

## BAB I PENDAHULUAN

### 1.1 Latar Belakang

Setiap instansi atau lembaga yang mengelola keuangan daerah terdapat begitu banyak data, sehingga menimbulkan kesulitan dalam hal pengelompokan data keuangan provinsi yang begitu besar. Penatausahaan data realisasi pendapatan pemerintah provinsi belum efektif dilaksanakan sesuai Peraturan Menteri Dalam Negeri Nomor 55 Tahun 2008 tentang Tata Cara Penatausahaan dan Penyusunan Laporan Pertanggungjawaban Bendahara serta Penyampaiannya, disebabkan banyaknya kendala dalam penatausahaan tersebut, diantaranya keterbatasan sumber daya manusia.

Kendala yang dihadapi oleh Lembaga Badan Pengelola Keuangan dan Aset Daerah Kantor Gubernur adalah bagaimana mengelola dan mengidentifikasi data pendapatan provinsi yang begitu banyak. Lembaga Badan Pengelola Keuangan juga kesulitan mengidentifikasi data pendapatan provinsi yang memiliki kesamaan dalam karakteristik tertentu. Namun seiring dengan perkembangan Teknologi Informasi (TI) muncul berbagai cara dan solusi untuk mengatasi dan meminimalisir kesulitan dalam mengidentifikasi data realisasi pendapatan pemerintah provinsi tersebut, diantaranya dengan menggunakan metode pada teknik Data Mining (DM), metode yang digunakan yaitu clustering, salah satunya dengan Algoritma K-Means. Metode ini dimaksudkan dapat mengelompokkan dan mengidentifikasi data Pendapatan Provinsi yang memiliki kesamaan dalam karakteristik tertentu.

Adapun penelitian terdahulu yaitu penelitian yang pertama adalah penelitian yang dilakukan oleh Firdausi dkk pada tahun 2011 tentang *Analisis Financial Distress Dengan Pendekatan Data Mining Pada Industri Manufaktur Go-Public Di Indonesia* hasilnya adalah Nilai SSE terkecil pada KMeans Clustering menunjukkan bahwa total kesalahan kuadrat yang terjadi pada pengelompokan metode tersebut kecil sehingga metode tersebut dapat dikatakan memiliki nilai error terkecil dan lebih baik dibandingkan metode Fuzzy C-Means. Penelitian yang kedua adalah penelitian yang dilakukan oleh Wahyuni pada tahun 2009 dalam jurnalnya yang berjudul *Penggunaan Cluster-Based Sampling Untuk Penggalan*

*Kaidah Asosiasi Multi Obyektif* hasilnya adalah Pada nilai minimum support 20 dan 40 nilai rata-rata confidence dari kaidah yang dicluster menggunakan K-Means lebih tinggi daripada Fuzzy C-Means untuk masing-masing cluster. Penelitian yang ketiga adalah penelitian yang dilakukan oleh Hamimi pada tahun 2014 dalam jurnalnya yang berjudul *Analisis Data Anggaran Pendapatan Belanja Daerah Menggunakan Clustering K-Means dan Forecasting* hasilnya adalah pengelompokan data APBD (2009-2013) menggunakan *clustering k-means* membentuk 3 *cluster*. Penelitian yang ke empat adalah penelitian yang dilakukan oleh Mayang Sari pada tahun 2015 dalam jurnalnya yang berjudul *Prediksi Data Anggaran Pendapatan Belanja Daerah Menggunakan Algoritma K-Means* hasilnya adalah dari 10 jenis belanja didapat 6 belanja yang besar. Penelitian yang ke lima yang dilakukan oleh Prasetyo pada tahun 2017 dalam jurnalnya yang berjudul *Penerapan Algoritma K-Means Untuk Cluster Data Anggaran Pendapatan Dan Belanja Desa ( Apbdes )* hasilnya adalah Semakin kecil batasan data yang digunakan dalam proses perhitungan dengan algoritma K-Means Clustering dapat diproses dengan lebih cepat.

Berdasarkan permasalahan tersebut maka penulis mengusulkan sebuah penulisan yang berjudul “Penerapan *Data Mining* Pada Data Anggaran Pendapatan Belanja Daerah Menggunakan *Clustering K-Means*”.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang maka dapat dirumuskan sebagai berikut:

1. Bagaimana hasil range pendapatan APBD Provinsi pada setiap cluster pada tahun 2014 sampai dengan 2017 ?
2. Bagaimana tingkat keakuratan dalam menentukan partisi menggunakan SSE (Sum Squared of Error) pada data Anggaran Pendapatan Belanja Daerah?

## **1.3 Batasan Masalah**

Untuk lebih memfokuskan pengerjaan penelitian ditetapkan pembahasan yang hanya dibatasi pada:

1. Dataset yang digunakan adalah data pendapatan tahun 2014 - 2017.

2. Atribut yang digunakan dalam dataset ini adalah data realisasi pendapatan pemerintah provinsi seluruh Indonesia yang meliputi Pendapatan Asli Daerah, Dana Perimbangan, dan Lain-lain Pendapatan yang Sah.
3. Jumlah *cluster* yang digunakan pada penelitian ini dibatasi sampai 3 *cluster*.

#### **1.4 Tujuan**

Tujuan penelitian pada tugas akhir ini adalah:

1. Mengetahui range pendapatan APBD Provinsi pada setiap cluster pada tahun 2014 sampai dengan 2017.
2. Mengetahui tingkat keakuratan dalam menentukan partisi menggunakan SSE (Sum Squared of Error) pada data Anggaran Pendapatan Belanja Daerah.

#### **1.5 Manfaat**

1. Dapat mengelompokkan range Pendapatan APBD Provinsi dalam setiap tahun.
2. Dapat membantu kesulitan dalam hal pengelompokan data keuangan daerah yang begitu besar.
3. Dapat membantu pembaca untuk mengetahui keakuratan jumlah centroid setelah perhitungan menggunakan SSE.
4. Dapat digunakan sebagai indikator untuk menilai potensi dan kemampuan dari masing-masing provinsi dalam usaha meningkatkan kesejahteraan masyarakat di wilayah masing-masing.