelompok k dan kelompok l

C_k = centroid kelompok k

C_l = centroid kelompok l

Indeks *Davies Bouldin* dihitung dengan menggunakan rumus:

$$DB(nc) = \frac{1}{nc} \sum_{k=1}^{nc} \max \left(\frac{sc(k)+sc(l)}{d_{kl}} \right) \quad (2.19)$$

dengan

nc = banyak kelompok

Skema pengelompokan yang optimal menurut Indeks *Davies Bouldin* ialah yang memiliki nilai Indeks *Davies Bouldin* minimal.

D. Indikator Pendidikan

Pemenuhan atas hak untuk mendapatkan pendidikan yang bermutu merupakan ukuran keadilan dan pemerataan atas hasil pembangunan dan sekaligus merupakan investasi sumber daya manusia yang diperlukan untuk mendukung keberlangsungan pembangunan. Pemerataan, akses dan peningkatan mutu pendidikan akan membuat warga negara Indonesia memiliki kecakapan dalam rangka pembangunan manusia seutuhnya.

Dalam beberapa tahun ke depan pembangunan pendidikan nasional masih dihadapkan pada berbagai tantangan serius, terutama dalam upaya meningkatkan kinerja yang mencakup :

1. Pemerataan dan perluasan akses.
2. Peningkatan mutu, relevansi, dan daya saing.
3. Penataan tata kelola, akuntabilitas, dan citra publik.
4. Peningkatan pembiayaan.

Beberapa indikator *output* yang dapat menunjukkan kualitas pendidikan SDM antara lain Angka Melek Huruf (AMH), Tingkat Pendidikan, Angka Partisipasi Sekolah (APS), Angka Partisipasi Kasar (APK) dan Angka Partisipasi Murni (APM). Indikator *input* pendidikan salah satunya adalah fasilitas pendidikan (Said, et al., 2016).

1. Angka Melek Huruf (AMH)

Kegiatan membaca merupakan proses awal memasuki dunia pengetahuan yang begitu luas menuju masyarakat maju. Membaca akan mempermudah seseorang untuk memahami informasi terkait bidang kerja dan berbagai aspek yang menyangkut peningkatan kualitas hidup. Kemampuan baca-tulis dianggap penting karena melibatkan pembelajaran berkelanjutan oleh seseorang untuk dapat mencapai tujuan hidupnya. Hal ini berkaitan langsung dengan bagaimana seseorang mendapatkan pengetahuan, menggali potensinya dan berpartisipasi dalam pembangunan.

Salah satu indikator mendasar yang digunakan untuk melihat tingkat kemampuan membaca dan menulis adalah angka melek huruf (*literacy rate*). Kata “melek huruf” dapat diartikan sebagai kemampuan untuk dapat membaca dan menulis huruf latin/lainnya pada tingkat yang baik untuk berkomunikasi dengan orang lain atau dapat menyampaikan idenya dalam masyarakat yang mampu baca tulis. AMH merupakan salah satu indikator penting untuk mengukur keberhasilan implementasi kebijakan bidang pendidikan, dan kualitas sumber daya manusia suatu daerah. Penduduk usia 15 tahun keatas

merupakan masyarakat dewasa yang sudah seharusnya dapat membaca dan menulis huruf latin.

2. Angka Partisipasi Sekolah (APS)

Salah satu tujuan dari *Millenium Development Goals* (MDGs) adalah menjamin bahwa sampai dengan tahun 2015 semua anak, baik laki-laki maupun perempuan dapat menyelesaikan pendidikan dasar (*primary schooling*). Salah satu indikator yang dapat digunakan adalah Angka Partisipasi Sekolah (APS) untuk menilai pencapaian MDGs dengan melihat akses pendidikan pada penduduk usia sekolah. Semakin tinggi APS semakin besar jumlah penduduk yang mempunyai kesempatan untuk mengenyam pendidikan, namun bukan berarti meningkatnya APS juga meningkatnya pemerataan kesempatan masyarakat untuk mengenyam pendidikan.

Angka Partisipasi Sekolah (APS) adalah proporsi dari semua anak yang masih sekolah pada suatu kelompok umur tertentu terhadap penduduk dengan kelompok umur yang sesuai. Sejak Tahun 2009, Pendidikan Non Formal (Paket A, Paket B, dan Paket C) turut diperhitungkan. APS yang tinggi menunjukkan terbukanya peluang yang lebih besar dalam mengakses pendidikan secara umum. Pada kelompok umur mana peluang tersebut terjadi dapat dilihat dari besarnya APS pada setiap kelompok umur.

$$APS\ 7 - 12\ tahun = \frac{\text{Jumlah penduduk usia 7 - 12 tahun yang masih bersekolah}}{\text{jumlah penduduk usia 7 - 12 tahun}} \times 100\%$$

$$APS\ 13 - 15\ tahun = \frac{\text{Jumlah penduduk usia 13 - 15 tahun yang masih bersekolah}}{\text{jumlah penduduk usia 13 - 15 tahun}} \times 100\%$$

$$APS\ 16 - 18\ tahun = \frac{\frac{Jumlah\ penduduk\ usia\ 16 - 18\ tahun}{yang\ masih\ bersekolah}}{jumlah\ penduduk\ usia\ 16 - 18\ tahun} \times 100\%$$

$$APS\ 19 - 24\ tahun = \frac{\frac{Jumlah\ penduduk\ usia\ 19 - 24\ tahun}{yang\ masih\ bersekolah}}{jumlah\ penduduk\ usia\ 19 - 24\ tahun} \times 100\%$$

3. Angka Partisipasi Kasar (APK)

Angka Partisipasi Kasar (APK) adalah proporsi anak sekolah pada suatu jenjang tertentu terhadap penduduk pada kelompok usia tertentu. Sejak tahun 2007 Pendidikan Non Formal (Paket A, Paket B, dan Paket C) turut diperhitungkan. APK berguna untuk menunjukkan tingkat partisipasi penduduk secara umum pada suatu tingkat pendidikan. APK yang tinggi menunjukkan tingginya tingkat partisipasi sekolah, tanpa memperhatikan ketepatan usia sekolah pada jenjang pendidikannya. Jika nilai APK mendekati atau lebih dari 100 persen menunjukkan bahwa ada penduduk yang sekolah belum mencukupi umur dan atau melebihi umur yang seharusnya. Hal tersebut juga dapat menunjukkan bahwa wilayah tersebut mampu menampung penduduk usia sekolah lebih dari target yang sesungguhnya. Berikut adalah rumus yang digunakan untuk menghitung APK.

$$APK\ SD = \frac{Jumlah\ murid\ SD/sederajat}{jumlah\ penduduk\ usia\ 7 - 12\ tahun} \times 100\%$$

$$APK\ SMP = \frac{Jumlah\ murid\ SMP/sederajat}{jumlah\ penduduk\ usia\ 13 - 15\ tahun} \times 100\%$$

$$APK\ SM = \frac{Jumlah\ murid\ SM/sederajat}{jumlah\ penduduk\ usia\ 16 - 18\ tahun} \times 100\%$$

$$APK\ PT = \frac{Jumlah\ murid\ PT/sederajat}{jumlah\ penduduk\ usia\ 19 - 24\ tahun} \times 100\%$$

4. Angka Partisipasi Murni (APM)

Angka Partisipasi Murni (APM) merupakan persentase jumlah anak yang sedang bersekolah pada jenjang pendidikan yang sesuai dengan usianya terhadap jumlah seluruh anak pada kelompok usia sekolah yang bersangkutan. Sejak tahun 2007, Pendidikan Non Formal (Paket A, Paket B, dan Paket C) turut diperhitungkan. APM digunakan untuk mengukur daya serap sistem pendidikan terhadap penduduk usia sekolah. APM menunjukkan seberapa banyak penduduk usia sekolah yang sudah dapat memanfaatkan fasilitas pendidikan sesuai pada jenjang pendidikannya. Jika $APM = 100$, berarti seluruh anak usia sekolah dapat bersekolah tepat waktu. Berikut adalah rumus yang digunakan untuk menghitung APM.

$$APM_{SD} = \frac{\text{Jumlah murid SD usia 7 – 12 tahun}}{\text{jumlah penduduk usia 7 – 12 tahun}} \times 100\%$$

$$APM_{SMP} = \frac{\text{Jumlah murid SMP usia 13 – 15 tahun}}{\text{jumlah penduduk usia 13 – 15 tahun}} \times 100\%$$

$$APM_{SM} = \frac{\text{Jumlah murid SM usia 16 – 18 tahun}}{\text{jumlah penduduk usia 16 – 18 tahun}} \times 100\%$$

$$APM_{PT} = \frac{\text{Jumlah murid PT usia 19 – 24 tahun}}{\text{jumlah penduduk usia 19 – 24 tahun}} \times 100\%$$

5. Putus Sekolah, Angka Mengulang dan Angka Kelulusan

Angka putus sekolah mencerminkan persentase anak-anak usia sekolah yang sudah tidak lagi bersekolah/tidak menamatkan jenjang pendidikan tertentu. Indikator ini digunakan sebagai barometer pencapaian rencana strategi dalam rangka meningkatkan mutu, relevansi dan daya saing pendidikan. Banyak faktor yang menyebabkan anak terpaksa putus sekolah,

diantaranya adalah karena kurangnya kesadaran orang tua akan pentingnya pendidikan anak, keterbatasan ekonomi, keadaan geografis yang kurang menguntungkan, keterbatasan akses menuju ke sekolah, jarak sekolah yang jauh atau minimnya fasilitas pendidikan di suatu daerah.

Angka Mengulang dari tahun ajaran ke tahun ajaran juga menunjukkan tren yang terus menurun setiap tahun, hal ini menunjukkan bahwa siswa makin hari makin baik prestasi belajarnya sehingga dapat menurunkan Angka Mengulang.

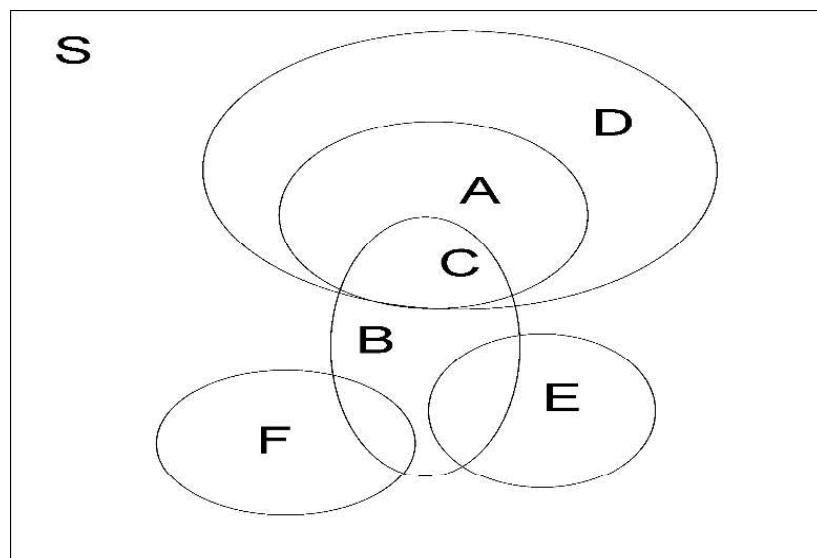
Angka Kelulusan SD hingga SMK sudah mencapai di atas 97 persen, bahkan Angka Kelulusan SD mencapai 99,70 persen (2015/2016) tertinggi dibandingkan jenjang pendidikan yang lain. Namun Angka Kelulusan SMK pada tahun 2015/2016 merupakan Angka Kelulusan terendah dibanding jenjang pendidikan yang lain, yaitu hanya 97,32 persen.

6. Rasio Murid-Guru dan Rasio Murid-Kelas

Salah satu indikator yang dapat digunakan untuk mengukur pemerataan dan perluasan akses pendidikan adalah Rasio Murid-Guru dan Rasio Murid-Kelas. Hal tersebut merupakan perbandingan antara jumlah murid dibandingkan dengan jumlah guru pada setiap jenjang pendidikan. Rasio Murid-Guru dapat memberikan gambaran besarnya beban guru dalam mengajar, sehingga apabila Rasio Murid-Guru semakin tinggi maka semakin banyak siswa yang harus dididik oleh seorang guru. Sedangkan Rasio Murid-Kelas menggambarkan daya tampung siswa dalam satu kelas. Semakin tinggi Rasio Murid-Kelas, maka semakin banyak jumlah siswa dalam satu kelas.

Keadaan ini berdampak pada suasana dan keberhasilan dalam kegiatan proses belajar mengajar.

Angka partisipasi pendidikan (APS, APK, dan APM) dapat direpresentasikan menggunakan Diagram Venn seperti pada Gambar 7. Berdasarkan gambar tersebut maka rumus APS, APK, dan APM secara sederhana dapat disajikan sebagai berikut.



Gambar 7. Diagram Venn Penduduk, Penduduk Usia Sekolah, dan Murid

$$APS = \frac{B}{D}$$

$$APK = \frac{A}{D}$$

$$APM = \frac{C}{D}$$

Keterangan:

A = penduduk berusia X yang masih bersekolah

B = Jumlah murid pada jenjang Y

C = jumlah murid berusia X pada jenjang Y

D = penduduk berusia X

E = penduduk berusia kerang dari X

F = penduduk berusia lebih dari X

S = penduduk semua usia