

**Pengelompokan Provinsi Di Indonesia Berdasarkan Index Kesehatan Masyarakat  
Menggunakan Algoritma Partitioning Around Medoids (PAM)  
Dan Metode *Davies Bouldin Index* (DBI)**

**Groupment of Provinces in Indonesia Based on the Public Health Index Using the  
Partitioning Around Medoids (PAM) Algorithm and the Davies Bouldin Index (DBI) Method**

**Dedi Setiawan<sup>1</sup>, Agung Nilogiri<sup>2</sup>, Moh. Dasuki<sup>3</sup>**

<sup>1</sup>Mahasiswa Fakultas Teknik, Universitas Muhammadiyah Jember  
email: [sdedi900@gmail.com](mailto:sdedi900@gmail.com)

<sup>2</sup>Fakultas Teknik, Universitas Muhammadiyah Jember email:  
[agungnilogiri@unmuhjember.ac.id](mailto:agungnilogiri@unmuhjember.ac.id)

<sup>3</sup>Fakultas Teknik, Universitas Muhammadiyah Jember email:  
[moh.dasuki22@gmail.com](mailto:moh.dasuki22@gmail.com)

**Abstrak**

Rendahnya ketersediaan air minum dan sanitasi yang memadai di Indonesia disebabkan oleh disparitas antara wilayah barat dan timur Indonesia. Di Provinsi Papua, misalnya, sekitar separuh penduduknya tidak memiliki air minum yang layak dan dua dari tiga rumah tangga memiliki sanitasi yang buruk. Sementara itu, hampir semua rumah tanggadi Jakarta memiliki akses terhadap air minum yang memadai. Untuk mengatasi kesenjangan dan ketidasetaraan ini, diperlukan penelitian yang mendalam terkait indikator kesehatan lingkungan, termasuk akses air minum yang memadai, sanitasi yang memadai, dan perumahan yang layak. Dalam hal ini, diperlukan upaya untuk menentukan tingkat sanitasi di setiap wilayah atau provinsi sebagai langkah untuk mengatasi ketimpangan dalam kesehatan lingkungan. Salah satu pendekatan yang dapat digunakan adalah menggunakan metode Partitioning Around Medoids dengan menggunakan metode Davies Bouldin Index Optimization untuk menghitung cluster optimal yang dapat menentukan kelompok yang paling baik. Hasil penerapan metode Partitioning Around Medoids dalam pengelompokan provinsi di Indonesia berdasarkan indikator kesehatan lingkungan menunjukkan bahwa terdapat 5 kelompok optimal dengan nilai indeks Davies Bouldin sebesar 1,3473 dengan variasi jumlah kelompok dari 2 hingga 10.

**Keywords:** Sanitasi, Cluster, Partitioning Around Medoids, Davies Bouldin Index, Optimal

**Abstract**

The low availability of drinking water and adequate sanitation in Indonesia is caused by disparities between the western and eastern regions of Indonesia. In Papua Province, for example, about half the population does not have proper drinking water and two out of three households have poor sanitation. In contrast, almost all households in Jakarta have access to safe drinking water. Considering these disparities and inequalities, in-depth studies are needed on environmental health indicators derived from variables such as access to safe drinking water, proper sanitation, and adequate housing. Efforts are required to assess the level of sanitation in a region or province in order to address environmental health disparities. The prioritized provinces can be determined by performing clustering, one of the methods being Partitioning Around Medoids using the Davies Bouldin Index Optimization method to calculate the optimal clusters for determining the best cluster. The application of the Partitioning Around Medoids method to group provinces in Indonesia based on environmental health indicators resulted in an optimum cluster of 5 clusters with a Davies Bouldin index value of 1.3473, considering scenarios with 2 to 10 clusters.

**Keywords:** Sanitation, Cluster, Partitioning Around Medoids, Davies Bouldin Index, Optimum

## 1. PENDAHULUAN

Sanitasi merupakan sebuah bentuk upaya untuk mempromosikan dan menciptakan situasi yang kondusif di bidang kesehatan, dengan penekanan pada pola dan perilaku hidup yang bersih dan sehat, khususnya kesehatan masyarakat. Sanitasi berkaitan dengan kesehatan lingkungan yang memiliki manfaat mengurangi dampak dari malnutrisi, meningkatkan taraf khususnya kesehatan masyarakat serta menjamin ketersediaan air bersih dan terbebas dari pencemaran limbah rumah tangga dan industri. Tiga parameter diatas merupakan kebutuhan dasar manusia (Lubis, 2019)

Pemukiman berbanding lurus dengan kesehatan dimana kebutuhannya dijamin oleh konstitusi. Upaya ini semakin ditingkatkan setiap tahunnya. Komitmen pemerintah untuk memenuhi kebutuhan yang mendasar tadi juga berbanding lurus dengan agenda global dengan tujuan pembangunan berkelanjutan pada tahun 2030 (*Sustainable Development Goals*) atau disebut *SDGs*. Targetnya meju pada sektor lingkungan hidup, yaitu dipastikannya warga dapat mengakses sanitasi serta air bersih secara menyeluruh (Lubis, 2019).

Karena perbedaan dan ketimpangan antara Indonesia bagian barat dan timur, akses terhadap air minum dan sanitasi yang layak di Indonesia masih buruk. Di wilayah Papua misalnya, sekitar separuh penduduk tidak memiliki air minum yang layak dan dua dari tiga rumah tangga memiliki fasilitas sanitasi yang tidak memadai. Pada saat yang sama, hampir seluruh rumah tangga di Jakarta memiliki akses terhadap air minum yang layak. Karena adanya kesenjangan dan ketimpangan tersebut, maka perlu dilakukan penelitian mendalam tentang indikator lingkungan yang diturunkan dari variabel yang terkait dengan akses air minum yang layak, sanitasi yang layak, dan perumahan yang layak. Oleh karena itu, diperlukan penelitian lebih lanjut terhadap ketiga variabel tersebut. Ini dapat digunakan untuk mengidentifikasi provinsi membutuhkan perhatian meningkatkan indikator kesehatan lingkungan. Jadi, penting untuk memerhatikan pengklasifikasian provinsi di Indonesia berdasarkan pada indikator kesehatan lingkungan (Mayasari, 2020)

Algoritma Partitioning Around Medoids terkait dengan K-Medoids, namun ada perbedaan penting antara kedua algoritma tersebut, yaitu di K-Medoids sebuah cluster direpresentasikan oleh pusat cluster, sedangkan di K-Medoids sebuah cluster direpresentasikan. . melalui objek yang paling dekat dengan pusat cluster. Medoid adalah objek yang berada di tengah cluster dan karenanya tidak terpengaruh oleh drift. Cluster dibentuk dengan mempertimbangkan kedekatan antara objek rata-rata dan tidak rata-rata. Partitioning Around Medoids (PAM) lebih kuat terhadap noise dan outlier daripada algoritma K-Means. Ini karena medoid kurang rentan terhadap orang luar dan ekstrem lainnya daripada rata-rata. Diantara sekian banyak metode yang dapat menggunakan metode Davies-Buldin Index (DBI) adalah metodenya. Menguji metode ini melakukan beberapa kalkulasi pada cluster yang diberikan dan mengembalikan hasil untuk beberapa nilai cluster. Nilai terendah dari Davies-Bouldin Index merupakan nilai optimal (Abdurrahman dkk, 2021).

## 2. STUDI PUSTAKA

### A. Data Mining

Data Mining adalah proses yang menggunakan teknik seperti statistik, matematika, dan kecerdasan buatan untuk mengekstraksi dan mengidentifikasi informasi baru dari kumpulan data besar. Pola yang ditemukan dapat berupa aturan bisnis, kesamaan, korelasi, tren atau pola prediksi. Definisi umum dari penambangan data mendefinisikannya sebagai "proses kompleks untuk mengidentifikasi pola yang valid, baru, berpotensi bermanfaat, dan dapat dipahami dari data yang disimpan dalam basis data terstruktur". Dalam hal ini, data disusun dalam baris yang disusun berdasarkan kategori, urutan, dan variabel yang berurutan. (Albert. 2016)

### B. Clustering

Mengelompokkan data untuk menggali persamaan pada data serta menemposisikan data yang mirip dalam beberapa kelompok. *Clustering* membagi *dataset* kedalam beberapa kelompok yang mana memiliki kesamaan satu dengan yang lain daripada kelompok yang lain. Tujuan d a r i *clustering* adalah Minimalkan jarak dalam cluster, serta memaksimalkan jarak antar cluster, proses pengelompokan atau klasifikasi kelas yang memiliki objek yang sama. Oleh karena itu konsep clustering adalah kumpulan dari semua kelompok. Perbedaan antara *clustering* dan klasifikasi adalah tidak ada variabel target saat mengelompokkan data pada proses *clustering* (Abdilah, 2016)

### C. Partitioning Around Medoids (PAM)

Algoritma Partitioning Around Medoids (PAM), sering disebut K-Medoids, adalah algoritma clustering yang berhubungan dengan algoritma K Means yang dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw pada tahun 1987. Algoritma ini pada dasarnya sangat mirip dengan K-Means. karena dua algoritma umum. Dengan kata lain, kedua algoritme membagi kumpulan data menjadi beberapa kelompok dan kedua algoritme mencoba meminimalkan kesalahan. Namun, algoritma PAM (Partitioning Around Medoids) bekerja dengan medoid, entitas dalam kumpulan data yang mewakili grup untuk disertakan. Bedanya, algoritma kmedoids menggunakan objek sebagai median di tengah setiap cluster, sedangkan K-Means menggunakan mean atau median sebagai pusat cluster (Kauretal, 2014).

Kelebihan dari algoritma K-Medoids adalah mengatasi kelemahan dari algoritma K-Medoids. K-medoid sensitif terhadap noise/outlier dan outlier besar, tetapi keuntungan lainnya adalah hasil sekuens tidak bergantung padanya. dari proses akumulasi. dalam urutan di mana catatan dimasukkan . (Furqon, 2015)

Total nilai ditunjukkan dalam persamaan:

$$Total Cost = \sum_{k=1}^n \sqrt{\sum^n (x_k - y_k)^2} \dots\dots\dots(2.1) \text{ Dengan:}$$

- $n$  = jumlah dari sebuah data
- $k$  = indeks suatu data
- $x_k$  = nilai atribut ke- $k$  dari  $x$
- $y_k$  = nilai atribut ke- $k$  dari  $y$

Nilai  $S$  dinyatakan dalam persamaan:

$$S = Total Cost \text{ baru} - Total Cost \text{ Lama} \dots\dots\dots(2.2) \text{ Dengan:}$$

- $S$  = Selisih
- $Total Cost \text{ Baru}$  = jumlah  $cost \text{ non-medoids}$
- $Total Cost \text{ Lama}$  = jumlah  $cost \text{ medoids}$

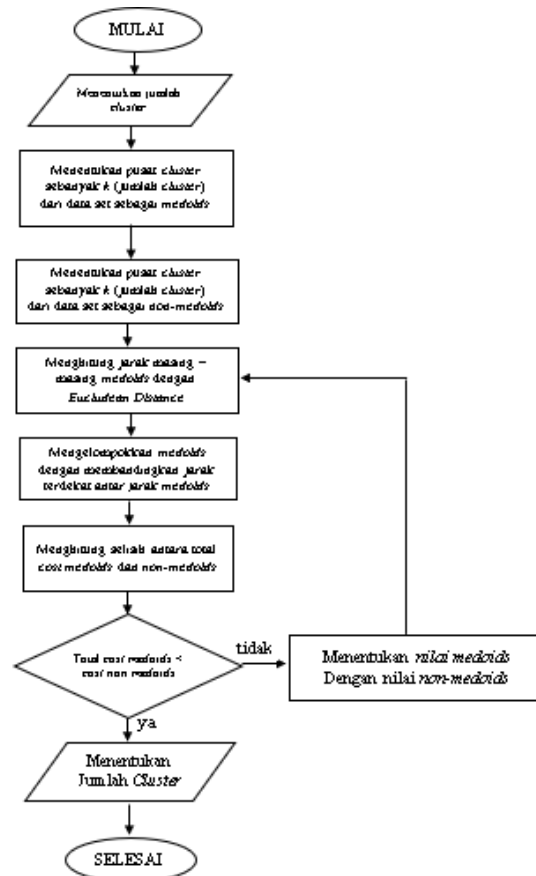
5. Apabila  $S < 0$  maka tukar  $o_j$  dengan  $o_{random}$  untuk membentuk sekumpulan  $k$  objek baru sebagai  $medoids$ .

Ulang pada langkahh 2 sehingga hasil dari langkah 5 tidak ada perubahan

## 3. METODOLOGI PENELITIAN

### A. Pengumpulan Data

Kumpulan data tersedia untuk 34 provinsi di Indonesia berdasarkan persentase rumah tangga dengan sanitasi yang layak, sumber air minum yang layak, dan rumah layak huni pada tahun 2020-2021. Data tersebut memiliki enam atribut yaitu Sanitasi Layak (SL), Air Minum Layak (AML), dan Rumah Layak Huni (RLH). Ini dapat dicapai dengan menggabungkan data menggunakan algoritma Partitioning Around Medoids (PAM) dan Davies Bouldin. Index Method (DBI) sebagai metode optimasi cluster untuk menentukan jumlah cluster optimal yang digunakan dalam proses clustering (Wiwit, 2015). Penelitian ini menggunakan metode analisis data Partitioning Around Medoids (PAM). Dimana K-Medoids merupakan teknik clustering, yaitu teknik untuk mengelompokkan sekumpulan objek menjadi satu perwakilan pusat (medoid) dari setiap cluster. Pada penelitian ini, algoritma Partitioning Around Medoids digunakan untuk mengklasifikasikan data sekunder dari 34 provinsi di Indonesia berdasarkan persentase rumah tangga yang memiliki akses sanitasi dan air minum yang layak. Clustering menggunakan teknik indeks Davies-Bouldin untuk menentukan cluster yang optimal. Metode penelitian dilakukan dengan menggunakan algoritma clustering Partitioning Around Medoids (PAM). Clustering adalah tentang efisiensi dengan menggunakan Davies-Bouldin Index dalam beberapa langkah yaitu penelitian pendahuluan, pengumpulan data dan proses clustering untuk mengidentifikasi cluster terbaik untuk sanitasi yang layak, air minum yang layak dan perumahan yang layak. Algoritma PAM (Partitioning Around Medoids) umumnya memiliki langkah-langkah operasional sebagai berikut:



Gambar 1 Flowchart Algoritma PAM

## B. Partitioning Around Medoids

Algoritma Partitioning Around Medoids (PAM), juga dikenal sebagai K-Medoids, adalah algoritma pengelompokan yang terkait dengan algoritma K-Medo yang dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw pada tahun 1987. Algoritma ini mirip dengan K-Means, terutama karena keduanya merupakan algoritma partisi. Dengan kata lain, kedua algoritma mengelompokkan data ke dalam kelompok dan keduanya bertujuan untuk meminimalkan kesalahan.

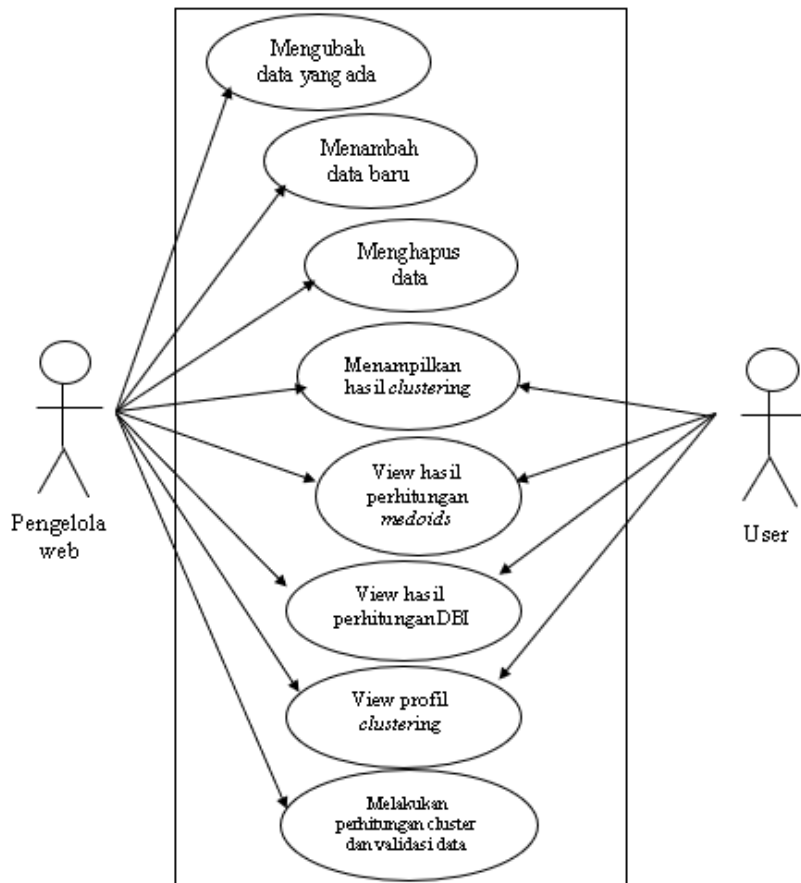
Namun, algoritma PAM (Partitioning Around Medoids) menggunakan medoids, yaitu entitas dalam kumpulan data yang mewakili kelompok yang terbentuk. Perbedaan utama adalah bahwa algoritma K-Medoids menggunakan item dalam kumpulan data sebagai median, yang merupakan pusat dari setiap grup, sedangkan K-Means menggunakan rata-rata atau median sebagai pusat grup. (Kauret al, 2014)

### 1. Unified Modelling Language

#### Use Case Diagram

Deskripsi Use Case Diagram merupakan penjelasan tentang fungsionalitas sistem dari sudut pandang pengguna. Use case diagram menggambarkan interaksi tipikal antara pengguna dan sistem menggunakan cerita yang menjelaskan bagaimana sistem digunakan. Singkatnya, diagram use case adalah sekumpulan skenario yang disatukan oleh tujuan umum pengguna.

Dalam konteks diagram use case, pengguna sering disebut sebagai aktor. Aktor adalah peran yang dapat diambil oleh pengguna saat berinteraksi dengan sistem. Use case mewakili interaksi antara sistem dan aktor, jadi memilih abstraksi yang tepat adalah penting. Use case dibuat sesuai dengan kebutuhan operator. Kasus penggunaan harus menjelaskan apa yang dilakukan perangkat lunak/aplikasi, bukan bagaimana hal itu dilakukan. Setiap use case harus diberi nama yang mencerminkan tujuan yang dicapai

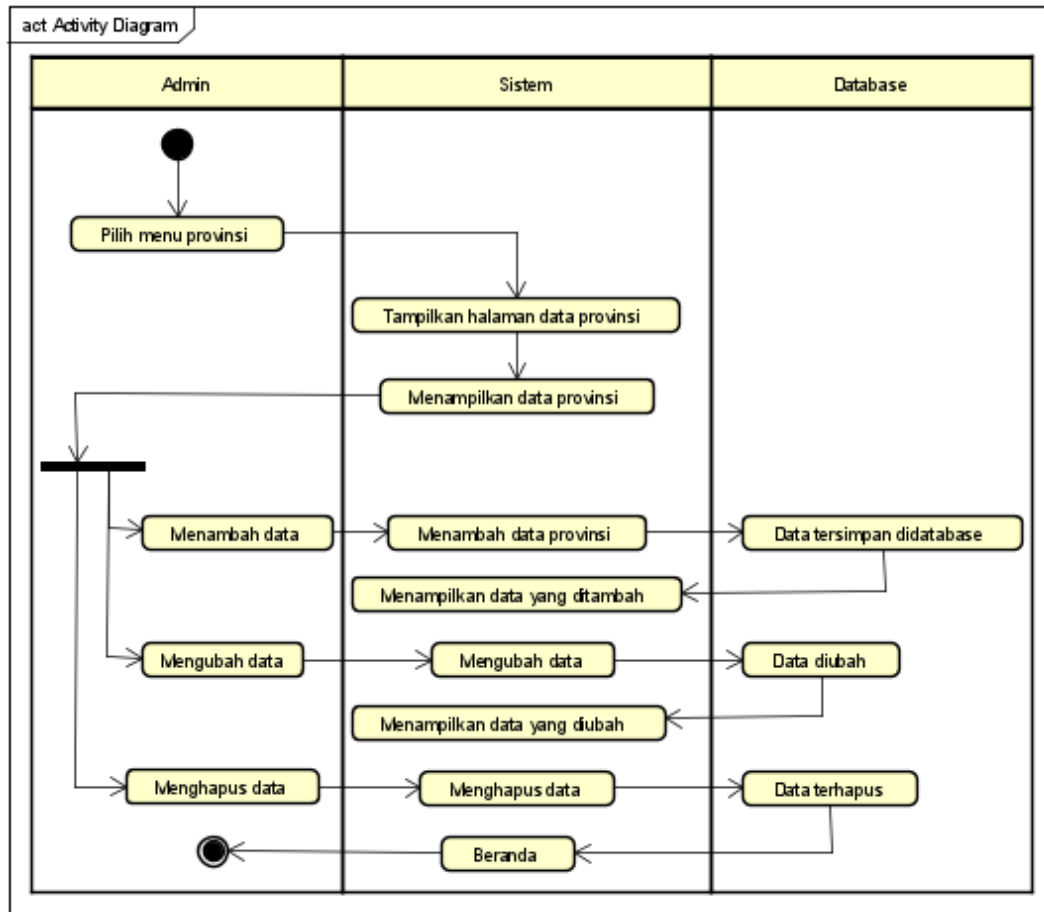


**Gambar 2** Use Case Diagram

### 1.1. Activity Diagram

Diagram aktivitas merupakan sebuah diagram dalam UML yang menunjukkan aspek dinamis dari suatu sistem. Sistem. Diagram aktivitas juga menggambarkan aliran sistem atau proses bisnis, atau menu tugas atau fungsi perangkat lunak. Diagram fungsional harus memperhitungkan bahwa fungsi yang dilakukan oleh sistem tidak sama dengan yang dilakukan oleh operator. . Elemen utama dari diagram aktivitas adalah aktivitas itu sendiri Operasi adalah operasi yang dilakukan oleh sistem. Setelah fungsi diidentifikasi, perlu diketahui bagaimana semua elemen ini berhubungan dengan kendala dan kondisi.

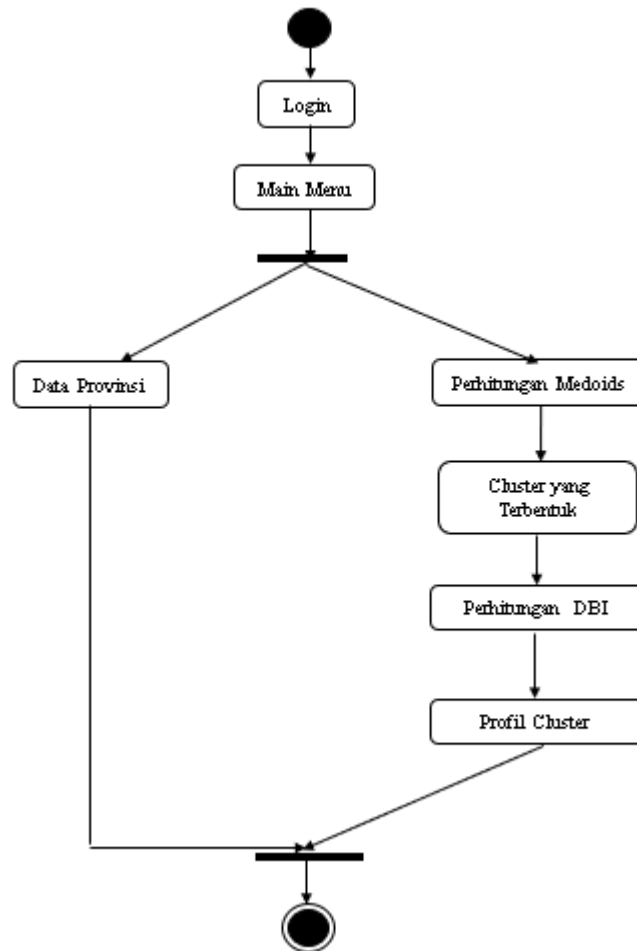
Diagram fungsional juga digunakan untuk mendefinisikan yang pertama untuk menentukan desain proses sistem, dan yang kedua untuk menentukan penggunaan sistem atau pengelompokan layar antarmuka pengguna, dengan asumsi bahwa setiap fungsi memiliki struktur layar antarmuka pengguna, dan kemudian diuji di mana setiap fungsi . digunakan. Fungsi tersebut memerlukan pengujian yang terdiri dari definisi kasus uji dan struktur menu yang terlihat terakhir dari perangkat lunak. Keuntungan Menggunakan Flowchart Saat menggunakan flowchart, penting untuk memperhatikan hubungan antara fungsi dan aktivitas dan memastikan bahwa diagram mudah dipahami dan mengikuti proses yang diuraikan. Di bawah ini adalah diagram fungsional



**Gambar 3** Activity Diagram Alur Mengedit Data

Berikut ini gambaran activity diagram kesehatan masyarakat provinsi indonesia meliputi sanitasi layak, air minum layak dan rumah layak huni. Activity diagram mengelola data provinsi merupakan rancangan proses yang akan dilakukan berupa aktifitas pada sistem bagian admin. Berikut penjelasannya

- Pada bagian awal admin melakukan akses pada *website*
- Admin melakukan akses pada *website* kesehatan lingkungan masyarakat
- Jika aktor berhasil login akan muncul tampilan menu utama dan apabila gagal akan kembali pada menu login
- Setelah berhasil log in, selanjutnya admin bisa mengakses menu yang berada pada *website* tersebut yaitu pengolahan data provinsi
- Untuk menambah, menghapus dan mengubah data yang ada admin masuk ke menu provinsi setelah itu admin bisa melakukan aktivitas pengelolaan data pada *website*.



**Gambar 4** Activity Diagram

Berikut ini gambaran activity diagram kesehatan masyarakat provinsi indonesia meliputi sanitasi layak, air minum layak dan rumah layak huni. Activity diagram mengelola data provinsi merupakan rancangan proses yang akan dilakukan berupa aktifitas pada sistem bagian admin. Berikut penjelasannya

- a. Pada bagian awal admin dan user melakukan akses pada *website*
- b. Admin atau user dapat melakukan akses pada *website* kesehatan lingkungan masyarakat
- c. Jika user atau admin berhasil login akan muncul tampilan menu utama dan apabila gagal akan kembali pada menu login
- d. Setelah berhasil log in, selanjutnya admin dan user dapat mengakses beberapa menu antara lain hasil clustering, perhitungan medoids, perhitungan DBI, profil cluster, dan cluster yang terbentuk
- e. Apabila telah menyelesaikan aktivitas pada website, admin dan user bisa logout meninggalkan websiter tersebut

#### 4. HASIL DAN PEMBAHASAN

Bab ini membahas hasil tes. Data yang diperoleh diolah dengan menggunakan metode Partitioning Around Medoids. Saat menguji hasil kinerja metode Partitioning Around Medoids, indeks pengukuran kinerja Davies Bouldin digunakan untuk menentukan kluster optimal dan hasil implementasi serta pembahasannya: Implementasi metode Davies-Bouldin Index (DBI) pada algoritma PAM (Partitioning Around Medoids). Data yang digunakan berdasarkan 34 provinsi di Indonesia sanitasi yang layak, kelayakan air minum, dan rumah layak huni pada tahun 2020-2021 yang diambil dari [Badan Pusat Statistik \(bps.go.id\)](https://bps.go.id)

##### A. Penentuan Cluster Optimum

Setelah dilakukan proses clustering menggunakan algoritma PAM (Partitioning Around Medoids), maka digunakan metode Davies-Bouldin Index untuk mencari cluster optimal dan menentukan cluster terbaik. Tabel 2 di bawah ini adalah hasil dari metode *Davies' Bouldin Index*:

**Tabel 2.** Perhitungan *Davies Bouldin-Index*

Cluster	<i>Davies Bouldin Index</i>
2	1,6837
3	2,5181
4	1,8873
5	1,3473
6	4,9553
7	2,2941
8	2,1649
9	2,4721
10	2,2973

**Sumber :** Hasil Perhitungan *cluster* Optimum

Berdasarkan hasil tes Presentasi tersebut mengidentifikasi kluster yang optimal untuk pengelompokan provinsi Indonesia berdasarkan indeks kesehatan masyarakat menggunakan algoritma Partitioning Around Medoide (PAM) dan metode Devies-Bouldin Index, pada kluster 5.

##### B. Analisis Cluster Optimum

Dari proses clustering menggunakan algoritma Partitioning Around Medoids dan pencarian cluster optimal menggunakan teknik Davies-Buldin Index, ditentukan jumlah cluster optimal dan berbagai properti cluster. Seperti diuraikan di bawah ini, ciri-ciri anggota provinsi terbentuk sebagai berikut:

#### 5. KESIMPULAN DAN SARAN

##### A. Kesimpulan

Berdasarkan Berdasarkan penelitian yang dilakukan dan hasil penelitian yang diperoleh, dapat ditarik kesimpulan sebagai berikut:

1. Dengan menggunakan metode Partitioning Around Medoids dan teknik Davies-Bouldin Index, dilakukan pengujian untuk mengetahui jumlah kluster optimal yang terbentuk dengan mengklasifikasikan provinsi Indonesia berdasarkan indikator kesehatan lingkungan. Dengan menggunakan metode partitioning around medoids, berdasarkan nilai indeks Davies-Bouldin sebesar 1,3473 didapatkan jumlah cluster yang optimal adalah 5 cluster. Rangkaian pengujian dijalankan dengan skenario kluster mulai dari 2 hingga 10 kluster.
2. Setelah melakukan pengelompokan dengan 5 cluster, hasilnya adalah sebagai berikut:
  - Cluster 1: Provinsi Bengkulu
  - Cluster 2: Provinsi Kepulauan Bangka Belitung, DKI Jakarta, Jawa Barat, dan Nusa Tenggara Timur



- Cluster 3: Provinsi Papua
- Cluster 4: Provinsi Lampung, Kepulauan Riau, Kalimantan Tengah, Kalimantan Selatan, dan Sulawesi Barat
- Cluster 5: Provinsi Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, Nusa Tenggara Barat, Kalimantan Barat, Kalimantan Timur, Kalimantan Utara, Kalimantan Tengah, Kalimantan Selatan, Sulawesi Tenggara, Gorontalo, Maluku, Maluku Utara, dan Papua Barat.

## B.SARAN

Tujuan dari penelitian ini adalah mengimplementasikan algoritma Partitioning Around Medoids. Agar hasilnya lebih bermanfaat, disarankan untuk membuat database sebagai wadah untuk mengimplementasikan algoritma Partitioning Around Medoids.

## 6. REFERENSI

- Badan Pusat Statistik. 2018. *Indikator Perumahan dan Kesehatan Lingkungan Tahun 2018*. Badan Pusat Statistik. Jakarta
- Fayyad, U. 1996. *Advances in Knowledge Discovery and Data Mining*. MIT Press.
- Han, J., & Kamber, M. 2006. *Data Mining : Concept and Techniques*, 2nd ed. San Fransisco: Morgan Kauffman.
- Hermawati, F.A. 2013. *Data Mining*. Surabaya: Andi Offset.
- Hermawati, F. A. 2013. *Data Mining*. Yogyakarta: Andi.
- Intelligent Systems. Yogyakarta: Andi Offset. Jollyta, D., Efendi, S., Zarlis, M., & Mawengkang, H. 2019. *Optimasi Cluster Pada Data Stunting: Teknik Evaluasi Cluster Sum of Square Error dan Davies Bouldin Index*. Medan: Prosiding Seminar Nasional Riset Information Science (SENARIS).
- Kaur, N. U., & Singh, D. 2014. *K-Medoids Clustering Algorithm – A Review*. [pdf] International Journal of Computer Application and Technology (IJCAT). ISSN. 2349-1841 Vol. 1, Issue 1. April 2014.
- Larose, D.T. 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Willey & Sons, Inc.
- Larose, D.T. 2006. *DATA MINING METHODS AND MODELS*, New Jersey: John Wiley & Sons, Inc.
- Lubis, I., Aisyah, N., & Mardikanto, A. K. 2019. *Pedoman Pengukuran Capaian Pembangunan Perumahan dan Permukiman Berbasis Hasil (Outcome)*. Jakarta: Kementerian Perencanaan Pembangunan Nasional.
- Mayasari, T. R. 2020. *Pengelompokkan Berdasarkan Variabel Kesehatan Lingkungan dan Pengaruhnya Terhadap Kemiskinan di Indonesia Tahun 2018*. Lampung: Jurnal Siger Matematika
- Ningsi, L. N., Poningsih, & Tambunan, H. S. 2021. *Implementasi Data Mining Kluster pada Rumah Tangga yang Memiliki Akses Hunian Layak Berdasarkan Provinsi*. Pematangsiantar: KLIK: Kajian Ilmiah Informatika dan Komputer.
- RuiXu and Donald C.W. 2009, *Clustering*, A John Wiley & Sons, Inc., Publication