

## **BAB III**

### **LANDASAN TEORI**

#### **3.1. Pendapatan**

##### **3.1.1. Pengertian pendapatan**

Pendapatan merupakan suatu hasil yang diterima oleh seseorang atau rumah tangga dari berusaha atau bekerja. Jenis masyarakat bermacam ragam, seperti bertani, nelayan, beternak, buruh, serta berdagang dan juga bekerja pada sektor pemerintah dan swasta (Nazir, 2010).

Pendapatan menurut ilmu ekonomi diartikan sebagai nilai maksimum yang dapat dikonsumsi oleh seseorang dalam satu periode seperti keadaan semula. Definisi tersebut menitik beratkan pada total kuantitatif pengeluaran terhadap konsumsi selama satu periode. Dengan kata lain pendapatan merupakan jumlah harta kekayaan awal periode ditambah keseluruhan hasil yang diperoleh selama satu periode, bukan hanya yang dikonsumsi. Secara garis besar pendapatan didefinisikan sebagai jumlah harta kekayaan awal periode ditambah perubahan penilaian yang bukan diakibatkan perubahan modal dan hutang (Irawan Suparmoko, 2008).

Pendapatan adalah keseluruhan penghasilan yang diterima baik dari sektor formal maupun nonformal yang dihitung dalam jangka waktu tertentu. BPS (2011), mengukur pendapatan masyarakat bukanlah pekerjaan yang mudah, oleh karena itu BPS melakukan perhitungan pendapatan dengan menggunakan pengeluaran atau konsumsi masyarakat. Hal ini didasari oleh paradigma bahwa bila pendapatan mengalami kenaikan maka akan diikuti oleh berbagai kebutuhan yang semakin banyak sehingga menuntut pengeluaran yang tinggi pula.

Kebutuhan dan keinginan tidak terbatas jumlahnya, hanya saja kebutuhan dan keinginan tersebut dibatasi dengan jumlah pendapatan yang diterima oleh seseorang. Pendapatan yang diterima oleh masyarakat tentu berbeda antar satu dengan yang lainnya, hal ini disebabkan berbedanya jenis pekerjaan yang dilakukannya. Perbedaan pekerjaan tersebut dilatar belakangi oleh tingkat pendidikan, *skill* dan pengalaman dalam bekerja. Indikator tingkat kesejahteraan dalam masyarakat dapat diukur dengan pendapatan yang diterimanya. Peningkatan taraf hidup masyarakat dapat digambarkan dari kenaikan hasil *real income* perkapita, sedangkan taraf hidup tercermin dalam tingkat dan pola konsumsi yang meliputi unsur pangan, pemukiman, kesehatan, dan pendidikan untuk mempertahankan derajat manusia secara wajar.

Kesimpulannya pendapatan adalah penghasilan yang diterima seseorang atau rumah tangga dari berusaha atau bekerja, yang berupa uang maupun barang yang diterima atau dihasilkan dalam jangka waktu tertentu.

### **3.1.2. Variabel-variabel yang mempengaruhi pendapatan**

Pada hakikatnya pendapatan yang diterima oleh seseorang maupun badan usaha tentunya dipengaruhi oleh banyak faktor, seperti tingkat pendidikan dan pengalaman seorang, semakin tinggi tingkat pendidikan dan pengalaman maka makin tinggi pula tingkat pendapatannya, kemudian juga tingkat pendapatan sangat dipengaruhi oleh modal kerja, jam kerja, akses kredit, jumlah tenaga kerja, tanggungan keluarga, jenis barang dagangan (produk) dan faktor lainnya. Pada umumnya masyarakat selalu mencari tingkat pendapatan tinggi untuk memenuhi kebutuhan rumah tangganya, akan tetapi dibatasi oleh beberapa faktor tersebut (Nazir, 2010).

Menurut Arfida BR (2003: 157-159) berbagai tingkat upah atau pendapatan terkait dalam struktur tertentu yaitu:

1. Sektoral

Struktur upah sektoral mendasarkan diri pada kenyataan bahwa kemampuan satu sektor berbeda dengan yang lain. Perbedaan karena alasan kemampuan usaha perusahaan. Kemampuan finansial perusahaan ditopang oleh nilai produk pasar.

2. Jenis jabatan

Dalam batas-batas tertentu jenis-jenis jabatan sudah mencerminkan jenjang organisatoris atau keterampilan. Perbedaan upah karena jenis jabatan merupakan perbedaan formal.

3. Geografis

Perbedaan upah lain mungkin disebabkan karena letak geografis pekerjaan. Kota besar cenderung memberikan upah yang lebih tinggi dari pada kota kecil atau pedesaan.

4. Keterampilan

Perbedaan upah yang disebabkan keterampilan adalah jenis perbedaan yang paling mudah dipahami. Biasanya jenjang keterampilan sejalan dengan jenjang berat-ringannya pekerjaan.

5. Seks

Perbedaan diakibatkan jenis kelamin, di mana seringkali upah golongan wanita lebih rendah daripada apa yang diterima laki-laki, *ceteris paribus*.

6. Ras

Meskipun menurut hukum formal perbedaan upah karena ras tidak boleh terjadi, namun kenyataannya perbedaan itu ada. Hal ini mungkin karena produk kebudayaan masa lalu, sehingga terjadi *stereo type* tenaga menurut ras atau daerah asal.

7. Faktor lain

Daftar penyebab perbedaan ini mungkin dapat diperpanjang dengan memasukan faktor-faktor lain, seperti masa hubungan kerja, ikatan kerja dan lainnya.

Sedangkan menurut Sukirno (2008), faktor-faktor yang menimbulkan perbedaan upah antara lain:

1. Permintaan dan Penawaran Tenaga Kerja

Permintaan dan penawaran tenaga kerja dalam sesuatu jenis pekerjaan sangat besar peranannya dalam menentukan upah di sesuatu jenis pekerjaan. Di dalam sesuatu pekerjaan di mana terdapat penawaran tenaga kerja yang cukup besar tetapi tidak banyak permintaannya, upah cenderung rendah. Sebaliknya didalam sesuatu pekerjaan dimana terdapat penawaran tenaga kerja yang terbatas tetapi permintaannya sangat besar, upah cenderung tinggi.

2. Perbedaan corak pekerjaan

Kegiatan ekonomi meliputi berbagai jenis pekerjaan. Ada diantara pekerjaan tersebut merupakan pekerjaan ringan dan sangat mudah dikerjakan. Tetapi ada pula pekerjaan yang harus dikerjakan dengan mengeluarkan tenaga fisik yang besar, dan ada pula pekerjaan yang harus dilakukan dalam lingkungan yang kurang menyenangkan.

3. Perbedaan kemampuan, keahlian, dan pendidikan

Kemampuan, keahlian, ketrampilan para pekerja di dalam sesuatu jenis pekerjaan adalah berbeda. Jika hal tersebut lebih tinggi maka produktivitas akan lebih tinggi sehingga upah yang didapat pun akan lebih tinggi. Tenaga kerja yang lebih berpendidikan memperoleh pendapatan yang lebih tinggi karena pendidikan mempertinggi kemampuan kerja dan kemampuan menaikkan produktivitas.

4. Pertimbangan Bukan Uang

Daya tarik sesuatu pekerjaan bukan saja tergantung kepada besarnya upah yang ditawarkan. Ada tidaknya perumahan yang tersedia, jauh dekatnya rumah pekerja, apakah berada di kota besar atau di tempat yang terpencil, dan pertimbangan lainnya. Faktor-faktor bukan keuangan seperti ini mempunyai peranan yang cukup penting pada waktu seseorang memilih

pekerjaan. Seseorang sering kali menerima upah yang rendah apabila pertimbangan bukan keuangan sesuai dengan keinginannya.

## 5. Mobilitas Pekerja

Upah dari sesuatu pekerjaan di berbagai wilayah dan bahkan di dalam sesuatu wilayah tidak selalu sama. Salah satu faktor yang menimbulkan perbedaan tersebut adalah ketidaksempurnaan dalam mobilitas tenaga kerja. Ketidaksempurnaan mobilitas pekerja disebabkan oleh faktor geografis dan institusional.

### **3.2. Pendidikan**

#### **3.2.1. Pengertian pendidikan**

Menurut UU Nomor 20 Tahun 2003, pendidikan adalah usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta ketrampilan yang diperlukan dirinya, masyarakat, bangsa dan negara. Pendidikan yaitu tuntunan di dalam hidup tumbuhnya anak-anak. Adapun maksudnya pendidikan yaitu, menuntun segala kekuatan kodrat yang ada pada anak-anak itu, agar mereka sebagai manusia dan sebagai anggota masyarakat dapatlah mencapai keselamatan dan kebahagiaan yang setinggi-tingginya (Siswoyo, 2007).

Dari beberapa pengertian di atas, maka dapat disimpulkan bahwa pendidikan adalah suatu usaha yang dilakukan secara sadar dan sengaja untuk mengubah tingkah laku manusia, baik secara individu maupun kelompok untuk mendewasakan manusia melalui upaya pengajaran dan pelatihan.

#### **3.2.2. Jenjang pendidikan**

Menurut Undang-undang Republik Indonesia No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional Bab VI Pasal 14-19, jenjang pendidikan di Indonesia adalah sebagai berikut :

1. Pendidikan Dasar

Pendidikan dasar merupakan jenjang paling dasar pendidikan di Indonesia yang mendasari pendidikan menengah Anak usia 7 – 15 tahun diwajibkan mengikuti pendidikan dasar. Bentuk pendidikan dasar adalah Sekolah Dasar (SD) atau Madrasah Ibtidaiyah (MI) dan Sekolah Menengah Pertama (SMP) atau Madrasah Tsanawiyah (MTs).

2. Pendidikan Menengah

Pendidikan menengah merupakan lanjutan dari jenjang pendidikan dasar. Pendidikan menengah diselenggarakan selama 3 tahun dan terdiri atas Sekolah Menengah Umum (SMA) atau Sekolah Menengah Kejuruan (SMK).

3. Pendidikan Tinggi

Pendidikan tinggi merupakan jenjang pendidikan setelah pendidikan menengah yang mencakup program pendidikan diploma, sarjana, magister, spesialis, dan doktor yang diselenggarakan oleh perguruan tinggi. Pendidikan tinggi diselenggarakan untuk menyiapkan peserta didik menjadi anggota masyarakat yang memiliki kemampuan akademik atau profesional yang dapat menerapkan, mengembangkan, dan bahkan menciptakan ilmu pengetahuan, teknologi atau kesenian.

Jenjang pendidikan tersebut adalah jenjang pendidikan yang secara resmi dan wajib diikuti oleh peserta didik dalam jalur pendidikan formal, tetapi ada tahap pendidikan yang tidak wajib dilaksanakan yaitu pendidikan anak usia dini sebelum mengikuti pendidikan dasar. Pendidikan Anak Usia Dini (PAUD) antara lain adalah Taman Kanak-kanak (TK), dan *Raudatful Atfal* (RA) yang berada dibawah naungan Departemen Agama.

### **3.3. Kelompok Umur Tenaga Kerja**

Tenaga kerja adalah orang yang siap masuk dalam pasar kerja sesuai dengan upah yang ditawarkan oleh penyedia pekerjaan. Jumlah tenaga kerja dihitung dari penduduk usia produktif (usia 15 tahun-65 tahun) yang masuk kategori angkatan kerja (*labour force*). Penduduk yang bekerja pada umumnya didominasi oleh penduduk yang berusia 25 tahun-54 tahun yang merupakan usia prima (BPS, 2008: 19).

Kondisi di negara berkembang pada umumnya memiliki tingkat pengangguran yang jauh lebih tinggi dari angka resmi yang dikeluarkan oleh pemerintah. Hal ini terjadi karena ukuran sektor informal masih cukup besar sebagai salah satu lapangan nafkah bagi tenaga kerja tidak terdidik. Sektor informal tersebut dianggap sebagai katup pengaman bagi pengangguran.

Angka resmi tingkat pengangguran umumnya menggunakan indikator pengangguran terbuka, yaitu jumlah angkatan kerja yang secara sungguh-sungguh tidak bekerja sama sekali dan sedang mencari kerja pada saat survei dilakukan. Sementara yang setengah pengangguran dan pengangguran terselubung tidak dihitung dalam angka pengangguran terbuka, karena mereka masih menggunakan waktu produktifnya selama seminggu untuk bekerja meskipun tidak sampai 35 jam penuh. Sehingga kelompok umur sangatlah penting dalam prasyarat pencari kerja, karena yang cukup umur (15 tahun-65 tahun) adalah orang-orang yang telah siap terjun dalam lapangan kerja.

### **3.4. Lapangan Pekerjaan**

#### **3.4.1. Pengertian pekerjaan informal dan formal**

Menurut Undang-Undang (UU) Ketenagakerjaan No. 13 tahun 2003, pekerja informal mengacu pada orang yang bekerja tanpa relasi kerja, yang berarti tidak ada perjanjian yang mengatur elemen-elemen kerja, upah dan kekuasaan. Kementerian Tenaga Kerja dan Transmigrasi secara umum mendefinisikan sektor informal sebagai semua bisnis komersial dan non-komersial (atau aktivitas ekonomi) yang tidak terdaftar, yang tidak memiliki struktur organisasi formal dan

secara umum memiliki ciri-ciri: dimiliki oleh keluarga, kegiatan berskala kecil, padat karya, menggunakan teknologi yang diadaptasi dan bergantung pada sumber daya lokal. Konferensi Internasional Statistik Tenaga Kerja ke-17 (dalam BPS, 2013) mendefinisikan pekerja sektor informal sebagai “karyawan dengan hubungan kerja yang tidak tercakup dalam perundang-undangan atau dalam praktiknya, tidak tunduk pada undang-undang tenaga kerja, pajak, pendapatan, perlindungan sosial atau hak tertentu untuk jaminan kerja tertentu (pemberitahuan pemecatan sebelumnya, pembayaran yang buruk, dibayar tahunan atau ijin sakit, dll).”

Menurut Todaro (2011), sektor informal dicirikan dengan adanya sejumlah besar aktivitas produksi dan jasa skala kecil yang dilakukan individu atau memiliki keluarga dan menggunakan teknologi sederhana padat karya. Sektor informal cenderung beroperasi seperti perusahaan-perusahaan yang bersaing secara monopolistik yang memiliki kemudahan memasuki pasar, kapasitas berlebih, dan adanya persaingan yang menurunkan laba sampai ke tingkat harga penawaran tenaga kerja.

Dalam menghitung pekerja informal, BPS melakukan pendekatan khusus dalam menentukan penduduk yang bekerja di sektor formal/informal yaitu berdasarkan status pekerjaan dalam pekerjaan utama dan Jenis Pekerjaan/Jabatan (BPS, 2013). Menurut statusnya, pekerja dikategorikan menjadi tujuh:

1. Berusaha sendiri.
2. Berusaha dibantu buruh tidak tetap/buruh tidak dibayar.
3. Berusaha dibantu buruh tetap/buruh dibayar.
4. Buruh/karyawan/pegawai.
5. Pekerja bebas di pertanian.
6. Pekerja bebas di non pertanian.
7. Pekerja keluarga/tak dibayar.



Kategori 3 dan 4 umumnya mengacu pada pekerja di sektor formal, sementara kategori lainnya adalah sektor informal. Dari jenis pekerjaan utama, BPS telah menetapkan 10 kategori antara lain:

1. Tenaga profesional.
2. Tenaga kepemimpinan.
3. Pejabat pelaksana dan tata usaha.
4. Tenaga penjualan.
5. Tenaga usaha jasa.
6. Tenaga usaha pertanian, peternakan, kehutanan, nelayan, dan pemburu.
7. Tenaga produksi dan terkait.
8. Tenaga operasional.
9. Pekerja kasar.
10. Lain-lain.

### **3.5. Analisis Deskriptif**

Analisis deskriptif dapat diartikan sebagai prosedur pemecahan masalah yang diselidiki dengan menggambarkan keadaan subjek atau objek dalam penelitian dapat berupa orang, lembaga, masyarakat dan yang lainnya yang pada saat sekarang berdasarkan fakta-fakta yang tampak atau apa adanya.

Menurut Nazir (1988), metode deskriptif merupakan suatu metode dalam meneliti status sekelompok manusia, suatu objek, suatu kondisi, dan suatu sistem pemikiran ataupun suatu kelas peristiwa pada masa sekarang. Tujuan dari penelitian deskriptif ini adalah untuk membuat deskripsi, gambaran, atau lukisan secara sistematis, faktual dan akurat mengenai fakta-fakta, sifat-sifat serta hubungan antarfenomena yang diselidiki.

Menurut Sugiyono (2005) menyatakan bahwa metode deskriptif adalah suatu metode yang digunakan untuk menggambarkan atau menganalisis suatu hasil penelitian tetapi tidak digunakan untuk membuat kesimpulan yang lebih luas. Sedangkan menurut Whitney (1960) metode deskriptif adalah pencarian fakta

dengan interpretasi yang tepat. Sehingga dapat dikatakan bahwa penelitian deskriptif merupakan penelitian yang berusaha mendeskripsikan suatu gejala, peristiwa yang terjadi pada saat sekarang atau masalah aktual.

Terdapat ciri-ciri yang pokok pada metode deskriptif, antara lain adalah:

1. Memusatkan perhatian pada permasalahan yang ada pada saat penelitian dilakukan atau permasalahan yang bersifat aktual.
2. Menggambarkan fakta tentang permasalahan yang diselidiki sebagaimana adanya, diiringi dengan interpretasi rasional yang seimbang.
3. Pekerjaan peneliti bukan saja memberika gambaran terhadap fenomena-fenomena, tetapi juga menerangkan hubungan, menguji hipotesis, membuat prediksi, serta mendapatkan makna dan implikasi dari suatu masalah.

#### **3.5.1. Metode-metode penyajian analisis deskriptif**

Pada umumnya, terdapat tiga metode yang biasa digunakan untuk menjelaskan karakteristik suatu data, yaitu:

1. Tabel.

Penyajian data dalam bentuk tabel bertujuan untuk mengelompokkan nilai nilai pengamatan ke dalam beberapa kelompok yang masing-masing mempunyai karakteristik yang sama. Bentuk tabel yang sering digunakan adalah tabel distribusi frekuensi, tabel distribusi frekuensi relatif dan tabel kontingensi untuk data kualitatif dengan banyak kategori dalam baris maupun kolom.

2. Grafik atau Diagram.

Penyajian data dalam bentuk grafik atau diagram bertujuan untuk memvisualisasikan data secara keseluruhan dengan menonjolkan karakteristik-karakteristik tertentu dari data tersebut. Jenis grafik atau

diagram yang sering digunakan diantaranya adalah histogram, diagram batang dan daun, diagram garis, diagram lingkaran dan diagram kotak.

### 3. Statistik sampel.

Statistik sampel digunakan untuk menjelaskan pemusatan dan penyebaran nilai-nilai pengamatan dari suatu set data.

Menurut Howell (1982), Ukuran pemusatan atau ukuran lokasi adalah beberapa ukuran yang menyatakan dimana distribusi data tersebut terpusat. Ukuran pemusatan berupa nilai tunggal yang bisa mewakili suatu kumpulan data dan karakteristiknya (menunjukkan pusat dari nilai data). Jenis-jenis ukuran pemusatan antara lain:

1. Rata-rata (*mean*). *Mean* merupakan ukuran pemusatan yang sering digunakan. Keuntungan dari menghitung rata-rata adalah angka tersebut dapat digunakan sebagai gambaran atau wakil dari data yang diamati. Perhitungan dalam mencari nilai rata-rata adalah sebagai berikut:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

dengan,

$\bar{x}$  : rata-rata

$x$  : data ke1,2, dst

$n$  : jumlah sampel

2. Nilai tengah (*median*). Median merupakan suatu nilai ukuran pemusatan yang menempati posisi tengah setelah data diurutkan. Jika data ganjil, data yang ditengah-tengah adalah mediannya. Sedangkan jika data genap, rata-rata kedua pengamatan yang ditengah adalah mediannya. Median selalu ditentukan dengan membagi kumpulan data menjadi dua bagian fraksi yang sama.
3. Nilai yang sering muncul (*modus*). Modus merupakan sekumpulan pengamatan data yang nilai terjadinya sering muncul atau yang mempunyai frekuensi paling tinggi didalam satu kelompok nilai. Didalam

kasus suatu distribusi tidak selalu memiliki modus. Hal ini bila didalam pengamatan hanya mempunyai satu frekuensi saja, sehingga hanya bermanfaat untuk dijadikan penciri kumpulan data yang berukuran besar.

### 3.6. Analisis Cluster

Menurut Joseph, dkk (1998) Analisis *cluster* adalah teknik multivariat yang tujuan utamanya adalah untuk mengelompokan objek berdasarkan kesamaan sifat yang dimiliki. Analisis *cluster* mengklasifikasikan objek sehingga setiap objek sangat mirip dengan objek lain akan ada dalam satu *cluster* dengan beberapa kriteria seleksi yang telah ditentukan. *Cluster* yang dihasilkan dari objek harus menunjukkan kesamaan yang tinggi di dalam *cluster* (homogenitas) dan antar *cluster* (heterogenitas). Menurut Santoso (2002) homogenitas (kesamaan) yang tinggi antar anggota dalam satu *cluster* (*Within-cluster*), sedangkan heterogenitas (perbedaan) yang tinggi antar *cluster* yang satu dengan *cluster* yang lainnya (*between-cluster*).

Secara umum metode *clustering* dapat diklasifikasikan menjadi kategori-kategori sebagai berikut (Han dan Kamber, 2006):

1. Metode partisi. Misalkan ada sebuah basis data berisi  $n$  (objek). Metode partisi membangun  $k$  partisi pada basis data tersebut, dengan tiap partisi merepresentasikan *cluster* dan  $k \leq n$ . Partisi yang terbentuk harus memenuhi syarat yaitu setiap *cluster* harus berisi minimal satu objek dan setiap objek harus termasuk tepat satu *cluster*.
2. Metode hirarkhi, yaitu membuat sebuah dekomposisi berhirarki dari himpunan data atau objek menggunakan beberapa kriteria. Metode ini memiliki dua jenis pendekatan yaitu:
  - a. *Agglomerative*, dimulai dengan titik-titik sebagai *cluster* individu. Pada setiap tahap dilakukan penggabungan setiap pasangan titik pada *cluster* sampai hanya satu titik (atau *cluster*) yang tertinggal.

- b. *Divisive*, dimulai dengan satu *cluster* besar yang berisi semua titik data. Pada setiap langkah, dilakukan pemecahan sebuah *cluster* sampai setiap *cluster* berisi sebuah titik (atau terdapat  $k$  *cluster*).
3. Metode berdasarkan kepekatan, merupakan pendekatan yang berdasarkan pada konektivitas dan fungsi kepadatan.
  4. Metode berdasarkan grid, merupakan pendekatan yang berdasarkan pada struktur *multiple-level granularity*.
  5. Metode berdasarkan model, yaitu: sebuah model yang dihipotesis untuk tiap *cluster* dan ide dasarnya adalah untuk menemukan model yang cocok untuk tiap *cluster*.

#### 3.6.1. Tujuan *cluster*

Analisis *cluster* dilakukan untuk tujuan:

1. Menggali data atau eksplorasi data.
2. Mereduksi data menjadi kelompok data baru dengan jumlah lebih kecil atau dinyatakan dengan pengkelasan (klasifikasi) data.
3. Menggeneralisasi suatu populasi untuk memperoleh suatu hipotesis dan menduga karakteristik data-data.

Menurut Joseph, dkk (1998) Tujuan utama dari analisis *cluster* adalah untuk membagi satu set objek menjadi dua atau lebih kelompok berdasarkan kesamaan obyek untuk satu set karakteristik tertentu (*cluster variate*). Dalam membentuk kelompok homogen, peneliti dapat mencapai salah satu dari tiga tujuan:

1. Deskripsi Taksonomi. Penggunaan tradisional analisis *cluster* sudah sebagian besar dilakukan untuk tujuan eksplorasi dan pembentukan taksonomi berdasarkan klasifikasi objek. Seperti dijelaskan sebelumnya, analisis *cluster* telah digunakan dalam berbagai aplikasi karena kemampuannya partisinya. Namun, walaupun dilihat terutama sebagai

teknik eksplorasi, analisis *cluster* dapat digunakan untuk tujuan konfirmasi. Jika struktur yang diusulkan dapat menentukan untuk satu set objek, analisis *cluster* dapat diterapkan, dan mengusulkan tipologi (secara teoritis berdasarkan klasifikasi) dapat dibandingkan dengan yang berasal dari analisis *cluster*.

2. Penyederhanaan Data. Dalam taksonomi, analisis *cluster* mencapai perspektif yang disederhanakan dari pengamatan. Dengan struktur yang didefinisikan, pengamatan dapat dikelompokkan untuk analisa lebih lanjut. Sedangkan analisis faktor mencoba untuk memberikan "dimensi" atau struktur untuk variabel, analisis *cluster* melakukan tugas yang sama untuk pengamatan. Jadi, analisis *cluster* melihat semua pengamatan yang unik, semua objek dapat dilihat sebagai anggota *cluster* dan diprofilkan oleh karakteristik umum.
3. Identifikasi Hubungan. Dengan mendefinisikan *cluster* dan struktur yang mendasari data yang mempresentasikan *cluster*, peneliti memiliki sarana untuk mengungkapkan hubungan antara pengamatan yang mungkin tidak mungkin dengan pengamatan individu. Apakah analisis seperti analisis diskriminan digunakan untuk mengidentifikasi hubungan secara empiris, atau kelompok yang mengalami metode yang lebih kualitatif, struktur disederhanakan dari analisis *cluster* akan menggambarkan hubungan atau persamaan dan perbedaan yang sebelumnya tidak terungkap.

### **3.6.2. Memilih variabel pengelompokan**

Menurut Supranto dalam Pratama (2015) Tujuan utama analisis kelompok adalah mengklasifikasi objek seperti orang, produk, toko, perusahaan ke dalam kelompok-kelompok yang relatif homogen didasarkan pada suatu set variabel yang dipertimbangkan untuk diteliti.

Salah satu cara menentukan jumlah kelompok adalah dengan visualisasi, digunakan karena kemudahan dan kejelasannya. Contoh visual terkadang digunakan untuk menggambarkan ilustrasi jalannya algoritma untuk ekspektasi hasil *clustering*. Salah satu cara dalam mengestimasi jumlah *cluster* dalam suatu set adalah sebagai berikut:

1. *Within Cluster Sum of Squares (WCSS)* adalah Jumlah kuadrat di dalam gerombol. Secara umum jika jumlah kuadrat di dalam gerombol itu kecil, maka biasanya jumlah objek yang berada di dalam gerombol juga kecil. *Within Cluster Sum of Squares (WCSS)* juga mampu membantu penentuan berapa kelompok yang ingin dibuat oleh peneliti (Turner Lynn I, 2014).
2. Standarisasi data dilakukan jika variabel pengamatan memiliki satuan yang berbeda. Proses standarisasi data dalam analisis *cluster* ada dua yaitu standarisasi berdasarkan variabel dan standarisasi berdasarkan observasi. Rumus standarisasi data dapat dicari dengan menggunakan persamaan dibawah ini:

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (\text{Johnson dan Wichern, 2007})$$

dengan:

$Z_{ij}$ : Data hasil standarisasi observasi ke-i variabel ke-j.

$X_{ij}$ : Observasi ke-i variabel ke-j.

$\bar{X}_j$ : Rata-rata variabel ke-j.

$S_j$ : Simpangan baku variabel ke-j.

### 3.6.3. Asumsi-asumsi analisis cluster

Menurut Joseph, dkk (1998) Analisis *cluster* adalah metodologi objektif untuk mengukur karakteristik struktural dari serangkaian observasi. Dengan demikian, analisis *cluster* memiliki sifat matematika yang kuat. Persyaratan normalitas, linieritas, dan homoskedastisitas sangat penting dalam teknik lainnya,

kecuali pada analisis *cluster*. Pada analisis *cluster*, peneliti harus memperhatikan asumsi-asumsi sebagai berikut:

1. Uji *outlier*. *Outlier* adalah merupakan data yang mempunyai nilai jauh di atas atau jauh dibawah rata-rata suatu data (Singgih, 2010). Metode yang digunakan harus berdistribusi normal, jika data tidak berdistribusi normal maka hasil analisis dikhawatirkan menjadi bias sedangkan untuk menentukan data tersebut outlier atau tidak maka perlu diuji yang kemudian dilakukan pemeriksaan dari data yang telah diuji. Ketentuannya  $-2,5 > X > 2,5$  bila melewati dari batas yang telah ditentukan maka dapat dikatakan data mengalami *outlier*. Tujuan uji *outlier* adalah untuk melihat bias pencilan dari nilai-nilai yang ekstrim. Jika data diketahui terdapat satu atau lebih data *outlier*, pada *outlier* tersebut dapat dilakukan penanganan sebagai berikut (Yulianto dan Hidayatullah, 2014):
  - a. Data *outlier* dihilangkan, jika data *outlier* tersebut didapat karena kesalahan pengambilan data, kesalahan *inputing* pada komputer, dan sebagainya.
  - b. Data *outlier* tetap dipertahankan dan tidak perlu dihilangkan, jika tidak terdapat kesalahan pada proses *sampling* maupun *inputing* data.
2. Uji *kaiser meyer olkin* (KMO). Uji KMO bertujuan untuk mengetahui apakah semua data yang telah terambil telah cukup untuk difaktorkan. Hipotesis dari KMO adalah sebagai berikut:

Hipotesis

$H_0$  : Jumlah data cukup untuk difaktorkan

$H_1$  : Jumlah data tidak cukup untuk difaktorkan

Statistik uji:



$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2}$$

dengan:

$i = 1, 2, 3, \dots, p$  dan  $j = 1, 2, \dots, p$

$r_{ij}$  = Koefisien korelasi antara variabel  $i$  dan  $j$

$a_{ij}$  = Koefisien korelasi parsial antara variabel  $i$  dan  $j$

Apabila nilai KMO lebih besar dari 0,5 maka gagal tolak  $H_0$  sehingga dapat disimpulkan jumlah data telah cukup difaktorkan (Rencher, 2002).

3. Uji *bartlett sphericity*. Uji *bartlett* ini bertujuan untuk mengetahui korelasi antar variabel prediktor. Variabel  $X_i$  bersifat saling bebas (*independent*), jika matriks korelasi antar variabel membentuk matriks identitas. Untuk menguji kebebasan antar variabel ini dapat dilakukan uji *bartlett sphericity*. Berikut hipotesisnya:

$H_0$  :  $R = 1$  (ada korelasi antar variabel prediktor)

$H_1$  :  $R \neq 1$  (tidak ada korelasi antar variabel prediktor)

Statiistik uji :

$$\chi_{hitung}^2 = - \left\{ n - 1 - \frac{2p + 5}{6} \right\} \ln |R|$$

Daerah kritis : tolak  $H_0$  jika  $\chi_{hitung}^2 > \chi_{\frac{1}{2}p(p-1)}^2$

Apabila didapatkan gagal tolak  $H_0$  artinya antar variabel bersifat saling bebas. Jika hipotesis ini yang diterima maka penggunaan metode *multivariate* tidak layak terutama metode analisis komponen utama dan analisis faktor (Morrison, 2005).

### 3.7. Analisis K-Means

*K-Means* merupakan salah satu metode pengelompokan data *non-hierarki* yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok (Supranto, 2004).

Analisis *K-Means* memiliki keuntungan sebagai berikut:

1. Dalam implementasi menyelesaikan masalah, algoritma *K-Means* sangat *simple* serta fleksibel. Artinya perhitungan komputasinya tidak terlalu rumit dan analisis ini dapat diimplementasikan pada segala bidang.
2. Analisis *K-Means* sangat mudah untuk dipahami, terutama dalam implementasi data yang sangat besar serta dapat mengurangi kompleksitas data yang dimiliki (Bangoria, 2013).

Sedangkan kelemahan yang dimiliki oleh analisis *K-Means* sebagai berikut:

1. Analisis *K-Means user* memerlukan angka yang tepat dalam menentukan jumlah *cluster* sebanyak  $k$  karena terkadang pusat *cluster* awal dapat berubah sehingga kejadian ini bisa mengakibatkan pengelompokan data menjadi tidak stabil (Joshi dan Nalwade, 2013).
2. Analisis *K-Means* tidak bisa maksimal dalam menentukan atau menginisialkan nilai *centroid* awalnya, karena pada pengelompokan data dengan algoritma *K-Means* sangat bergantung pada nilai *centroid*nya (Ahmed dan Ashour, 2011).
3. *Output* dari *K-Means* tergantung pada nilai – nilai pusat yang dipilih pada *clustering*. Sehingga pada algoritma ini nilai awal titik pusat *cluster*

menjadi dasar dalam penentuan *cluster*. Pemilihan *centroid cluster* awal secara acak akan memberikan pengaruh terhadap kinerja *cluster* tersebut (Singh dan Kaur, 2013).

Konsep dasar metode *K-Means* menurut Efrain dalam Reza Pratama (2015) adalah pencarian pusat kelompok, langkah-langkah pengelompokan *K-Means* adalah sebagai berikut:

- a. Menentukan pusat kelompok secara acak.
- b. Kemudian dihitung jarak antara setiap objek dengan setiap pusat kelompok. Untuk melakukan penghitungan jarak objek ke-*i* pada pusat kelompok ke-*k* dapat digunakan rumus jarak *euclidean*. Jarak *euclidean* ialah akar dari jumlah kuadrat perbedaan atau deviasi di dalam nilai untuk setiap variabel (Johanes Supranto, 2004). Jarak *euclidean* juga biasa disebut sebagai metode perhitungan jarak yang didasarkan pada ruang berdimensi terbatas bernilai riil (Archana Kumari dan Bhagat, 2013). Adapun persamaan untuk menghitung jarak *euclidean* adalah sebagai berikut (Eko Prasetyo, 2014):

$$d_{ik} = \sqrt{\sum_{j=1}^m (x_{ij} - c_{kj})^2}$$

dengan:

$d_{ik}$ : jarak objek ke-*i* pada pusat kelompok ke-*k*.

$x_{ij}$ : nilai objek ke-*i* pada variabel *j*.

$c_{kj}$ : pusat kelompok ke-*k* pada variabel *j*

*m* : jumlah variabel yang digunakan

*i* : menyatakan objek (*i* = 1,2,3.....,n)

*k* : menyatakan kelompok ke-*k* (*k* = 1,2,3....nk)

*j* : menyatakan variabel penelitian (*j* = 1,2,3.....,m)

- c. Suatu objek akan menjadi anggota dari kelompok ke- $k$  apabila jarak objek tersebut ke pusat kelompok ke- $k$  bernilai paling kecil jika dibandingkan dengan jarak ke pusat kelompok lainnya.
- d. Selanjutnya, kelompokan objek-objek yang menjadi anggota pada setiap kelompok.
- e. Menentukan nilai pusat kelompok yang baru dapat dihitung dengan cara mencari nilai rata-rata dari objek yang menjadi anggota pada kelompok tersebut, dengan rumus sebagai berikut:

$$C_{kj} = \frac{\sum_{i=1}^n x_{ij}}{p}; x_{ij} \in \text{cluster ke-}k$$

dengan:

$C_{kj}$  : pusat kelompok ke- $k$  pada variabel  $j$ .

$x_{ij}$  : nilai objek ke- $i$  pada variabel  $j$  yang masuk dalam *cluster* ke- $k$ .

$p$  : jumlah anggota kelompok terbentuk ke- $k$ .

- f. Ulangi langkah 2 sampai 5 hingga sudah tidak ada lagi objek yang berpindah ke kelompok yang lain.

### 3.8. Peta dan Pemetaan

Pemetaan adalah proses pengukuran, perhitungan dan penggambaran permukaan bumi dengan menggunakan cara atau metode tertentu sehingga didapatkan hasil berupa *softcopy* atau *hardcopy*. Sedangkan peta adalah suatu penyajian atau gambaran unsur-unsur kenampakan nyata yang dipilih dipermukaan bumi yang digambarkan dalam bidang datar dan diperkecil dengan skala (ICA: *International Cartograp Assocation*).

Menurut Suryadi (Eko Prasetyo, 2012), peta merupakan penyajian secara grafis dari kumpulan data yang mentah maupun yang telah dianalisis atau informasi sesuai lokasinya. Dengan kata lain peta adalah bentuk sajian informasi

spasial mengenai permukaan bumi untuk dapat dipergunakan dalam pembuatan keputusan. Agar dapat bermanfaat, suatu peta harus dapat menampilkan informasi secara jelas, mengandung ketelitian yang tinggi, walaupun tidak dihindari harus bersifat selektif, dengan mengalami pengolahan, biasanya terlebih dahulu ditambah dengan ilmu pengetahuan agar lebih dapat dimanfaatkan langsung oleh pengguna.

