

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Dalam rangka menyediakan data yang dapat mendukung pembangunan di Kabupaten Banyumas, Badan Pusat Statistik (BPS) Kabupaten Banyumas memiliki peran penting dalam menyediakan data statistik yang lengkap, akurat, dan terkini. Untuk mencapai tujuan tersebut, BPS Banyumas secara rutin melakukan sensus dalam jangka waktu tertentu. Meskipun demikian, karena berbagai alasan, tidak selalu memungkinkan untuk melakukan sensus setiap saat. Sebagai alternatif, BPS melakukan survei [1], seperti Survei Potensi Desa untuk memberikan gambaran yang representatif tentang kondisi sosial ekonomi, infrastruktur, dan potensi sumber daya di tingkat desa.

Salah satu data yang diperoleh dari survei potensi desa adalah data sosial ekonomi. Data sosial ekonomi memberikan gambaran lengkap tentang kondisi masyarakat di desa-desa, termasuk tingkat pendidikan, jenis pekerjaan, tingkat pendapatan, dan akses terhadap layanan dasar seperti kesehatan dan pendidikan [2]. Di Kabupaten Banyumas, pusat pemerintahan dan pusat ekonomi terpusat di Kecamatan Purwokerto, sehingga perekonomian lebih berkembang di wilayah tersebut, sedangkan desa-desa di pinggiran Kabupaten Banyumas masih banyak yang menghadapi tantangan ekonomi [3]. Dengan data sosial ekonomi dari survei potensi desa yang akurat, desa-desa di pinggiran yang mengalami keterbelakangan dapat diidentifikasi dan diintervensi dengan program pembangunan yang tepat, sehingga dapat mengurangi kesenjangan antara wilayah pusat dan pinggiran.

Salah satu tantangan yang dihadapi dalam pelaksanaan survei potensi desa adalah kasus non-respon. Non-respon yaitu kondisi di mana responden yang dipilih untuk berpartisipasi dalam survei atau penelitian tidak memberikan respons atau tanggapan terhadap pertanyaan yang diajukan atau tidak mengisi kuesioner dengan lengkap contoh wilayah yang sering terjadi ada di wilayah terpencil. Non-respon dapat menyebabkan data yang diperoleh menjadi bias [1]. Untuk mengatasi masalah

ini, BPS Banyumas harus mempertimbangkan strategi untuk mengurangi dampak desa-desa tertinggal yang tidak tercakup dalam sampel survei. Salah satu solusi yang dapat diterapkan adalah dengan mengelompokkan desa-desa tersebut berdasarkan karakteristiknya.

Algoritma pengelompokan membantu mengidentifikasi pola-pola tersembunyi pada data berdasarkan kesamaan [4]. Dalam pengelompokan data salah satu metode yang populer adalah *K-Means*. *K-Means* adalah salah satu metode analisis kluster non-hirarki dan bersifat *unsupervised learning*. *K-Means* memiliki keunggulan dalam hal efisiensi dan kecepatannya. Metode ini relatif lebih terukur dan efisien untuk digunakan pada pengolahan objek dalam skala yang besar karena memiliki ketelitian yang optimal terhadap ukuran objek [5]. Kompleksitas komputasi *K-Means* juga relatif rendah, sehingga dapat diterapkan dengan baik pada data berskala besar.

Namun dibalik keunggulannya, Algoritma *K-Means* mempunyai beberapa kelemahan. Diantaranya adalah tahapan inisialisasi *centroid* awal dilakukan dengan cara acak. Padahal dalam penerapannya, Algoritma *K-Means* sangat bergantung pada *centroid* atau titik pusat kluster. Pemilihan *centroid* awal yang acak akan berakibat pada hasil yang tidak optimal.

Berdasarkan penelitian sebelumnya [6], kelemahan dari *K-Means* bisa diselesaikan dengan menggunakan Algoritma *Pillar*. Algoritma *Pillar* [7] terinspirasi dari proses berpikir dalam menentukan lokasi sekumpulan pilar untuk membuat rumah atau bangunan stabil. Algoritma ini mempertimbangkan penempatan pilar yang harus ditempatkan sejauh mungkin satu sama lain agar tahan terhadap distribusi tekanan atap. Algoritma *Pillar* melakukan penentuan *centroid cluster* dengan menggunakan cara pemilihan pada suatu data dengan nilai *euclidean* yang paling jauh dari titik pusat *cluster*. Pendekatan ini mampu menempatkan semua *centroid* secara terpisah sejauh mungkin di antara *centroid* awal dalam distribusi data. Oleh karena itu, penelitian ini akan menerapkan Algoritma *Pillar* pada *K-Means* untuk mengatasi masalah inisialisasi *centroid* awal [7].

Pendekatan ini diharapkan dapat meningkatkan representasi data dalam Survei Podes dan menghasilkan hasil yang lebih akurat serta dapat dipercaya.

Dengan demikian, informasi yang diperoleh dari survei potensi desa dapat digunakan secara efektif untuk mendukung upaya pemerintah dalam mencapai tujuan pembangunan nasional dan internasional, termasuk Sustainable Development Goals (SDGs).

1.2 Perumusan Masalah

Pemilihan *centroid* awal yang acak pada Algoritma *K-Means* dapat menyebabkan hasil yang tidak stabil/suboptimal, maka perlu menerapkan Algoritma *Pillar* untuk menentukan *centroid* awal sehingga hasil pengelompokan menjadi optimal.

1.3 Pertanyaan Peneliti

Pertanyaan peneliti yang dapat ditulis adalah sebagai berikut:

1. Apakah Algoritma *Pillar* mampu menemukan *centroid* awal yang dapat menghasilkan pengelompokan *K-Means* menjadi optimal?
2. Bagaimana perbandingan hasil evaluasi *K-Means* antara pemilihan *centroid* secara random dan menggunakan Algoritma *Pillar*?

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Menguji *K-Means* yang dibantu dengan Algoritma *Pillar* dalam menentukan *centroid* awal.
2. Membandingkan hasil penentuan *centroid* awal pada *K-Means* secara random dan menggunakan Algoritma *Pillar*.

1.5 Batasan Masalah

Batasan masalah yang ditentukan sebagai berikut:

1. Dataset yang dianalisis bersumber dari data Survei Potensi Desa (Podes) BPS Banyumas.
2. Dataset yang digunakan yaitu data Potensi Desa Kabupaten Banyumas tahun 2014-2018.
3. Variabel yang digunakan dalam penelitian ini sebanyak 10.

1.6 Manfaat Penelitian

Manfaat dari penelitian yang akan dilakukan adalah sebagai berikut:

1. Dapat memberikan kontribusi pada pengembangan pengetahuan yaitu alternatif inisialisasi *centroid* awal melalui Algoritma *Pillar* untuk meningkatkan stabilitas dan optimalitas hasil analisis *clustering K-means*.
2. Penyediaan panduan praktis untuk BPS Banyumas dalam memilih strategi analisis pengelompokan desa berdasarkan karakteristiknya yang efisien.
3. Mendukung pencapaian Sustainable Development Goals (SDGs), khususnya SDG 10 - *Reduced Inequalities* dengan mengelompokkan desa-desa berdasarkan kondisi sosial-ekonomi.