

BAB I

PENDAHULUAN

1.1 Latar Belakang

Lingkungan memiliki peranan yang sangat penting dalam mewujudkan derajat kesehatan masyarakat yang optimal, disamping faktor-faktor lain seperti kualitas pelayanan kesehatan dan perilaku masyarakat. Kesehatan lingkungan merupakan bagian dari pada kesehatan masyarakat pada umumnya. Menyadari pentingnya kesehatan lingkungan diperlukan program-program penyehatan lingkungan yang tujuannya adalah untuk membina dan mempercepat terwujudnya derajat kesehatan masyarakat yang optimal, baik fisik, mental, maupun sosial.

Program-program kesehatan lingkungan di suatu daerah biasanya direncanakan sesuai dengan kondisi dari indikator-indikator kesehatan lingkungan di daerah tersebut. Beberapa program penyehatan lingkungan yang dilakukan berupa penyehatan air dan sanitasi dasar, penyehatan permukiman dan tempat-tempat umum, penyehatan kawasan dan sanitasi darurat, higienis sanitasi pangan dan pengamanan limbah udara dan radiasi melalui kegiatan teknis penyehatan, pengamanan dan pengendalian pada media air, udara, tanah, pangan, sarana bangunan dan vektor atau binatang pembawa penyakit.

Setiap daerah memiliki keadaan kesehatan lingkungan yang berbeda-beda jika dikaitkan dengan indikator kesehatan lingkungan tersebut. Oleh karena itu prioritas program penyehatan lingkungan pun berbeda pada setiap daerah. Suatu hal yang menarik untuk diketahui adalah bagaimana kesamaan/kemiripan dari masing-masing daerah tersebut berdasarkan indikator kesehatan lingkungan.

Kemiripan tersebut selanjutnya dapat dijadikan dasar untuk melakukan pengelompokan daerah –daerah tersebut, sehingga daerah yang memiliki kondisi kesehatan lingkungan yang hampir sama akan berada pada satu kelompok dan sebaliknya, daerah-daerah dengan kondisi kesehatan lingkungan yang tidak sama akan berada pada kelompok yang berbeda. Dengan adanya pengelompokan tersebut akan mempermudah pemerintah untuk menentukan prioritas bagi pembangunan kesehatan lingkungan di daerah-daerah tersebut.

Analisis statistika yang digunakan untuk mengelompokkan objek-objek adalah analisis *cluster*. Analisis *cluster* merupakan metode analisis peubah ganda untuk mengelompokkan n objek ke dalam m *cluster* ($m < n$) berdasarkan peubah-peubah karakteristiknya. Dengan analisis *cluster*, objek-objek tersebut akan dikelompokkan berdasarkan kemiripan sehingga objek-objek yang memiliki kemiripan berada di satu *cluster*, dan objek-objek yang tidak memiliki kemiripan berada pada *cluster* berbeda. Dengan terbentuknya kelompok-kelompok tersebut, identifikasi karakteristik kelompok lebih mudah dilakukan sehingga diperoleh gambaran karakteristik setiap kelompok. Hal ini akan sangat berarti pada saat berhadapan dengan data yang banyak, terlebih lagi di era *big data* ini.

Analisis *cluster* saat ini semakin berkembang pesat seiring dengan kemajuan teknologi dan informasi yang melahirkan data besar (*big data*). Banyak metode *cluster* yang telah dikembangkan oleh para pakar dan telah banyak pula diterapkan pada berbagai bidang. Selain metode konvensional, yaitu analisis *cluster* berhirarki dan tak berhirarki, salah satu metode yang hingga saat ini banyak dikembangkan adalah metode *cluster ensemble*. Metode ini diperkenalkan oleh Strehl dan Gosh pada tahun 2002 [9]. Ide dasar dari *cluster ensemble* adalah

mengombinasikan sekumpulan hasil *cluster* yang dibentuk berdasarkan metode-metode yang biasa dilakukan. Menurut [9], *cluster ensemble* dapat memberikan hasil pengklasteran yang lebih berkualitas.

Berdasarkan uraian tersebut, maka dalam tugas akhir ini akan dibahas tentang **“Penerapan Analisis Cluster Ensemble Untuk Mengelompokkan Provinsi di Indonesia Berdasarkan Indikator Kesehatan Lingkungan”**.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka yang akan menjadi masalah pada penelitian ini adalah bagaimana mengelompokkan provinsi di Indonesia berdasarkan indikator kesehatan lingkungan dengan menggunakan metode *cluster ensemble*.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah:

1. Indikator yang digunakan, adalah gugus peubah aspek kesehatan lingkungan yang terdiri dari aspek sanitasi total berbasis masyarakat (STBM), tatanan kawasan sehat, air minum, akses sanitasi layak, tempat-tempat umum (TTU) yang memenuhi syarat kesehatan, tempat pengelolaan makanan (TPM), rumah tangga yang menempati rumah layak huni dan rumah tangga kumuh.
2. Pengklasteran berdasarkan data yang diperoleh dari Kementerian Kesehatan Republik Indonesia tahun 2016, yang meliputi 34 provinsi [3].

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah menerapkan metode *cluster ensemble* untuk mengelompokkan provinsi di Indonesia berdasarkan indikator kesehatan lingkungan.

1.5 Sistematika Penulisan

Penulisan penelitian ini terdiri dari lima bab. Bab I berisi pendahuluan yang memuat latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, serta sistematika penulisan. Bab II berisi landasan teori yang menjelaskan dan menguraikan kesehatan lingkungan, analisis *cluster*, *cluster ensemble*, dan *F* Rasio. Bab III metode penelitian yang memuat sumber data, dan analisis data. Bab IV berisi hasil dan pembahasan yang memuat hasil eksplorasi data dan hasil pengelompokan dengan *cluster ensemble*. Bab V penutup yang memuat kesimpulan dari hasil yang diperoleh pada pembahasan, dan diakhiri dengan daftar pustaka.

