

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis multivariat dalam statistika merupakan analisis yang digunakan untuk analisis data dengan peubah yang banyak. Beberapa analisis yang dapat digunakan untuk data multivariat diantaranya adalah analisis *cluster*, analisis diskriminan, analisis komponen utama dan analisis faktor.

Skripsi ini membahas mengenai pengembangan analisis *cluster* dalam analisis multivariat. Analisis *cluster* pertama kali digunakan oleh Tyron pada tahun 1939. Analisis *cluster* bertujuan untuk mengalokasikan sekelompok objek pada suatu kelompok-kelompok yang saling bebas sehingga objek-objek di dalam satu kelompok yang mirip satu sama lain. Dalam pengelompokannya digunakan suatu ukuran yang dapat menerangkan kemiripan atau kedekatan antar data, yaitu ukuran jarak atau similaritas. Ukuran jarak yang sering digunakan adalah ukuran jarak yang disebut jarak *Euclidean* [15].

Dalam analisis *cluster* asumsi yang harus diperhatikan adalah data bebas dari pencilan dan tidak ada kolinieritas. Dalam melakukan pemilihan objek ke dalam kelompok-kelompok, analisis *cluster* sensitif terhadap pencilan. *Cluster-cluster* yang diperoleh tidak sesuai dengan struktur yang sebenarnya dari populasi jika pencilan dilibatkan dalam pengolahan data. Sedangkan jika terdapat kolinieritas antar variabel sebelum dilakukan analisis *cluster*, data awal terlebih dahulu ditransformasi melalui teknik komponen utama menjadi Z-Score.

Penggunaan analisis *cluster* dapat ditemukan dalam berbagai bidang antara lain pemasaran, asuransi, tata kota, pendidikan, psikologi, bahasa dan lain-lain. Dalam bidang pemasaran analisis *cluster* dapat digunakan untuk membentuk pihak pemasaran menentukan *cluster* khusus dan membuat program khusus untuk *cluster* ini. Dalam bidang tata kota, analisis *cluster* dapat digunakan untuk mengidentifikasi rumah-rumah berdasarkan tipe, harga, dan lokasi. Dalam pendidikan analisis ini dapat digunakan untuk mengelompokan data mahasiswa. Misalnya data berupa siswa,

orang tua, jenis kelamin, atau nilai IPK. Sehingga dapat dikelompokkan mahasiswa-mahasiswa yang memiliki nilai IPK tinggi, sedang dan rendah.

Pendidikan dinilai sangat penting dalam mendukung pertumbuhan ekonomi suatu negara [26]. Beberapa peneliti terdahulu banyak yang telah membuktikan pentingnya pendidikan bagi masyarakat di antaranya berkaitan dengan pengurangan kemiskinan, peningkatan status kesehatan, peningkatan partisipasi sosial dan politik, dan sebagainya [5]. Becker (1964) dalam studinya *Human Capital*, bahkan menyebutkan bahwa pendidikan menjadi salah satu bentuk investasi. Pendidikan dikatakan dapat mempengaruhi pendapatan seseorang dimasa yang akan datang dan dipercaya menjadi kunci dari suatu pembangunan ekonomi [20]. Pendidikan memiliki peran dalam menanggulangi kemiskinan melalui peningkatan pendapatan, peningkatan kesehatan dan nutrisi [26].

Angka partisipasi dalam suatu kegiatan penting diketahui, dengan mengetahui angka partisipasi tersebut dapat dinilai apakah kegiatan tersebut disukai masyarakat atau tidak disukai. Semakin besar angka partisipasi suatu program pendidikan berarti, program, lembaga, daerah tersebut berkualitas, sebaliknya kurang dan peserta banyak berhenti dalam proses pelaksanaan program berarti program, lembaga dan daerah tersebut tidak berkualitas. Angka partisipasi sekolah terdiri dari 2 jenis pengukuran yaitu angka partisipasi murni (APM) dan angka partisipasi kasar (APK). Perbedaan dari kedua hal tersebut yaitu jika APM mengukur perbandingan antara proporsi penduduk pada kelompok umur jenjang pendidikan tertentu yang masih bersekolah pada jenjang pendidikan yang sesuai dengan kelompok umurnya terhadap penduduk pada kelompok umur tersebut. Sedangkan APK mengukur proporsi penduduk pada kelompok umur jenjang pendidikan tertentu yang masih bersekolah terhadap penduduk pada kelompok umur tertentu [3]. Dengan demikian, data APK SD merupakan data yang menghitung jumlah penduduk yang berpartisipasi tidak hanya di sekolah dasar (SD) dan madrasah ibtida'iyah (MI), namun juga yang berpartisipasi pada paket A yang muridnya dapat berusia di atas usia sekolah dasar.

Skripsi ini dilakukan untuk mengelompokkan kota/kabupaten di Jawa Barat menjadi kelompok-kelompok sehingga mudah diamati untuk menentuka

kota/kabupaten mana yang memiliki APK yang tinggi dan yang rendah. Dan juga dapat melihat seberapa jauh perbedaan APK pada kelompok-kelompok tersebut. Penelitian ini dilakukan dengan mengelompokkan objek-objek terhadap variabel-variabel angka partisipasi kasar SD , SMP, SMA dan perguruan tinggi.

Berdasarkan uraian di atas, penulis, tertarik untuk mengkaji lebih dalam dan milih tema skripsi ini dengan judul “ **Analisis *Cluster* Data Panel Multivariabel dengan Metode Ward (Studi : Data Angka Partisipasi Kasar (APK) Pendidikan Jawa Barat Tahun 2012-2014)**”

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas beberapa masalah di bahas dalam skripsi ini meliputi :

- 4 Bagaimana langkah analisis *cluster* data panel multivariabel ?
- 5 Bagaimana penerapan analisis *cluster* pada data panel multivariabel pada data APK di Jawa Barat?

1.3 Tujuan Penelitian

Tujuan dari penulisan skripsi ini adalah :

1. Menentukan tahapan-tahapan analisis yang dilakukan dalam analisis *cluster* untuk data panel multivariabel.
2. Menerapkan penggunaan analisis *cluster* data panel multivariabel pada data APK di Jawa Barat.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan pengetahuan tentang sejauh mana penerapan teori-teori statistika khususnya analisis *cluster* dalam bidang pendidikan. Memberikan informasi tentang karakteristik *cluster-cluster* pada objek sehingga dapat menerapkan strategi pendidikan yang tepat untuk data APK di Jawa Barat.

1.5 Batasan Masalah

Pembatasan masalah dalam skripsi ini adalah :

1. Penulis hanya mengasumsikan bahwa data panel yang dihasilkan adalah data lengkap (bukan data yang hilang).
2. Penerapan analisis *cluster* dilakukan pada data angka partisipasi kasar (APK) Provinsi Jawa Barat jenjang pendidikan tahun 2011-2014.

1.6 Metodologi Penelitian

Tahapan yang digunakan dalam skripsi ini agar mencapai tujuan adalah sebagai berikut :

1. Studi pustaka, meliputi pengkajian dan pembelajaran buku-buku, jurnal yang berkaitan dengan analisis *cluster*, data panel, metode Ward dan pembahasan lain yang dikaji dalam skripsi ini.
2. Penerapan analisis *cluster* pada data angka partisipasi kasar (APK) di Jawa Barat yaitu mengelompokkan setiap kota di Jawa Barat yang memiliki APK terbaik.
3. Membuat kesimpulan dari metode yang dipakai proses pengerjaannya dibantu dengan bantuan *software* matematika.

1.7 Sistematika Penulisan

Sistematika penulisan skripsi ini dibagi menjadi 5 bab serta daftar pustaka, dimana setiap bab terdapat subbab yaitu :

Bab I : PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, tujuan penelitian dan manfaat penulisan, batasan masalah, metode penelitian dan sistematika penulisan.

Bab II : LANDASAN TEORI

Bab ini membahas mengenai teori-teori dasar yang melandasi penulisan skripsi ini yaitu : data panel, analisis multivariat, analisis *cluster*, ukuran jarak, klasifikasi prosedur pengelompokkan, metode Ward, angka patisipasi kasar (APK).

4



Bab III :ANALISIS *CLUSTER* DATA PANEL MULTIVARIABEL
DENGAN METODE WARD

Bab ini membahas dan mempelajari tentang analisis *cluster* data panel multivariabel, metode Ward dan tahapan-tahapan yang dilakukan pada analisis *cluster*.

Bab IV : STUDI KASUS

Bab ini membahas tentang penerapan analisis *cluster* data panel multivariabel pada data angka partisipasi kasar (APK) di Jawa Barat.

Bab V : PENUTUP

Bab ini berisi tentang kesimpulan dan saran dari skripsi.





uin

UNIVERSITAS ISLAM NEGERI
SUNAN GUNUNG DJATI
BANDUNG