

## **BAB 2**

### **TINJAUAN PUSTAKA**

#### **2.1 Tinjauan Non Statistik**

##### **2.1.1 Pengertian Indeks Pembangunan Manusia (IPM)**

Indeks Pembangunan Manusia (IPM) merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat atau penduduk). Indikator pembangunan manusia merupakan salah satu alat ukur yang dapat digunakan untuk menilai kualitas pembangunan manusia, baik dari sisi dampaknya terhadap kondisi fisik manusia (kesehatan dan kesejahteraan) maupun yang bersifat non-fisik (intelektualitas) (BPS, 2019). Menurut Septiana (2008) Pendidikan memiliki peranan penting dalam pembangunan, pendidikan yang baik akan memberikan kontribusi terhadap kesejahteraan mereka.

##### **2.1.2 Pendidikan**

Pendidikan merupakan kebutuhan manusia yang sangat penting karena pendidikan mempunyai tugas untuk menyiapkan SDM bagi pembangunan bangsa dan negara untuk menunjukkan jati dirinya (Iswara dan Indrajaya, 2014). Pendidikan merupakan salah satu bidang penting yang terus diupayakan oleh pemerintah demi memperbaiki mutu masyarakat Indonesia. Namun realitanya, masyarakat sendiri enggan mendukung sepenuhnya dalam mengupayakan perbaikan kualitas pendidikan (Qomar, 2012). salah satunya sebagai berikut:

### 1. Angka Partisipasi Sekolah

Partisipasi berarti peran serta seseorang atau kelompok masyarakat dalam proses pembangunan baik dalam bentuk pernyataan maupun dalam bentuk kegiatan. Ukuran tingkat partisipasi masyarakat dilihat dari rasio jumlah penduduk yang memasuki jenjang pendidikan (SD/SMP/SMA/se-Derajat) terhadap jumlah penduduk secara keseluruhan setiap tahunnya. Indikator yang digunakan secara matematis oleh BPS dalam mengukur tingkat partisipasi disebut Angka Partisipasi Sekolah (Rodliyah, 2013).

### 2. Presentase Tamat SMP

SMP (Sekolah Menengah Pertama) merupakan pendidikan formal pada jenjang pendidikan dasar. Dilihat dari perkembangannya, anak usia SMP berada pada tahap perkembangan masa remaja (13-18 tahun) (Harlock, 1990).

#### 2.1.3 Kesehatan

Dalam Undang-Undang No. 23 Tahun 1992 dijelaskan bahwa pengertian Kesehatan adalah “keadaan sejahtera dari badan, jiwa, dan sosial yang memungkinkan setiap orang hidup produktif secara sosial dan ekonomis”.

#### 1. Rumah Tangga dan Akses Air Bersih

Kebutuhan akan akses air bersih yang mudah dan murah merupakan hak azasi bagi setiap warga negara. Pemerintah berkewajiban menyediakan kebutuhan akan air bersih bagi setiap penduduk. Namun dalam kenyataannya masih banyak penduduk yang kesulitan dalam hal akses air bersih, akibatnya banyak rumah tangga menggunakan semua sumber daya yang ada di rumah (Dian Wismar, 2008).

#### 2.1.4 Kependudukan

Kependudukan adalah perihal yang berkaitan dengan jumlah, struktur, umur, jenis kelamin, agama, kelahiran, perkawinan, kehamilan, kematian, persebaran, mobilitas dan kualitas serta ketahanannya yang menyangkut politik, ekonomi, sosial, dan budaya. Pengelolaan kependudukan untuk mengarahkan perkembangan untuk mewujudkan penduduk tumbuh seimbang dan mengembangkan kualitas penduduk. Perkembangan kependudukan adalah kondisi yang berhubungan dengan perubahan keadaan kependudukan yang dapat berpengaruh dan dipengaruhi oleh keberhasilan pembangunan berkelanjutan.

##### 1. Tingkat Partisipasi Angkatan Kerja

Tingkat partisipasi angkatan kerja sendiri merupakan suatu indikator ketenagakerjaan yang memberikan gambaran tentang penduduk yang aktif secara ekonomi dalam kegiatan sehari-hari merujuk pada suatu waktu dalam periode survei. Semakin besar jumlah penduduk yang tergolong bukan angkatan kerja, semakin kecil jumlah angkatan kerja yang mengakibatkan semakin kecil (Payaman J Simanjuntak, 2005:45).

##### 2. Tingkat Pengangguran Terbuka

Pengangguran adalah seseorang yang tergolong dalam angkatan kerja dan secara aktif mencari pekerjaan tetapi belum memperolehnya, pengangguran merupakan salah satu masalah kompleks yang dihadapi sebagian negara di dunia, baik di negara maju maupun negara berkembang. Dalam hal ini, negara berkembang tidak mampu menghadapi masalah tersebut dengan memberikan

tunjangan kepada warga yang menganggur, sedangkan negara maju mampu memberikan tunjangan tersebut (Sukirno,2004:28).

### 3. PDRB

Produk Domestik Regional Bruto pada dasarnya merupakan jumlah nilai tambah yang dihasilkan oleh seluruh unit usaha. Pembangunan daerah yang diharapkan akan membawa dampak positif pula terhadap pertumbuhan ekonomi. Pertumbuhan ekonomi daerah dapat dicerminkan dari perubahan Produk Domestik Regional Bruto dalam suatu wilayah (Suryono, 2010).

## 2.2 Tinjauan Statistik

### 2.2.1 Analisis Deskriptif

Statistika deskriptif adalah metode – metode yang berkaitan dengan pengumpulan dan penyajian suatu data sehingga memberikan informasi yang berguna (Walpole, 1995). Statistik deskriptif berfungsi untuk mendeskripsikan atau memberi gambaran terhadap objek yang diteliti melalui data sampel atau populasi, statistika deskriptif juga memberikan informasi hanya mengenai data yang dipunyai dan sama sekali tidak menarik inferensia atau kesimpulan (Sugiyono, 2007).

### 2.2.2 Analisis Cluster

Analisis Cluster adalah teknik pengelompokan untuk mengelompokkan objek berdasarkan karakteristik yang dimiliki oleh objek tersebut. Analisis cluster mengelompokkan objek yang paling mirip kedalam satu kelompok yang sama sehingga, karakteristik objek dalam satu cluster bersifat homogen sedangkan karakteristik objek antar cluster yang lain bersifat heterogen (Hair, *et al.* 2006).

Analisis Cluster terbagi menjadi dua, yaitu *metode hierarki* dan *non hierarki*.

Berikut penjelasan untuk masing-masing jenis :

1. *Hierarchical Methods* (Metode Hierarki)

Metode ini biasa digunakan untuk individu yang tidak terlalu banyak, dan jumlah kelompok yang hendak dibentuk belum diketahui. Pengelompokan ini disajikan dalam bentuk dendrogram, yang mirip dengan struktur diagram pohon atau *tree diagram* (Usman dan Sobari, 2013). Metode hierarki terbagi menjadi dua, yaitu *Agglomerative* atau Metode Penggabungan dan *Divisive* atau Metode Pembagian.

Metode *agglomerative* dimulai dengan setiap objek dalam satu kelompok yang terpisah. Kelompok dibentuk dengan mengelompokkan objek ke dalam kelompok yang semakin membesar. Proses ini dilanjutkan sampai semua objek menjadi anggota dari suatu kelompok tunggal (*a single cluster*). Sedangkan metode *divisive* dimulai dari semua objek dikelompokkan menjadi kelompok tunggal. Kemudian kelompok tersebut dibagi atau dipisah, sampai setiap objek berada dalam kelompok yang terpisah (Sangga, 2018). Beberapa metode pengklasteran hierarki antara lain.

- a. *Single linkage* (jarak terdekat atau tautan tunggal), pengklasteran ini memberikan hasil bila kelompok-kelompok digabungkan menurut jarak antara anggota-anggota yang terdekat di antara dua kelompok (Prasetyo, 2012).

- b. *Complete linkage* (jarak terjauh atau tautan lengkap), pengklasteran ini terjadi bila kelompok-kelompok digabungkan menurut jarak antara anggota-anggota yang terjauh di antara dua kelompok (Prasetyo, 2012).
  - c. *Average linkage* (jarak rata-rata atau tautan rata-rata), pengklasteran ini menggabungkan objek menurut jarak rata-rata pasangan-pasangan anggota masing-masing pada himpunan antara dua kelompok (Prasetyo, 2012).
  - d. *Ward's method*, metode ini menggunakan perhitungan yang lengkap dan memaksimalkan homogenitas di dalam satu kelompok (Dillon dan Goldstein, 1984).
2. *Non Hierarchical Methods* (Metode Non Hirarki)
- Metode Non Hierarki Analisis klaster dengan metode non hierarki merupakan metode klaster yang menentukan jumlah klaster secara manual.. Prosedur pada metode non hierarki dimulai dengan memilih sejumlah nilai cluster awal sesuai dengan jumlah yang diinginkan, teknik analisis klaster nonhierarki dirancang untuk mengelompokkan item bukan variabel menjadi kumpulan K klaster. Metode non hierarki ini meliputi metode sequential threshold, parallel threshold, dan optimizing partitioning (Gudono, 2011).
- a. *Metode Sequential Threshold*, pada metode Sequential Threshold dengan pemilihan satu cluster dan menempatkan semua obyek yang berada pada jarak terdekat ke dalam cluster tersebut.
  - b. *Metode Parallel Threshold*, secara prinsip sama dengan metode sequential threshold, hanya saja pada metode parallel threshold dilakukan pemilihan

terhadap beberapa obyek awal cluster sekaligus dan kemudian melakukan penggabungan obyek ke dalamnya secara bersamaan.

- c. *Metode Optimization*, Metode Optimization hampir mirip dengan metode Sequential Threshold dan metode Parallel Threshold yang membedakan adalah metode optimization ini memungkinkan untuk menempatkan kembali obyek-obyek ke dalam cluster yang lebih dekat atau dengan melakukan optimasi pada penempatan obyek yang ditukar untuk cluster lainnya dengan pertimbangan kriteria optimasi.

Ada dua masalah utama pendekatan non hierarki. Pertama, banyaknya cluster harus ditentukan terlebih dahulu. Kedua, pemilihan pusat cluster tidak menentu (pasti). Seterusnya, hasil pengclusteran tergantung pada bagaimana pusat cluster dipilih. Banyak program yang dimulai dengan memilih kasus pertama  $k$  ( $k$  = jumlah cluster) sebagai pusat cluster awal. Jadi, hasil pengclusteran tergantung pada observasi data. Kekurangan metode ini dapat dilakukan dengan cepat dan sangat bermanfaat kalau jumlah observasi besar (Simamora, 2005).

### 2.2.3 Pengertian Metode *K-Medoids*

*K-Medoids* atau *Partitioning Around Medoids* (PAM) adalah algoritma *clustering* yang mirip dengan *K-Means*. Algoritma *K-Medoids* ini diusulkan oleh Leonard Kaufmann dan Peter J. Rousseeuw pada tahun 1987 (Setiyawati, 2017). *K-Medoids* merupakan suatu algoritma yang digunakan untuk menemukan *medoids* didalam sebuah kelompok (*cluster*) yang merupakan titik pusat dari suatu kelompok (*cluster*).

Algoritma *K-Medoids* lebih baik dibandingkan dengan *K-Means* karena pada *K-Medoids* kita menemukan  $k$  sebagai objek yang representatif untuk meminimalkan jumlah ketidaksamaan objek data, sedangkan pada *K-Means* menggunakan jumlah jarak *euclidean distances* untuk objek data (Marlina D, 2008). Analisis *cluster* yang menggunakan prosedur partisi salah satunya adalah *Medoids*. *Medoids* adalah objek yang dianggap mewakili *cluster* sekaligus sebagai pusat *cluster*. *K-Medoids Clustering* hadir untuk mengatasi kelemahan *K-Means Clustering* yang sensitif terhadap *outlier* karena suatu objek dengan suatu nilai yang besar mungkin secara substansial menyimpang dari distribusi data (Han dan Kamber, 2006).

#### 2.2.4 Ukuran Jarak

Pada metode *cluster* membutuhkan suatu ukuran seberapa mirip objek (jarak) yang akan di kelompokkan. Adapun Jarak yang biasa digunakan dalam analisis pengklasteran diantaranya ada *Euclidian* (Charles, 2015) sebagai berikut:

##### 1. Jarak *Euclidean*

Jarak *Euclidean* adalah akar dari jumlah kuadrat perbedaan/deviasi di dalam nilai untuk setiap variabel (Supranto, 2004). Jarak *Euclidean* juga biasa disebut sebagai metode perhitungan jarak yang didasarkan pada ruang berdimensi terbatas bernilai riil (Kumari dan Bhagat, 2013). Adapun persamaan untuk menghitung jarak *Euclidean* sebagai berikut:

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

Keterangan :

$d_{ij}$  = Tingkat perbedaan jarak antara objek  $i$  dan objek  $j$

$n$  = Jumlah variabel cluster

$X_{ik}$  = Dari data variabel ke  $i$  pada variabel ke  $k$

$X_{jk}$  = Dari data variabel ke  $j$  pada variabel ke  $k$

### 2.2.5 Tahapan Metode *K-Medoids*

Tahapan - tahapan pengelompokan menggunakan metode *K-Medoids Clustering* dapat dijelaskan sebagai berikut:

1. Menentukan  $k$  sebagai jumlah variabel cluster);
2. Menentukan nilai centroid untuk pusat pada cluster;
3. Menempatkan objek *non medoids* ke dalam cluster yang paling dekat dengan *medoids*, dengan menggunakan rumus *Euclidian* . Berikut adalah persamaan *Euclidian*;

$$d_{x_i y_i} = \sqrt{(x_i - E_x)^2 + (y_i - E_y)^2}$$

4. Menghitung jarak setiap objek yang berada pada masing – masing cluster dengan menentukan objek *non medoids*;
5. Hitung total simpangan ( $S$ ) dengan menghitung nilai total cluster baru – total cluster lama. Jika  $S < 0$ , maka tukar objek dengan data cluster untuk membentuk sekumpulan  $k$  objek baru sebagai *medoids*.

Ulangi langkah 3 sampai 5 hingga tidak terjadi perubahan *medoids*, sehingga didapatkan cluster.

### 2.2.6 Validasi Cluster

Menurut Dalton, et al.(2009), salah satu pengukuran validasi internal adalah *Silhouette Coefisien*. *Silhouette Coefisien* mengevaluasi objek secara visual baik di dalam cluster maupun yang berada diluar, berdasarkan rentang nilai *silhouettenya*. Nilai *silhoutte* objek mendefinisikan kedekatannya dengan kelompok sendiri dengan kelompok lain. Rumus rentang *silhoutte* sebagai berikut:

$$SC = \frac{1}{N} \sum_{i=1}^N S(xi) \quad \text{dengan}$$

$$S(xi) = \frac{b-a}{\max[b,a]}$$

Keterangan :

a = rata – rata kemiripan antara objek ke-i dengan objek lainnyadalam satu cluster

b = minimum kemiripan antara objek ke-i dengan objek lain di masing – masing cluster

$S(xi)$  = lebar *Silhouette* cluster ke – i

Menurut Lewis (2010), jika nilai *silhoutte coefisien* ( $sci$ ) < 0 maka objek berada dalam kelompok yang salah sedangkan jika nilai *silhoutte coefisien* ( $sci$ ) > 0 maka objek berada dalam kelompok yang benar dan jika nilai *silhoutte coefisien* ( $sci$ ) = 0 maka objek berada diantara dua cluster sehingga objek tersebut belum dapat diputuskan masuk di dalam cluster yang mana. Ketika nilai rata-rata rentang *silhoutte coefisien* objek secara keseluruhan ( $SC$ ) > 0,51 menunjukkan bahwa struktur hasil pengelompokan sudah dianggap baik, sedangkan rata-rata rentang *silhoutte coefisien* objek secara keseluruhan ( $SC$ ) > 0,71 menunjukkan bahwa struktur hasil pengelompokan sangat baik.