

## Implementasi Importance Performance Analysis dan Algoritme K-means untuk Wilayah Indonesia

(Studi kasus : Data Presentase Rumah Tangga Menurut Sumber Air Minum Bersih dan Air Minum Layak Tahun 2016)

Bayu Pratama<sup>1</sup>, Hawila Sonya Savitri<sup>2</sup>, Mega Ayu Pramusinta<sup>3</sup>, Muhammad Duhau<sup>4</sup>,  
Nurlina Tri Wulandari<sup>5</sup>, Edy Widodo<sup>6</sup>

### Abstrak

*Air minum bersih dan air minum layak merupakan infrastruktur dasar yang seharusnya memperoleh perhatian khusus dari pemerintah Indonesia, akan tetapi sebagian besar pemerintah provinsi belum menganggap sumber air bersih dan layak sebagai prioritas pembangunan di daerahnya. Berdasarkan kondisi tersebut, pada penelitian ini akan dilakukan pengelompokan provinsi di Indonesia menjadi empat bagian untuk mengetahui provinsi yang perlu mendapatkan penyediaan akses air minum bersih dan air minum layak. Analisis yang digunakan adalah Importance Performance Analysis (IPA) dan metode K-means clustering sebagai pembandingan. Setelah diperoleh pengelompokan dari masing-masing metode, dapat disimpulkan bahwa provinsi Bengkulu dan Papua perlu menjadi perhatian lebih bagi pemerintah karena merupakan provinsi dengan penggunaan air minum bersih dan layak yang rendah.*

Kata Kunci: *Air Bersih, Air Layak, Indonesia, Importance Performance Analysis, K-Means Clustering, Provinsi*

## 1. Pendahuluan

Air menjadi kebutuhan primer yang digunakan untuk kebutuhan sehari-hari, seperti minum, masak, mandi dan lain sebagainya. Dilihat dari jumlahnya yang melimpah, hanya lima persen yang dapat diminum dan sisanya adalah air laut. Menurut sumber dari Liputan 6 bahwa sebanyak 72 juta penduduk Indonesia belum mempunyai akses minum air minum layak. Permasalahan ini seharusnya menjadi perhatian khusus dari pemerintah Indonesia karena akses air minum bersih dan layak belum mencakup ke berbagai wilayah. Organisasi kesehatan WHO atau *World Health Organization* menyatakan bahwa dalam satu dekade terakhir, rata-rata perhari 50.000 orang meninggal karena penyakit yang berkaitan dengan air tak bersih (Cahyu, 2017).

Berdasarkan kondisi tersebut maka dalam penelitian ini akan dilakukan pengelompokan wilayah provinsi di Indonesia dengan menggunakan analisis untuk mengetahui provinsi yang perlu mendapatkan penanganan penyediaan akses air minum bersih yang layak. Analisis yang digunakan adalah *Importance Performance Analysis* (IPA) yang digunakan untuk mengelompokkan provinsi-provinsi di Indonesia dalam empat kuadran. Selanjutnya digunakan analisis *k-means* untuk melihat perbedaan pengelompokan yang telah dilakukan dengan analisis IPA. Sebelumnya telah dilakukan

penelitian menggunakan metode *k-means* untuk mengelompokkan kemampuan keuangan daerah Pemkab dan Pemkot di Indonesia dengan membentuk 4 *cluster* (Fajar, 2013).

Berdasarkan penjabaran di atas, maka diketahui bahwa tujuan dari penelitian ini adalah mengetahui gambaran penggunaan sumber air minum bersih dan air minum layak tertinggi serta terendah di seluruh provinsi di Indonesia. Selanjutnya akan diketahui pula pengelompokkan wilayah tiap provinsi di Indonesia berdasarkan data presentase rumah tangga menurut sumber air minum bersih dan air minum layak.

## 2. Metodologi Penelitian

### 2.1 Populasi

Populasi dari penelitian ini adalah provinsi-provinsi yang ada di Indonesia yaitu sebanyak 34 provinsi.

### 2.2 Variabel Bebas (*Independent Variable*)

Variabel bebas adalah variabel yang mempengaruhi serta menyebabkan timbulnya atau berubahnya variabel terikat. Variabel bebas yang digunakan dalam penelitian ini adalah data presentase rumah tangga berdasarkan sumber air minum bersih dan air minum layak.

### 2.3 Devinisi Operasional Penelitian

Definisi operasional penelitian ini dapat dilihat pada tabel berikut ini :

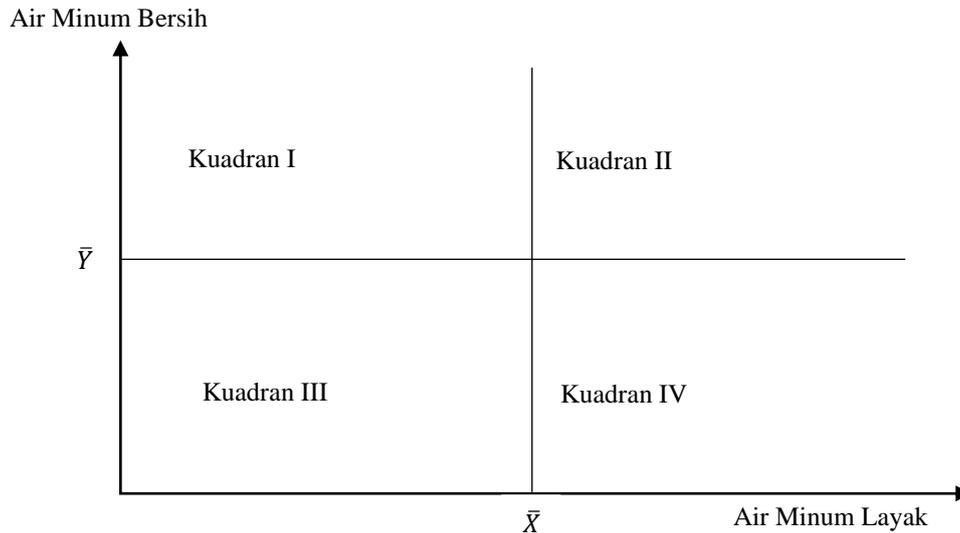
**Tabel 2.1** Definisi Operasional Penelitian

Jenis Variabel	Definisi
Presentase Rumah Tangga menggunakan Air Bersih	<p><b>Air minum berkualitas bersih</b> adalah air minum yang terlindung meliputi air ledeng (keran), keran umum, <i>hydrant</i> umum, terminal air, penampungan air hujan (PAH) atau mata air dan sumur terlindung, sumur bor atau sumur pompa, yang jaraknya minimal 10 meter dari pembuangan kotoran, penampungan limbah, dan pembuangan sampah. Termasuk air kemasan dan air isi ulang.</p> <p><b>Rumus yang digunakan :</b></p> $\frac{X_1}{N} \times 100\%$ <p>Keterangan :  <math>X_1</math> = Jumlah penduduk atau rumah tangga di wilayah tertentu dengan akses terhadap sumber air minum bersih pada periode tertentu.  <math>N</math> = Jumlah penduduk atau rumah tangga pada wilayah dan periode tertentu.</p>
Presentase Rumah Tangga menggunakan Air Layak	<p><b>Air minum yang layak</b> adalah air minum yang terlindung meliputi air ledeng (keran), air hujan, keran umum, <i>hydrant</i> umum, terminal air, penampungan air hujan (PAH) atau mata air dan sumur terlindung, sumur bor atau sumur pompa, yang jaraknya minimal 10 meter dari pembuangan kotoran, penampungan limbah dan pembuangan sampah. Tidak termasuk air kemasan.</p> <p><b>Rumus yang digunakan :</b></p> $\frac{X_2}{N} \times 100\%$ <p>Keterangan :  <math>X_2</math> = Jumlah penduduk atau rumah tangga di wilayah tertentu dengan akses terhadap sumber air minum layak pada periode tertentu.  <math>N</math> = Jumlah penduduk atau rumah tangga pada wilayah dan periode tertentu.</p>

## 2.4 Analisis

### a. Importance Performance Analysis (IPA)

Setelah mengetahui gambaran data presentase rumah tangga berdasarkan sumber air minum bersih dan layak, selanjutnya peneliti melakukan analisis IPA. Pada analisis ini peneliti akan membagi provinsi-provinsi di Indonesia menjadi empat kelompok dengan menggunakan diagram kartesius. Visualisasi dari kuadran IPA dalam bentuk diagram kartesius dapat dilihat pada **Gambar 2.1**.



**Gambar 2.1** Kuadran *Importance Performance Analysis*

Diketahui:

$\bar{X}$  = Nilai rata-rata tingkat air minum layak

$\bar{Y}$  = Nilai rata-rata tingkat air minum bersih

Diagram kartesius dibatasi oleh dua buah garis yang berpotongan tegak lurus pada titik  $\bar{X}$  dan  $\bar{Y}$ . Perpotongan tersebut membentuk empat kuadran. Kuadran I merupakan tingkat air minum bersih tinggi dan air minum layak rendah, kuadran II adalah tingkat air minum bersih tinggi dan air minum layak tinggi, kuadran III adalah tingkat air minum bersih rendah dan air minum layak rendah, dan kudran IV ialah tingkat air minum bersih rendah dan air minum layak tinggi.

### b. K-Means

Sebelum dilakukan analisis *cluster*, terlebih dahulu peneliti perlu melakukan uji asumsi agar dapat memenuhi syarat analisis *cluster*. Uji asumsi yang perlu dilakukan yaitu melakukan uji normalitas, *outlier* serta uji non-multikolinieritas. Menurut Giyanto dalam tulisan Utari (2016), *k-means* merupakan salah satu metode data *clustering* non hirarki yang mempartisi ke dalam *cluster* atau kelompok sehingga data yang memiliki karakteristik yang sama (*Low inter class similarity*) dikelompokkan ke dalam satu *cluster* yang sama dan yang memiliki karakteristik yang berbeda (*High intra class similarity*) dikelompokkan pada kelompok yang lain. Secara umum langkah-langkah dari *k-means clustering* adalah sebagai berikut :

1. Tentukan jumlah *cluster*.
2. Alokasikan data ke dalam *cluster* secara random.
3. Hitung *centroid* atau rata-rata dari data yang ada di masing-masing *cluster*, dengan menggunakan rumus :

$$v = \sum_{i=1}^n \frac{x_i}{n}; i = 1, 2, 3, \dots, n \quad (2.1)$$

4. Alokasikan masing-masing data ke *centroid* atau rata-rata terdekat.
5. Kembali ke *step* 3, apabila masih ada data yang berpindah *cluster* atau apabila perubahan nilai *centroid*, ada yang diatas nilai *threshold* yang ditentukan atau apabila perubahan nilai pada *objective function* yang digunakan diatas nilai *threshold* yang ditentukan.
6. Menghitung jarak setiap objek, *distance space* digunakan untuk menghitung jarak antara data dan *centroid*. Adapun persamaan yang dapat digunakan salah satunya yaitu *Euclidean Distance Space*. *Euclidean Distance Space* sering digunakan dalam perhitungan jarak, hal ini dikarenakan hasil yang diperoleh merupakan jarak terpendek antara dua titik yang diperhitungkan. Adapun persamaannya adalah sebagai berikut :

$$d_{ij} = \sqrt{\sum_{k=1}^p \{X_{ik} - X_{jk}\}^2} \quad (2.2)$$

7. Setelah didapat jumlah anggota dari tiap-tiap *cluster* maka dapat dilihat karakteristik dari *cluster* yang terbentuk, dengan menggunakan rumus:
 
$$x = \mu + (z \cdot \sigma) \quad (2.3)$$
 (Muliadinata, 2013)

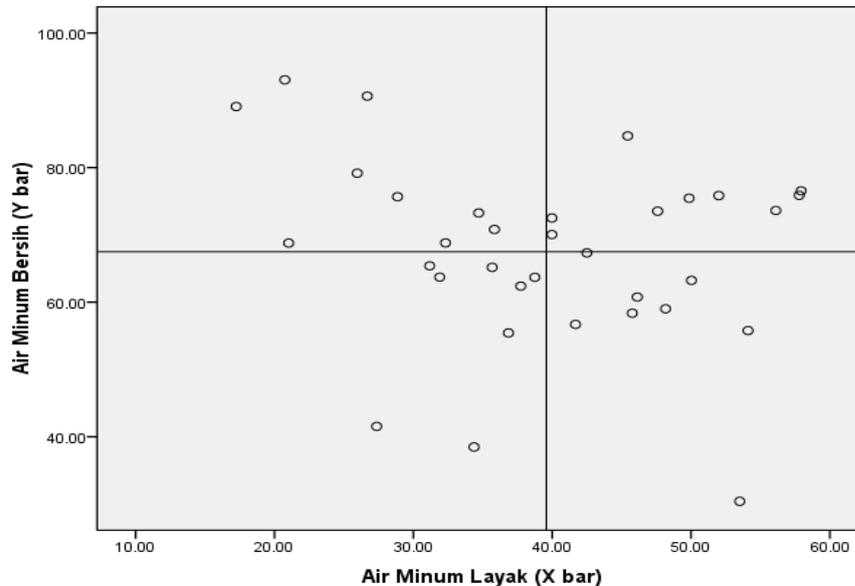
### 3. Hasil dan Pembahasan

Tabel 3.1 Summary Data

	Air Bersih		Air Layak	
	Frekuensi (%)	Provinsi	Frekuensi (%)	Provinsi
Minimum	30,39	Kalimantan Barat	17,24	Kepulauan Riau
Maksimum	93,05	DKI Jakarta	57,93	DI Yogyakarta

Data presentase rumah tangga menurut sumber air minum bersih tahun 2016 seperti pada **Tabel 3.1** memiliki nilai minimum 30,39% yaitu untuk provinsi Kalimantan Barat dan nilai maksimal sebesar 93,05%, yaitu untuk provinsi DKI Jakarta. Artinya 30 dari 100 rumah tangga di provinsi Kalimantan Barat dan 93 dari 100 rumah tangga di provinsi DKI Jakarta mengkonsumsi air minum bersih. Kemudian nilai terendah pada data presentase rumah tangga menurut sumber air layak tahun 2016 adalah 17,24% yang diperoleh provinsi Kepulauan Riau dan tertinggi adalah provinsi Daerah Istimewa Yogyakarta dengan nilai sebesar 57,93%. Hal ini membuktikan bahwa terjadi kesenjangan antar provinsi karena presentase penggunaan air minum bersih maupun air minum layak yang berbeda jauh. Oleh karena itu, diperlukan suatu pengelompokkan untuk memisahkan provinsi-provinsi dengan frekuensi penggunaan air minum yang rendah dan tinggi sehingga peneliti mudah dalam pengambilan kebijakan.

Salah satu caranya yaitu dengan menggunakan analisis IPA. Peneliti menggunakan analisis IPA, bertujuan untuk membagi provinsi-provinsi di Indonesia menjadi empat kuadran berdasarkan penggunaan air minum bersih dan air minum layak. Diketahui bahwa nilai rata-rata dari data penggunaan air minum layak tiap provinsi ( $\bar{X}$ ) adalah 39,58. Sedangkan nilai rata-rata dari data penggunaan air minum bersih tiap provinsi ( $\bar{Y}$ ) sebesar 67,49. Maka titik perpotongan kuadran adalah (39,58 : 67,49).



**Gambar 3.1** Hasil *Importance Performance Analysis*

Diketahui bahwa pada rata-rata nilai penggunaan air minum layak yaitu  $39,58 < \bar{X} < 39,58$ . Artinya pendistribusian air layak kepada masyarakat tiap provinsi akan dikatakan baik ketika nilainya telah melebihi rata-rata. Begitu pula dengan penggunaan air minum bersih, yaitu dengan nilai rata-rata nilai penggunaan air minum bersih sebesar  $67,49 < \bar{X} < 67,49$ . Sehingga didapatkan pemetaan atribut provinsi pada masing-masing kuadran adalah sebagai berikut :

a. Kuadran I

Kuadran I merupakan daerah provinsi dengan tingkat penggunaan air minum layak rendah dan air minum bersih tinggi. Provinsi-provinsi yang ada pada kuadran I yaitu Banten, Jawa Barat, Sumatera Barat, Sulawesi Utara, Kalimantan Utara, Kepulauan Bangka Belitung, Kepulauan Riau, Kalimantan Timur dan DKI Jakarta.

b. Kuadran II

Kuadran II merupakan daerah provinsi dengan tingkat penggunaan air minum layak tinggi dan air minum bersih tinggi. Provinsi-provinsi yang ada pada kuadran II diantaranya yaitu provinsi Sumatera Utara, Gorontalo, Sulawesi Selatan, Nusa Tenggara Barat, Jawa Timur, Sulawesi Tenggara, Jawa Tengah, Daerah Istimewa Yogyakarta, dan Bali.

c. Kuadran III

Kuadran III merupakan daerah provinsi dengan tingkat penggunaan air minum layak rendah dan air minum bersih rendah. Provinsi-provinsi yang ada pada kuadran III adalah Papua, Bengkulu, Lampung, Papua Barat, Sulawesi Tengah, Kalimantan Tengah, Riau, dan Aceh. Provinsi yang ada pada kuadran ini perlu perhatian khusus karena air minum bersih maupun air minum layaknya memiliki tingkat presentase rendah. Oleh karena itu, perlu dilakukan evaluasi kerja pemerintah terhadap penyaluran air bersih kepada masyarakat.

d. Kuadran IV

Kuadran IV merupakan daerah provinsi dengan tingkat penggunaan air minum layak tinggi dan air minum bersih rendah. Terdapat delapan provinsi yang ada pada kuadran IV, yaitu Kalimantan Barat, Nusa Tenggara Timur, Sulawesi Barat, Jambi, Maluku, Maluku Utara, Sumatera Selatan dan Kalimantan Selatan.

Setelah diketahui hasil pengelompokan dengan menggunakan analisis IPA, peneliti akan menggunakan metode *k-means* sebagai pembanding. Dibentuk empat *cluster* untuk mengetahui hasil pengelompokan provinsi di Indonesia berdasarkan penggunaan air minum bersih dan layak. Iterasi yang dihasilkan menggunakan metode *k-means* adalah sebanyak 2 kali iterasi. Hasil *cluster* provinsi-provinsi yang berada dalam satu kelompok berdasarkan data air minum layak dan air minum bersih adalah sebagai berikut:

**Tabel 3.4** Hasil Pengelompokan Menggunakan Analisis *Cluster*

Cluster	Banyak Anggota	Anggota
1	7	Jawa Tengah, Daerah Istimewa Yogyakarta, Jawa Timur, Bali, Nusa Tenggara Barat, Sulawesi Selatan, Sulawesi Tenggara.
2	18	Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Lampung, Jawa Barat, Nusa Tenggara Timur, Kalimantan Tengah, Kalimantan Selatan, Sulawesi Utara, Sulawesi Tengah, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat.
3	3	Bengkulu, Kalimantan Barat, Papua.
4	6	Kep. Bangka Belitung, Kep. Riau, DKI Jakarta, Banten, Kalimantan Timur, Kalimantan Utara.

Setelah didapatkan hasil pengelompokan menggunakan metode IPA dan *k-means*, peneliti dapat membedakan hasil pengelompokan wilayah pada masing-masing metode. Pengelompokan pada kuadran I menggunakan metode IPA memiliki hasil yang hampir sama dengan *cluster* keempat. Seluruh provinsi yang ada di *cluster* keempat berada pada kuadran I, namun beberapa provinsi pada kuadran I tidak ada pada *cluster* keempat. Oleh karena itu dapat dikatakan bahwa *cluster* keempat merupakan provinsi dengan penggunaan air minum layak rendah dan air minum bersih tinggi.

Selanjutnya pengelompokan wilayah pada kuadran II hampir sama dengan *cluster* pertama. Jadi, *cluster* pertama merupakan provinsi dengan tingkat penggunaan air minum layak tinggi dan air minum bersih tinggi. Pengelompokan wilayah ini adalah provinsi-provinsi dengan nilai presentase penggunaan air minum bersih maupun layak berada di atas rata-rata atau dapat dikatakan baik. Berikutnya beberapa provinsi pada kuadran IV memiliki hasil yang sama seperti *cluster* kedua. Pada *cluster* kedua ini memiliki hasil yang lebih beragam karena setiap hasil *cluster* yang tidak ada pada kuadran sebelumnya akan masuk pada *cluster* kedua. Pengelompokan wilayah ini merupakan provinsi dengan tingkat penggunaan air minum layak tinggi dan air minum bersih rendah.

Kemudian pada kuadran III terdapat dua provinsi yang sama dengan *cluster* ketiga, yaitu Bengkulu dan Papua. Pada *cluster* ketiga ini terdapat satu provinsi yang tidak berada pada kuadran III, yaitu Kalimantan Barat. Kuadran III maupun *Cluster* ketiga adalah provinsi dengan tingkat penggunaan air minum layak rendah dan air minum bersih rendah. Oleh karena itu, pada kasus ini provinsi Bengkulu dan Papua perlu menjadi perhatian lebih bagi pemerintah dalam pendistribusian air bersih maupun air layak. Hal ini dibuktikan dengan hasil analisis menggunakan IPA dan *k-means clustering* yang menunjukkan letak klasifikasi provinsi Papua dan Bengkulu berada pada tingkat penggunaan air minum bersih rendah dan air minum layak rendah. Hasil penelitian ini dapat digunakan sebagai bahan evaluasi bagi pemerintah dalam memfokuskan daerah dengan tingkat penggunaan air minum bersih dan layak yang rendah, karena hal ini akan berpengaruh kepada kesehatan masyarakat.

## 4. Kesimpulan

Presentase rumah tangga menurut sumber air minum bersih tahun 2016 memiliki nilai terendah 30.39% yaitu di provinsi Kalimantan Barat dan nilai tertinggi sebesar 93,05%, yaitu provinsi DKI Jakarta. Presentase rumah tangga menurut sumber air minum layak tahun 2016 memiliki nilai terendah sebesar 17,24% yang diperoleh provinsi Kepulauan Riau dan nilai tertinggi adalah provinsi Daerah Istimewa Yogyakarta dengan nilai sebesar 57,93%. Berikutnya pengelompokan menggunakan metode IPA dan algoritma *k-means*, memiliki hasil yang hampir sama. Setelah menggunakan kedua metode tersebut, diperoleh hasil bahwa provinsi Bengkulu dan Papua perlu menjadi perhatian lebih karena merupakan provinsi dengan tingkat penggunaan air minum bersih dan layak yang rendah.

## Daftar Pustaka

- [1] Cahyu. 2017. *Kementrian PUPR Dukung Kota Pontianak Capai Akses Aman Air Minum*. <http://news.liputan6.com/read/3144305/kementerian-pupr-dukung-kota-pontianak-capai-akses-aman-air-minum>. Diakses pada 8 November 2017.
- [2] Fajar, Muhamad. 2013. *Pemetaan Kemampuan Keuangan Daerah Pemkab dan Pemkot di Indonesia*. [https://www.academia.edu/21197964/Pemetaan\\_Kemampuan\\_Keuangan\\_Daerah\\_Pemkab\\_dan\\_Pemkot\\_di\\_Indonesia\\_Pendekatan\\_Cluster](https://www.academia.edu/21197964/Pemetaan_Kemampuan_Keuangan_Daerah_Pemkab_dan_Pemkot_di_Indonesia_Pendekatan_Cluster). Diakses pada 22 September 2017.
- [3] Muliadinata, Saban. 2013. *Algoritma K-Means Clustering*. <http://sharewy.blogspot.co.id/2013/04/algoritma-k-means-clustering.html>. Diakses pada 25 September 2017.
- [4] Utari, Prianita. 2016. *Aplikasi Analisis Cluster Dalam Mengetahui Karakteristik Pengelompokan Menggunakan Metode K-Means dan K-Medoids*. Skripsi Program Sarjana Statistika Universitas Islam Indonesia.
- [5] Wiryadi, Rudi. 2017. *Inilah Kriteria Air Bersih dan Cara Mengolahnnya*. [https://www.kompasiana.com/rudywiryadi12/kriteria-air-bersih-dan-cara-mengolahnnya\\_59520fab425ce00a1202a152](https://www.kompasiana.com/rudywiryadi12/kriteria-air-bersih-dan-cara-mengolahnnya_59520fab425ce00a1202a152). Diakses pada 8 November 2017.