



## Clustering tingkat pengangguran di Provinsi Maluku dengan algoritma K- Means

Fadli H. Wattiheluw<sup>1)</sup>

<sup>1)</sup>Teknik Informatika, Institut Teknologi Dan Bisnis Stikom Ambon  
Corresponding Author: <sup>1</sup> [fadliwattiheluw1994@gmail.com](mailto:fadliwattiheluw1994@gmail.com)

### Article Info

#### Article history:

Received: 01/09/2025,  
Revised: 29/09/2025  
Accepted: 07/10/2025  
Published: 20/10/2025

#### Keywords:

K-Means Clustering,  
pengangguran,  
provinsi Maluku.

### ABSTRAK

Tingkat pengangguran merupakan salah satu indikator penting dalam mengukur kondisi sosial ekonomi suatu wilayah. Provinsi Maluku, sebagai bagian dari wilayah Indonesia timur, memiliki karakteristik geografis dan sosial yang beragam, yang turut mempengaruhi tingkat pengangguran di setiap kabupaten/kota. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di provinsi Maluku berdasarkan tingkat pengangguran menggunakan algoritma K-Means. Data yang digunakan dalam penelitian ini diperoleh dari BPS (Badan Pusat Statistik). Hasil analisis menunjukkan bahwa wilayah di provinsi Maluku dapat diklasifikasikan ke dalam 3 cluster, yaitu: cluster 0 sebanyak 10 wilayah dengan tingkat pengangguran sedang; cluster 1 sebanyak 1 wilayah dengan tingkat pengangguran sedang; dan cluster 2 sebanyak 1 wilayah dengan tingkat pengangguran tinggi; Implementasi penelitian berupa website yang diharapkan dapat digunakan sebagai dasar dalam pengambilan kebijakan penanggulangan pengangguran secara lebih fokus dan efektif di provinsi Maluku.

### 1. PENDAHULUAN

Pengangguran merupakan salah satu permasalahan sosial ekonomi yang dihadapi hampir oleh setiap negara di dunia. Pengangguran adalah mereka yang termasuk dalam angkatan kerja dan secara aktif mencari pekerjaan, tetapi tidak dapat menemukan pekerjaan yang sesuai dengan keinginan mereka [1]. Keberadaan pengangguran mencerminkan ketidakseimbangan antara jumlah angkatan kerja yang tersedia dengan lapangan pekerjaan yang mampu menyerapnya. Menambahkan bahwa pengangguran juga mencakup mereka yang tidak sedang mencari pekerjaan saat ini, tetapi masih termasuk dalam kategori angkatan kerja [2]. Dampak dari pengangguran tidak hanya dirasakan secara individu, tetapi juga

dapat memengaruhi kestabilan ekonomi dan sosial masyarakat secara luas. Tingkat pengangguran yang tinggi seringkali berhubungan dengan rendahnya tingkat kesejahteraan, meningkatnya angka kemiskinan, serta timbulnya berbagai permasalahan sosial lainnya. Oleh karena itu, pemetaan dan analisis terhadap kondisi pengangguran menjadi langkah penting dalam mendukung perumusan kebijakan ketenagakerjaan yang lebih tepat sasaran.

Meskipun jumlah pengangguran bervariasi setiap tahun, hal tersebut tetap menjadi indikator penting dalam menilai kesejahteraan sosial dan ekonomi suatu daerah.

Di Provinsi Maluku, dinamika ketenagakerjaan dipengaruhi oleh berbagai faktor seperti akses pendidikan, kesempatan kerja, dan pembangunan wilayah. Ketimpangan antar wilayah seringkali menciptakan pola yang tidak merata dalam tingkat pengangguran. Oleh karena itu, diperlukan pendekatan yang tepat untuk mengidentifikasi kelompok wilayah dengan karakteristik serupa guna menunjang perumusan kebijakan yang lebih efektif.

Untuk memahami distribusi pengangguran secara lebih mendalam, diperlukan pendekatan analitis yang mampu mengelompokkan wilayah berdasarkan karakteristik penganggurannya. Metode clustering, seperti K-Means telah digunakan dalam berbagai penelitian untuk mengelompokkan daerah berdasarkan indikator sosial ekonomi, termasuk tingkat pengangguran. Menurut Xie dkk, Clustering K-Means bekerja dengan cara membagi objek data ke dalam beberapa kelompok/cluster yang berbeda sesuai dengan ukuran kesamaan dari data - data tersebut, sehingga untuk objek data yang berada dalam cluster mempunyai tingkat kesamaan terbesar sedangkan untuk objek data antar cluster yang berbeda mempunyai tingkat kesamaan terkecil [3]. Clustering atau analisis cluster adalah proses pengelompokan satu set benda-benda fisik atau abstrak ke dalam kelas objek yang sama [4]. Dengan menerapkan metode clustering pada data pengangguran di Provinsi Maluku, diharapkan dapat diidentifikasi pola-pola tertentu yang mencerminkan kondisi spesifik di masing-masing kabupaten/kota.

Beberapa penelitian terdahulu telah menunjukkan efektifitas implementasi dengan K-means. Dalam penelitiannya ditemukan bahwa penerapan K-means

pada Beberapa penelitian terdahulu telah menunjukkan efektivitas implementasi metode K-Means dalam menganalisis dan mengelompokkan data pengangguran. Studi-studi tersebut, yang dilakukan di berbagai wilayah seperti Jawa Barat, Jawa Tengah, dan tingkat nasional, membuktikan bahwa K-Means mampu memetakan wilayah berdasarkan kemiripan karakteristik pengangguran, sehingga memudahkan identifikasi daerah prioritas. Temuan ini memperkuat relevansi penggunaan K-Means dalam konteks Provinsi Maluku, yang juga memiliki dinamika pengangguran yang beragam di setiap kabupaten/kota.

Proyek ini bertujuan untuk melakukan analisis clustering terhadap tingkat pengangguran di Provinsi Maluku, guna mengidentifikasi kelompok wilayah dengan karakteristik pengangguran yang serupa. Hasil dari analisis ini diharapkan dapat memberikan informasi yang berguna bagi pemerintah daerah dan pemangku kebijakan dalam merumuskan strategi penanggulangan pengangguran yang lebih efektif dan tepat sasaran. Selain itu, pendekatan ini juga dapat menjadi dasar untuk penelitian lebih lanjut dalam bidang pembangunan ekonomi regional dan perencanaan tenaga kerja.

## 2. METODE PENELITIAN

K-Means adalah algoritma clustering berbasis centroid yang digunakan untuk mengelompokkan data ke dalam beberapa *cluster* berdasarkan kemiripan atribut [10]. Algoritma ini bekerja dengan cara menentukan pusat *cluster* secara acak, menghitung jarak terdekat menggunakan rumus Euclidean, mengelompokkan data berdasarkan jarak terdekat ke pusat *cluster*, memperbarui posisi pusat *cluster*, dan mengulangi proses tersebut hingga tidak ada perubahan anggota *cluster*. Jarak antara titik data dan pusat *cluster* biasanya dihitung menggunakan rumus Euclidean berikut:

$$d(x_i, \mu_j) = \sqrt{\sum (x_i - \mu_j)^2}$$

Keterangan:

$x_i$  = data kriteria

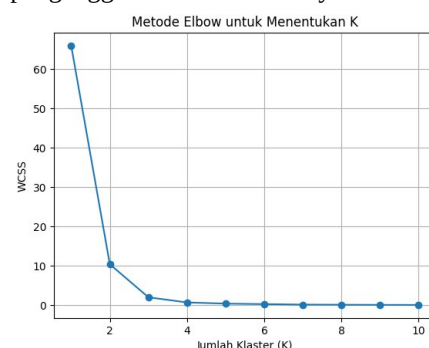
$\mu_j$  = centroid pada cluster ke - j

K-Means banyak digunakan karena algoritmanya yang sederhana namun efektif untuk analisis data berkelompok.

- A. Adapun tahapan dalam penelitian ini adalah

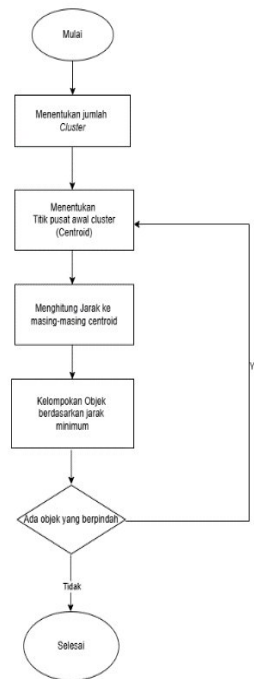
sebagai berikut: Dataset mencatat jumlah pengangguran di Provinsi Maluku berdasarkan kabupaten/kota untuk periode 2019–2024, disusun dalam format CSV. Data mencakup tingkat pengangguran terbuka (TPT).

- B. Data diimpor menggunakan bahasa pemrograman Python, kemudian dilakukan proses pra-proses data dengan bantuan library *pandas*. Normalisasi data dilakukan menggunakan *StandardScaler* untuk menyamakan skala variabel sebelum dilakukan proses clustering.
- C. Penentuan jumlah cluster optimal dilakukan menggunakan metode Elbow, dengan cara mencari titik siku (elbow point) pada grafik nilai SSE (Sum of Squared Error), untuk menentukan jumlah cluster yang paling representatif terhadap data.
- D. Algoritma K-Means diterapkan pada data yang telah dinormalisasi. Hasil clustering kemudian ditambahkan ke dalam dataset asli untuk mengelompokkan daerah berdasarkan tingkat pengangguran ke dalam beberapa kategori tertentu (misalnya: tinggi dan rendah).
- E. Hasil clustering divisualisasikan menggunakan scatter plot untuk melihat pola distribusi pengangguran berdasarkan tahun dan pie chart untuk menggambarkan proporsi kabupaten/kota dalam setiap cluster, sehingga memudahkan interpretasi pola pengangguran di wilayah Maluku.



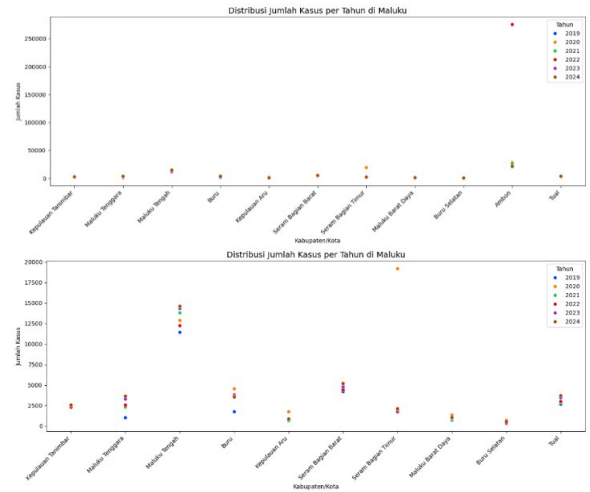
Untuk lebih jelasnya, alur proses penelitian menggunakan algoritma K-Means ditunjukkan

pada gambar di bawah.



## 1. Scatter Plot

Scatter Plot digunakan untuk menampilkan distribusi tingkat pengangguran per tahun di setiap kabupaten/kota di Provinsi Maluku. Setiap tahun direpresentasikan dengan warna yang berbeda, sehingga memudahkan untuk melihat perubahan tren dari tahun ke tahun serta pola penyebaran antar wilayah.



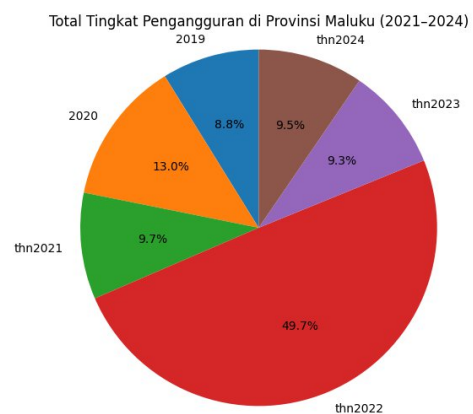
## 3. HASIL DAN PEMBAHASAN

### 1. Menentukan Jumlah Klaster

Setelah menerapkan algoritma K-Means pada data pengangguran di Provinsi Maluku, peneliti memperoleh hasil clustering dengan menggunakan jumlah cluster optimal yang telah ditentukan sebelumnya. Peneliti mendapat jumlah cluster  $k = 3$  berdasarkan hasil analisis menggunakan metode Elbow Method. Hasil clustering ini direpresentasikan dalam bentuk label cluster yang ditambahkan ke dalam dataset asli. Setiap cluster mencerminkan kelompok kabupaten/kota yang memiliki pola serupa dalam tingkat pengangguran selama periode tahun 2020 hingga 2023.

## 1. Pie Chart Total Pengangguran

Pie Chart (Diagram Pie) digunakan untuk menunjukkan ukuran sebagai perbandingan data. Pie Chart Total Pengangguran menunjukkan presentase tingkat pengangguran dari tahun 2019 – 2024. Ini memberikan gambaran komparatif tentang frekuensi tingkat pengangguran setiap tahunnya.

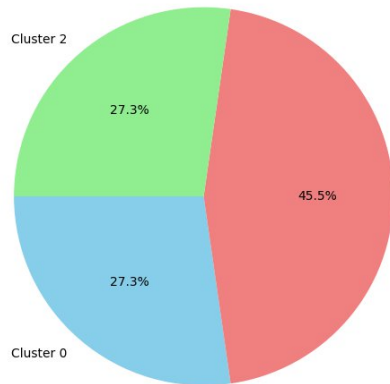


### 1.4. Pie Chart Distribusi Cluster

Pie Chart Distribusi Cluster menampilkan jumlah kabupaten

/ kota yang masuk ke dalam masing-masing cluster. Pie Chart ini membantu memahami proporsi relatif dari setiap cluster terhadap total tingkat pengangguran.

Presentase jumlah Kabupaten/Kota dalam klaster



### 1.5. Hasil Clustering

Hasil ini menunjukkan bahwa pembagian wilayah menjadi 3 cluster berbeda berdasarkan karakteristik tingkat pengangguran.

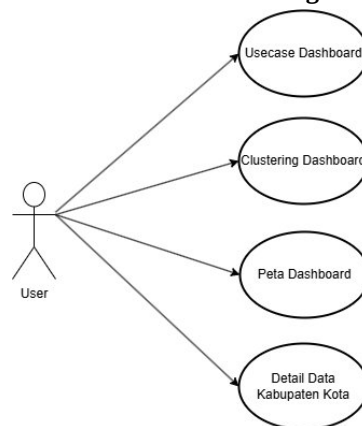
|                    |  |  |
|--------------------|--|--|
| Bagian Barat       |  |  |
| Seram Bagian Timur |  |  |
| Buru               |  |  |
| Buru Selatan       |  |  |

- Cluster 0 memiliki karakteristik jumlah pengangguran yang sedang ke rendah yaitu dengan jumlah 10000 pengangguran perkabupaten / kota setiap tahunnya dibandingkan dengan cluster satunya.
- Cluster 1 menunjukkan bahwa tingkat pengangguran paling tinggi hanya ditunjukkan oleh 1 kota saja karena data pengangguran di setiap tahunnya lebih dari 20000 pengangguran , yang mana bahkan tidak ada data satupun kabupaten / kota yang mendekati kota tersebut.
- Cluster 2 masuk dalam tingkat sedang yang hanya terdiri dari 1 kota saja dengan jumlah pengangguran di bawah 20000 dan di atas 10000.

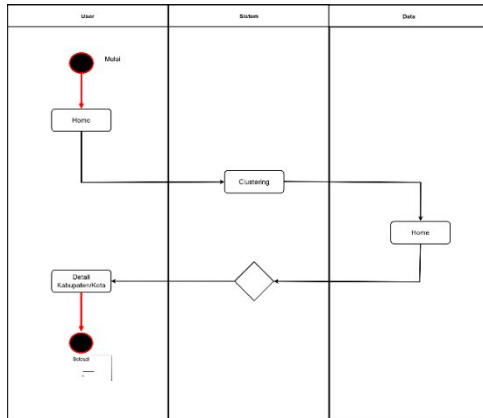
Tabel 1. Hasil Clustering

| Cluster 0          | Cluster 1  | Cluster 2     |
|--------------------|------------|---------------|
| Kepulauan Tanimbar | Kota Ambon | Maluku Tengah |
| Kepulauan Aru      |            |               |
| Maluku Tenggara    |            |               |
| Kota Tual          |            |               |
| Maluku Barat Daya  |            |               |
| Seram              |            |               |

### 1.5. Unified Modeling Language



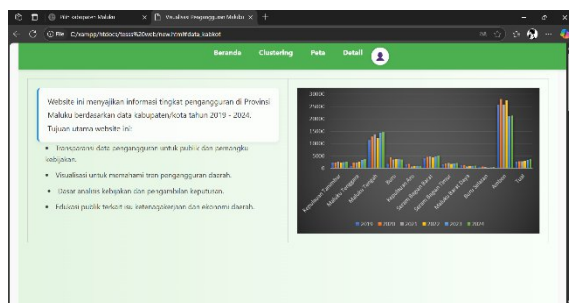
Dari gambar ini bahwa use case diagram mengilustrasikan fungsionalitas yang diinginkan dari sebuah sistem. Use case bertujuan untuk memodelkan interaksi antara aktor dan sistem. Aktor adalah entitas manusia yang terlibat dalam sistem untuk melakukan tugas-tugas tertentu.



Dari gambar ini bahwa Activity Diagram menggambarkan aliran kerja atau aktivitas dalam sistem yang akan dieksekusi. Diagram ini juga digunakan untuk mendefinisikan atau mengelompokkan proses dalam sistem. Gambar aktivitas memiliki komponen Activity Diagram menggambarkan aliran kerja atau aktivitas dalam sistem yang akan dieksekusi.

### 1.6. Home

Pada tampilan Home, pengguna diperlihatkan tujuan dari website ini dan hasil dari clustering yang sudah dilakukan



Gambar 9. Tampilan Home Website

### 1.7. Clustering

Tampilan clustering di rancang untuk menampilkan hasil analisis clustering yang dilakukan menggunakan k-means. Pada tampilan ini bertujuan untuk memberikan informasi terperinci mengenai pengelompokkan cluster di Provinsi Maluku berdasarkan tingkat pengangguran.

| Kabupaten           | 2019  | 2020  | 2021  | 2022  | 2023  | 2024  | Clustering |
|---------------------|-------|-------|-------|-------|-------|-------|------------|
| Kepulauan Sula-Buru | 2437  | 2307  | 2403  | 2991  | 2511  | 2629  | 3          |
| Maluku Tengah       | 9135  | 2238  | 2234  | 2227  | 2241  | 2259  | 3          |
| Maluku Utara        | 11726 | 5280  | 5385  | 5178  | 5111  | 5481  | 3          |
| Buru                | 3147  | 4535  | 2917  | 3875  | 3241  | 3511  | 3          |
| Kepulauan Aru       | 1179  | 1185  | 4217  | 384   | 999   | 814   | 3          |
| Sulawesi Barat      | 4721  | 4628  | 4833  | 4675  | 4798  | 5197  | 3          |
| Sulawesi Tengah     | 5175  | 5095  | 2733  | 5128  | 5262  | 4732  | 3          |
| Maluku Barat Daya   | 1728  | 1114  | 214   | 421   | 1365  | 1896  | 3          |
| Buru Selatan        | 482   | 438   | 518   | 586   | 371   | 351   | 3          |
| Ambon               | 27861 | 28332 | 26760 | 27835 | 27555 | 27526 | 1          |
| Mal                 | 3474  | 4817  | 3634  | 2491  | 3114  | 3784  | 3          |
| Maluku              | 54625 | 42348 | 58582 | 30737 | 37334 | 51963 |            |

Gambar 10. Tampilan Hasil Clustering

### 1.8. Peta

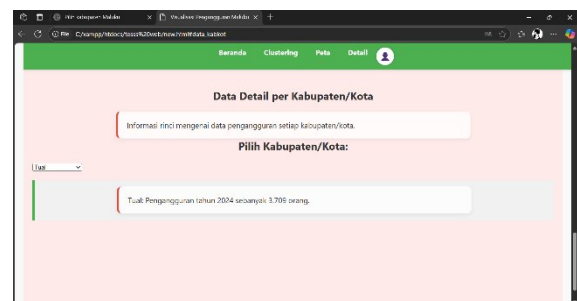
Pada tampilan peta, di desain untuk memberikan visualisasi geografis dari tingkat pengangguran di Provinsi Maluku berdasarkan pembagian cluster.



Gambar 11. Tampilan Sebaran Tingkat Pengangguran

### 1.9. Detail Kabupaten / Kota

Pada tampilan ini, menyajikan lebih detail tentang jumlah pengangguran di setiap kabupaten / kota pada Provinsi Maluku.



Gambar 12. Tampilan Data Tingkat Pengangguran per Kabupaten/Kota

## 4. KESIMPULAN

Penelitian ini menggunakan algoritma K-Means untuk mengelompokkan data clustering tingkat pengangguran di Provinsi Maluku. Untuk tingkat pengangguran di maluku selama periode 2019-2024.

berdasarkan hasil clustering, ditemukan nilai  $K=3$  yang mencerminkan pola distribusi di provinsi maluku. cluster pertama (cluster 0) yang terdiri dari Kepulauan Tanimbar, Kepulauan Aru, Maluku Tenggara, Kota Tual, Maluku Barat Daya, Seram Bagian Barat, Seram Bagian Timur, Buru dan Buru Selatan Dengan Jumlah Tingkat pengangguran Rendah. Adapun juga cluster 1 menunjukkan bahwa tingkat pengangguran yang tergolong Tinggi di tunjukkan oleh satu Kabupaten/kota yaitu kota ambon.dengan tingkat pengangguran Tinggi. Dan untuk cluster 2 dengan hanya terdiri dari Kabupaten/kota yaitu Kabupaten Maluku Tengah dengan Tingkat pengangguran Sedang hanya satu.

## DAFTAR PUSTAKA

- [1] A. I. Ramadhan, P. D. Atika, and K. F. Ramdhania, "Analisis clustering K-Means untuk pemetaan tingkat pengangguran terbuka di provinsi-provinsi Indonesia tahun 2013–2023," *Jurnal Statistika dan Riset Computasi*, vol. 5, no. 2, pp. 112–121, 2024. [Online]. Available: <https://doi.org/10.31599/wbpydb62>
- [2] C. N. Rianda, "Analisis dampak pengangguran berpengaruh terhadap individual," *At-Tasyri': Jurnal Ilmiah Prodi Muamalah*, vol. 12, no. 1, pp. 11–20, 2020. [Online]. Available: <http://ejournal.staindirundeng.ac.id/index.php/Tasyri>
- [3] I. R. Padiku and A. Lahinta, "Penerapan clustering K-Means untuk mendukung pengelolaan koleksi pada perpustakaan Fakultas Teknik Universitas Negeri Gorontalo," *Jurnal Teknologi*, vol. 20, no. 1, pp. 33–41, 2022. [Online]. Available: <https://doi.org/10.37031/jt.v20i1.206>
- [4] N. L. Anggreini, "Teknik clustering dengan algoritma K-Medoids untuk menangani strategi promosi di Politeknik TEDC Bandung," *Jurnal Teknologi Informasi dan Pendidikan (JTIP)*, vol. 12, no. 2, pp. 45–54, 2019. [Online]. Available: <https://www.neliti.com/publications/328497>
- [5] D. Nurhasanah and R. A. Wibowo, "Analisis Metode Clustering pada Data Kependudukan di Indonesia," *\*Jurnal Teknologi Informasi dan Komputer\**, vol. 7, no. 3, pp. 125–133, 2022.
- [6] Badan Pertanahan Nasional. (2023). Informasi Pertanahan Indonesia. Diakses dari <https://www.atrbpn.go.id>
- [7] Stack Overflow, "Forum Tanya Jawab Pemrograman," 2025. [Online]. Available: <https://stackoverflow.com>
- [8] Google, "Mesin Pencari Informasi," 2025. [Online]. Available: <https://www.google.com>
- [9] Tech With Tim, "Tutorial Python Pandas untuk Pemula [Video YouTube]," 2023. [Online]. Available: <https://www.youtube.com/watch?v=xyz123>
- [10] DQLab, "Kenali Fungsi Pandas Python untuk Analisis Data," 2023. [Online]. Available: <https://dqlab.id/kenali-fungsi-pandas-python-untuk-analisis->