

ANALISIS TINGKAT PENGANGGURAN DI KOTA PALOPO MENGGUNAKAN METODE *K-MEANS*

Julita Veronika Lenda^{1*}, Magdalena A. Ineke Pakereng²

^{1,2}Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Kristen Satya Wacana

Jl. Diponegoro 52-60, Salatiga 50711, Indonesia

e-mail: ¹672020266@student.uksw.edu, ²ineke.pakereng@uksw.edu

Abstrak

Pengangguran merujuk pada kondisi dimana seseorang belum memperoleh pekerjaan dalam artian masih mencari pekerjaan, tetapi dapat juga diartikan sebagai telah bekerja tetapi belum mendapatkan hasil yang berkecukupan. Pengangguran mempunyai dampak yang sangat negatif terhadap perekonomian dan masyarakat. Oleh karena itu, diperlukan upaya untuk mengatasi permasalahan tersebut. Tujuan dari penelitian ini yaitu untuk mengetahui perkembangan tingkat pengangguran di kota Palopo. Penelitian ini menggunakan teknik data mining atau yang biasa disebut dengan algoritma *K-Means*. Algoritma *K-Means* adalah teknik pengelompokan (*clustering*) yang berfungsi membagi suatu dataset menjadi beberapa kelompok variabel. Hasil penelitian menunjukkan pada tujuh tahun terakhir jumlah persentase tingkat pengangguran di Kota Palopo mengalami naik turunnya tingkat pengangguran. Hasil penelitian ini diharapkan dapat membantu pemerintah untuk mengetahui perkembangan tingkat pengangguran di Kota Palopo dari tahun ke tahun.

Kata kunci: Data Mining, Algoritma *K-Means*, Pengangguran, *Clustering*.

Abstract

Unemployment refers to a condition where a person has not obtained a job in the sense that he is still looking for work, but can also be interpreted as having worked but has not received sufficient results. Unemployment has a very negative impact on the economy and society. Therefore, efforts are needed to overcome these problems. The purpose of this study is to determine the development of the unemployment rate in Palopo City. This research uses data mining techniques or commonly referred to as the *K-Means* algorithm. The *K-Means* algorithm is a clustering technique that functions to divide a dataset into several groups of variables. The results showed that in the last seven years the percentage of unemployment rate in Palopo City has experienced ups and downs in the unemployment rate. The results of this study are expected to help the government to find out the development of the unemployment rate in Palopo City from year to year.

Keywords: Data Mining, *K-means* Algorithm, Unemployment, *Clustering*.

1. PENDAHULUAN

Indonesia merupakan negara dengan jumlah penduduknya mencakup nilai yang cukup besar [1]. Informasi yang diperoleh dari situs Badan Pusat Statistik (BPS) menyatakan bahwa perkembangan penduduk di setiap tahun mencakup persentase kenaikan 1%. Angka pengangguran di Indonesia pada tahun 2021 mencapai 6,49 atau setara dengan 9,1 juta individu yang tidak bekerja [2]. Pemerintah Indonesia sudah banyak mengeluarkan kebijakan – kebijakan yang berupa memutus dan mengurangi persentase pengangguran di Indonesia, salah satu upaya dari pemerintah ialah dengan adanya pelatihan – pelatihan kerja yang dapat mendongkrak *soft* maupun *hard skill* pada calon pekerja. Selain hal tersebut, pemerintah mengeluarkan kebijakan dengan kartu pra kerja yang dapat dikatakan berhasil dalam membantu penduduk yang belum bekerja dengan menerima bantuan perbulannya. Akan tetapi hal tersebut belum begitu

berdampak besar menurunkan persentase pengangguran di Indonesia, 7 dari 10 orang lebih memilih bekerja di bawah sektor instansi dibandingkan dengan memilih bekerja sebagai wirausaha [3].

Clustering merupakan metode yang digunakan dalam mengelompokkan data menjadi beberapa kelompok – kelompok data yang disebut sebagai *cluster* [4]. Hubungan antara clustering dan pengangguran terletak pada kemampuan clustering untuk dalam membantu menganalisis dan mengelompokkan data yang berhubungan dengan tingkat penghasilan. Clustering dapat digunakan untuk mengelompokkan daerah – daerah berdasarkan karakteristik ekonomi mereka, seperti tingkat pengangguran, tingkat pendidikan, bahkan tingkat pengangguran. Informasi ini dapat membantu pemerintah dalam menyediakan pelatihan atau lapangan kerja dengan menyesuaikan program – program yang dibutuhkan.

Menganalisis tingkat pengangguran di Kota Palopo dapat digunakan metode *K-Means* atau biasa disebut pengelompokan data [5]. *K-Means* adalah salah satu metode analisis data yang mengelompokkan data kedalam beberapa kelompok *cluster*. Hasil dari penelitian ini menggunakan metode *K-Means*, data pengangguran dari tahun 2008 sampai dengan tahun 2022 dibagi menjadi tiga *cluster*, yaitu *cluster 0* sebanyak dua data (2008 – 2009), *cluster 1* sebanyak tujuh data (2016 – 2022), dan *cluster 2* sebanyak enam data (2010 – 2015) dengan total keseluruhan data sebanyak 15 data. Pada tujuh tahun terakhir jumlah persentase tingkat pengangguran di Kota Palopo mengalami naik turunnya tingkat pengangguran. Dengan informasi ini, diharapkan dapat membantu pemerintah untuk mengetahui perkembangan tingkat pengangguran di Kota Palopo dari tahun ke tahun.

Penelitian sebelumnya telah memanfaatkan algoritma *K-Means* untuk menganalisis data pengangguran di berbagai wilayah Indonesia. Provinsi Jawa Barat, ditemukan bahwa dengan menggunakan nilai $k = 2$, berhasil membentuk dua cluster yang validasinya dilakukan dengan menggunakan metode Davies-Bouldin Index (DBI) dengan nilai sebesar 0,733 [6]. Penelitian lainnya meneliti perbandingan tingkat pengangguran antar provinsi berdasarkan data TPT Data Mining, mengidentifikasi perubahan tren pengangguran dari tahun 2019 hingga 2021, termasuk pergeseran status Sumatera Barat dari provinsi dengan tingkat pengangguran rendah menjadi tinggi [7].

Penelitian terkait dampak pandemi COVID-19 pada tingkat pengangguran di Indonesia menunjukkan bahwa provinsi Banten mengalami tingkat pengangguran tertinggi, sementara Sulawesi Barat memiliki yang terendah selama masa pandemi. Temuan ini memberikan gambaran tentang dampak yang signifikan dari pandemi terhadap ketenagakerjaan di berbagai wilayah Indonesia [8]. Berfokus pada penerapan metode *K-Means* untuk mengelompokkan data pengangguran di tingkat lokal di Jawa Barat, berhasil diidentifikasi dua cluster yang menunjukkan perbedaan yang signifikan dalam tingkat pengangguran antar-kabupaten. Hal ini memberikan informasi penting bagi pemerintah dalam merancang kebijakan ketenagakerjaan di tingkat lokal [9].

Pengukuran tingkat pengangguran di berbagai daerah di Indonesia dengan menggunakan metode *K-Means* membawa manfaat yang baik bagi pemerintah. Penelitian ini mampu mengelompokkan tingkat pengangguran kedalam beberapa kategori sehingga memudahkan dalam mengidentifikasi wilayah dengan tingkat penganggurannya, seperti di pulau Jawa dan Bali yang memiliki tingkat pengangguran yang lebih tinggi dibandingkan daerah lainnya [10]. Hal yang sama terjadi di Sulawesi Selatan menggunakan metode yang sama, dimana beberapa seperti Palopo, Parepare, dan Makassar memiliki tingkat pengangguran yang tinggi sedangkan daerah lainnya memiliki tingkat pengangguran yang lebih rendah [11].

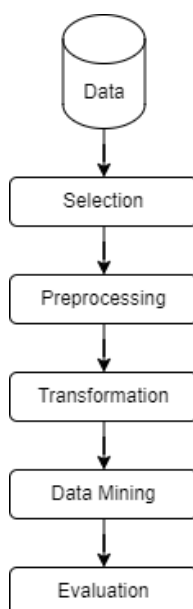
Penerapan metode *K-Means* pada data pengangguran di Indonesia memberikan gambaran lebih detail mengenai sebaran pengangguran tingkat lokal. Informasi ini memberikan peluang bagi pemerintah untuk mengarahkan upaya membangun ekonomi dan menciptakan lapangan kerja secara luas dengan tujuan mendorong pemerataan pertumbuhan ekonomi di seluruh Indonesia [12]. Penggunaan *K-Means* dalam mengelompokkan data pengangguran di Indonesia menjadi beberapa cluster berdasarkan tingkat pengangguran memberikan pemahaman yang mendalam tentang pola pengangguran di berbagai wilayah Indonesia yang menjadi landasan

bagi pemerintah dalam menyusun strategi untuk memperluas lapangan kerja dan mendorong pertumbuhan ekonomi secara menyeluruh.

Penelitian-penelitian tersebut memberikan gambaran komprehensif tentang kondisi pengangguran di Indonesia dengan menggunakan algoritma *K-Means* untuk analisis clustering. Melalui pendekatan yang beragam, seperti menggunakan data provinsi, menganalisis perubahan tren dari tahun ke tahun, dan menerapkan algoritma pada tingkat lokal, jurnal-jurnal tersebut memberikan pemahaman mendalam tentang pola pengangguran di berbagai wilayah. Mengamati dampak peristiwa eksternal, seperti pandemi COVID-19 terhadap tingkat pengangguran di berbagai wilayah, yang memperdalam pemahaman tentang dinamika pasar tenaga kerja dan pentingnya respons pemerintah dalam menghadapi krisis ekonomi. Dengan demikian, jurnal-jurnal tersebut menjadi sumber bagi peneliti dalam memahami, menganalisis, dan merespons permasalahan pengangguran di Indonesia.

2. METODE PENELITIAN

Berikut tahapan dalam penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

Data yang digunakan pada penelitian ini diambil dari kantor Badan Pusat Statistik (BPS) kota Palopo dari tahun 2008 – 2022. Masuk pada tahap *Preprocessing/Data Cleaning*, proses ini meliputi penghapusan variabel yang tidak digunakan dan memperbaiki kesalahan data seperti kesalahan cetak. Data yang tersedia selanjutnya akan masuk pada proses *transformation*, dimana data akan diubah ke dalam format yang sesuai untuk diproses pada tahapan *data mining*. Proses ini penting dikarenakan algoritma *K-Means* hanya bisa memproses data berupa data numerik. Masuk pada tahapan *Data Mining*, tahapan ini memuat pencarian pola atau informasi yang diinginkan [13]. Fase ini merupakan proses utama dalam mencari informasi berguna dari sejumlah besar data yang ada [14]. Data yang digunakan dalam penelitian ini adalah data pengangguran Kota Palopo pada tahun 2008 hingga tahun 2022. Tahapan terakhir yaitu masuk pada tahapan evaluasi (*evaluation*). Tahapan ini menggunakan metode *Silhouette*. Pada metode ini jumlah *cluster* yang dipilih adalah nilai *Silhouette* yang terbesar [15].

Hasil pengujian yang diperoleh dari data yang telah melalui proses visualisasi menghasilkan nilai *k* terbaik yaitu nilai $k = 3$ dimana jarak tiap *cluster* sebesar 6.051. Data pengangguran yang digunakan dari tahun 2008 sampai dengan tahun 2022 dibagi menjadi tiga *cluster*, yaitu *cluster* 0 sebanyak dua data (2008 – 2009), *cluster* 1 sebanyak tujuh data (2016 – 2022), dan *cluster* 2 sebanyak enam data (2010 – 2015) dengan total keseluruhan data sebanyak 15 data.

3. HASIL DAN PEMBAHASAN

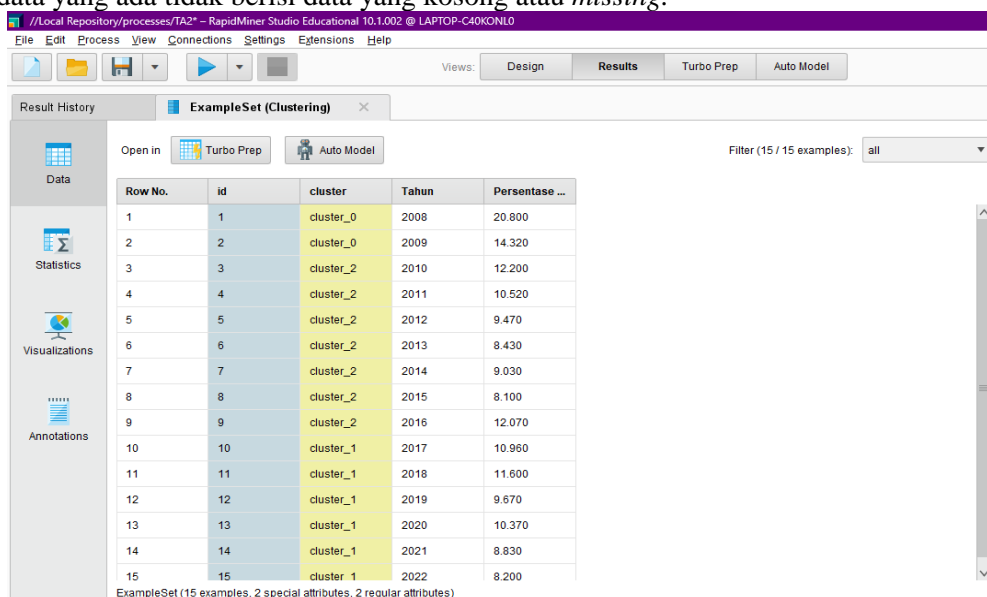
Data yang digunakan dalam penelitian ini adalah data tingkat pengangguran di Kota Palopo dalam rentan waktu 2008 hingga tahun 2022 dinyatakan dalam format persentase.

Tabel 1. Dataset Tingkat Pengangguran

Tahun	Persentase Tingkat Pengangguran Terbuka	Satuan
2008	20,8	Persen
2009	14,32	Persen
2010	12,2	Persen
2011	10,52	Persen
2012	9,47	Persen
2013	8,43	Persen
2014	9,03	Persen
2015	8,1	Persen
2016	12,07	Persen
2017	10,96	Persen
2018	11,6	Persen
2019	9,67	Persen
2020	10,37	Persen
2021	8,83	Persen
2022	8,2	Persen

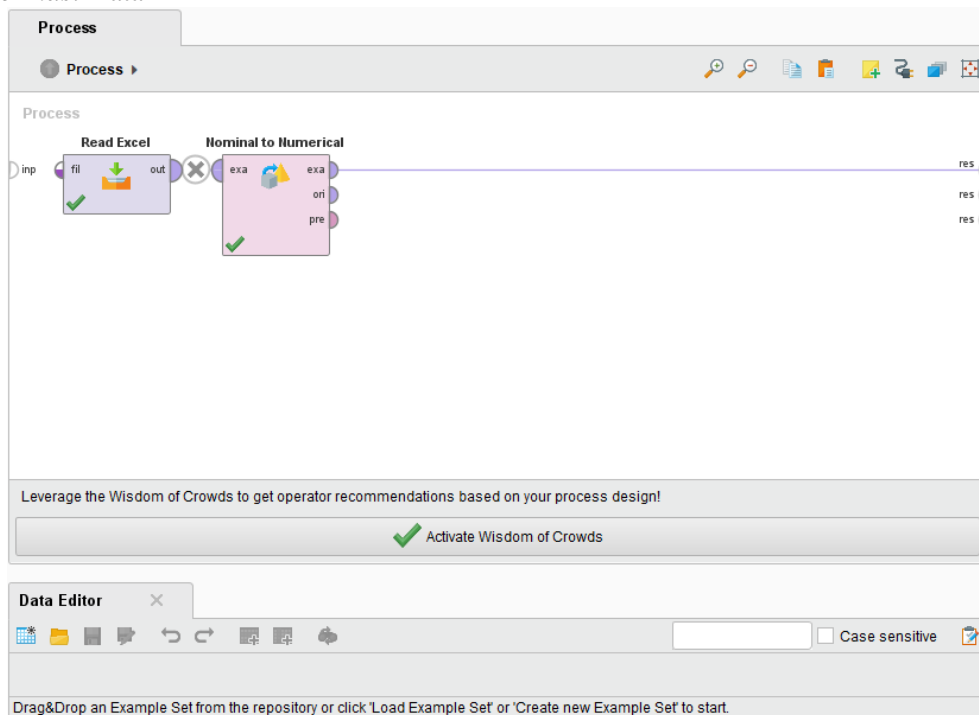
1. Data Cleaning

Tahapan ini menggunakan operator *Replace missing value* di *rapidminer* yang mana untuk mengetahui apakah data yang ada terdapat data kosong maupun *missing*. Hal ini dihasilkan bahwa data yang ada tidak berisi data yang kosong atau *missing*.



Gambar 2. Proses Data Cleaning

2. Transformasi Data



Gambar 3. Transformasi data

Setelah melewati tahap data *cleaning*, selanjutnya data akan ditransformasikan sesuai dengan ketentuan proses *data mining*. Yang dimana data tidak dapat diproses jika variabel data bukan numerik. Disini user menggunakan operator *Nominal To Numeric* untuk mengubah data dari atribut persen menjadi atribut numerik.

The screenshot shows the 'ExampleSet (Nominal to Numerical)' window in RapidMiner Studio. The window displays a table with 15 rows of data. The columns are 'Row No.', 'Satuan = Pe...', 'Tahun', and 'Persentase ...'. The data shows a decreasing trend in the 'Persentase ...' column over the years from 2008 to 2022. The 'Satuan = Pe...' column contains the value '1' for all rows.

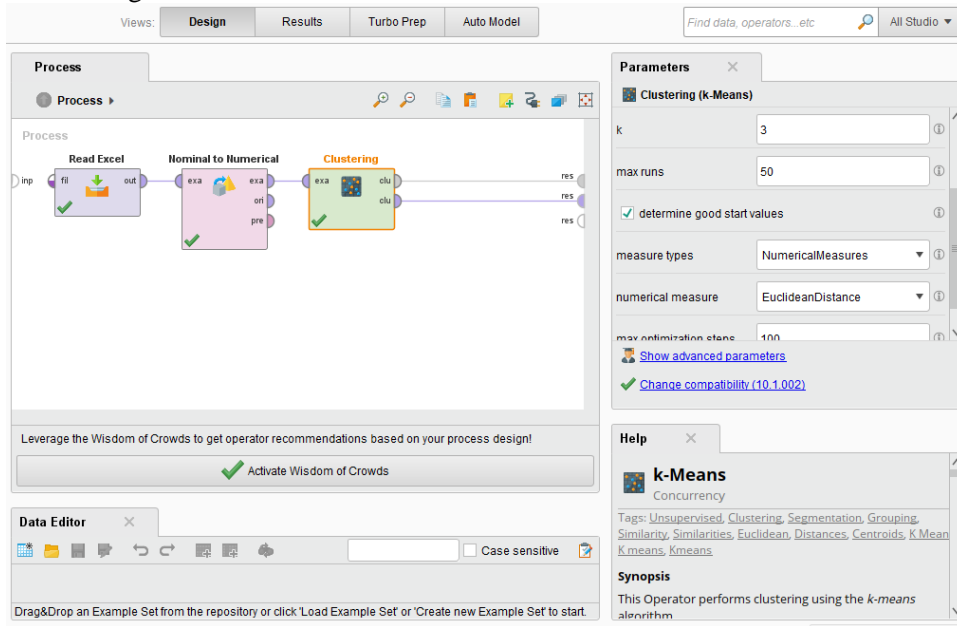
Row No.	Satuan = Pe...	Tahun	Persentase ...
1	1	2008	20.800
2	1	2009	14.320
3	1	2010	12.200
4	1	2011	10.520
5	1	2012	9.470
6	1	2013	8.430
7	1	2014	9.030
8	1	2015	8.100
9	1	2016	12.070
10	1	2017	10.960
11	1	2018	11.600
12	1	2019	9.670
13	1	2020	10.370
14	1	2021	8.830
15	1	2022	8.200

ExampleSet (15 examples, 0 special attributes, 3 regular attributes)

Gambar 4. Hasil Transformasi Data

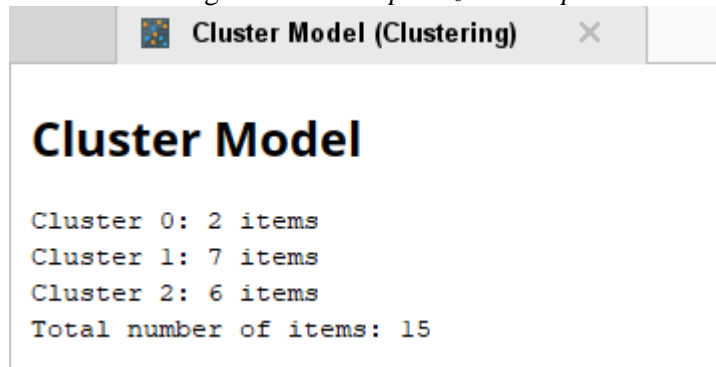
Pada gambar diatas menghasilkan data yang telah mengalami transformasi atau perubahan data pada bagian satuan yang awalnya berisikan persen sudah terganti menjadi variabel numerik yaitu 1.

3. Proses Data Mining



Gambar 5. Proses Data Mining

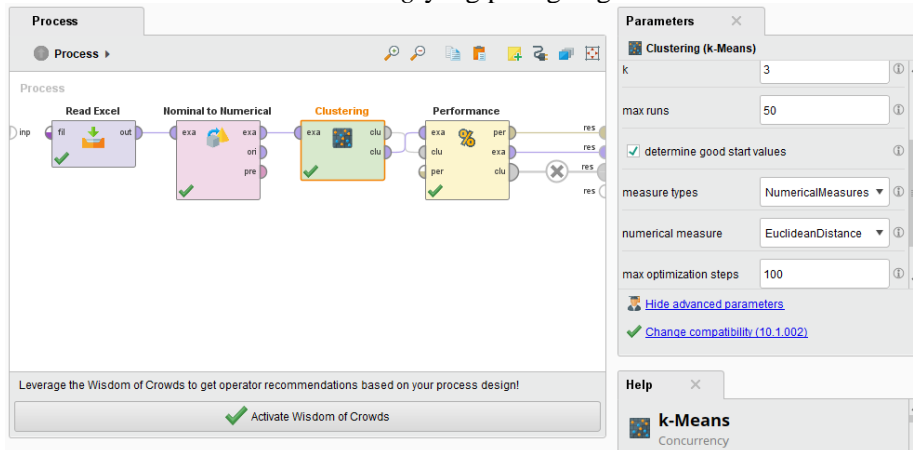
Pada proses ini, algoritma yang akan digunakan yaitu algoritma *K-Means*. pada bagian parameter *clustering* untuk nilai $k = 3$ dan maksimal run 50. Lalu pada bagian *measure type* menggunakan *Numerical Measure* dengan maksimal *optimization steps* sebesar 100.



Gambar 6. Hasil Proses Data Mining

4. Evaluasi

Pada tahapan ini, untuk mengetahui nilai *Silhouette* dibutuhkan panel operator *cluster performance* untuk memvalidasi nilai *clustering* yang paling bagus.



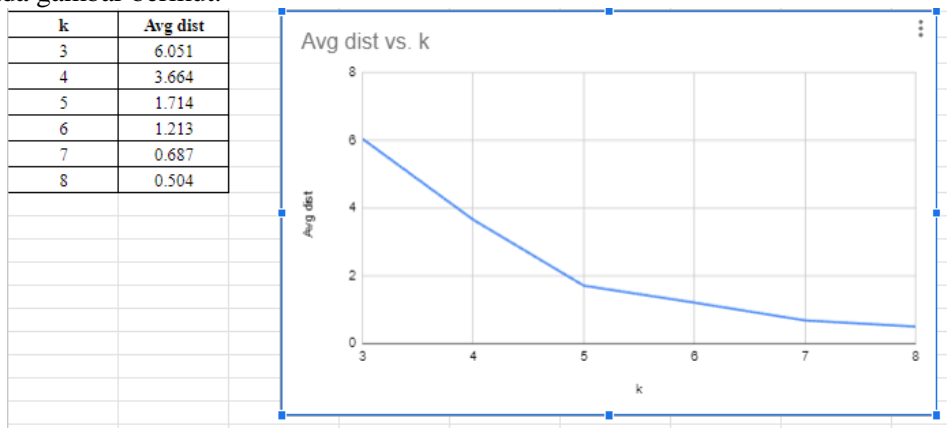
Gambar 7. Proses Evaluasi

Dalam pengujian ini, nilai k yang akan diuji yaitu nilai k = 3 sampai dengan nilai k = 8, diperoleh nilai seperti berikut.

Tabel 2. Hasil Pengujian nilai k = 3 hingga k = 8

k	Avg dist
3	6.051
4	3.664
5	1.714
6	1.213
7	0.687
8	0.504

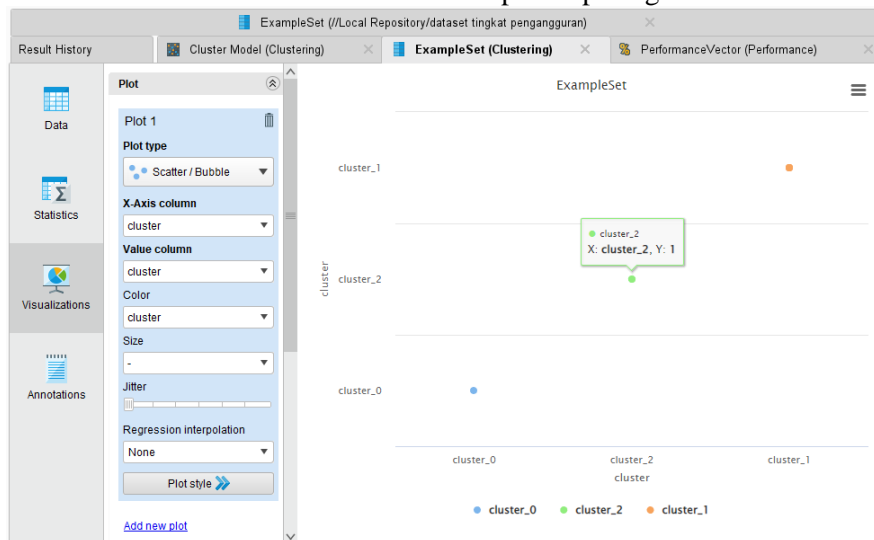
Untuk mengetahui nilai *cluster* terbaik maka akan dilakukan visualisasi menggunakan kurva seperti pada gambar berikut.



Gambar 8. Hasil Visualisasi

Dari gambar diatas, untuk menentukan nilai *Silhouette* terbaik perhatikan pada titik yang paling siku. Sesuai dengan ketentuan metode ini, nilai *cluster* terbaik yaitu terdapat pada nilai k = 3 dimana jarak tiap *cluster* sebesar 6.051. Ini membagi data menjadi tiga *cluster* yaitu *cluster* 0 berjumlah 2 data, *cluster* 1 berjumlah 7 data, dan *cluster* 2 berjumlah 6 data sehingga total keseluruhan sebanyak 15 data.

Berikut diagram *Scatter/Bubble* untuk nilai k = 3 ditampilkan pada gambar 9 berikut.



Gambar 9. Diagram Hasil Pengelompokan k = 3

Hasil pengujian data menghasilkan nilai k terbaik yaitu nilai k = 3 dimana jarak tiap *cluster* sebesar 6.051. Data pengangguran yang digunakan dari tahun 2008 sampai dengan tahun 2022 dibagi menjadi tiga *cluster*, yaitu *cluster* 0 sebanyak dua data (2008 – 2009), *cluster* 1 sebanyak

tujuh data (2016 – 2022), dan *cluster* 2 sebanyak enam data (2010 – 2015) dengan total keseluruhan data sebanyak 15 data.

4. KESIMPULAN

Dari hasil pengelompokan data pengangguran Kota Palopo dari tahun 2008 sampai dengan tahun 2022 menggunakan algoritma *clustering* yaitu *K-Means* dengan parameter $k = 3$ diperoleh hasil bahwa nilai *clustering* terbaik telah berhasil diidentifikasi. Hasil *clustering* tersebut validitas menggunakan metode *Silhouette* dengan nilai 6.051. Dari hasil percobaan tersebut, data dibagi menjadi tiga *cluster* yaitu *cluster* 0 sebanyak 2 data (tahun 2008 – 2009), *cluster* 1 sebanyak 7 data (tahun 2016 – 2022), dan *cluster* 2 sebanyak 6 data (tahun 2010 – 2015) dengan total keseluruhan ada sebanyak 15 data.

Dengan demikian ditarik kesimpulan pada tujuh tahun terakhir jumlah persentase tingkat pengangguran di Kota Palopo mengalami naik turunnya tingkat pengangguran. Dengan informasi ini, diharapkan dapat membantu pemerintah untuk mengetahui perkembangan tingkat pengangguran di Kota Palopo dari tahun ke tahun.

Daftar Pustaka

- [1] Maria Titah Jatiningrum, dkk. 2022. "Pengelompokan Kabupaten dan Kota di Provinsi Jawa Timur Berdasarkan Tingkat Kesejahteraan dengan Metode K-Means dan Density-Based Spatial Clustering Of Application With Noise." *Jurnal Derivat* Vol. 9, No. 1: 70-81. DOI : <https://doi.org/10.31316/j.derivat.v9i1.2832>.
- [2] Siti Aprizikiyandri, dkk. 2023. "Implementasi K-Means Cluster untuk Menentukan Persebaran Tingkat Pengangguran." *Empiricism Journal* Vol. 4, No. 2: 400-406. DOI : <https://doi.org/10.36312/ej.v4i2.1518>.
- [3] Candra Naya, Arif Siswandi. 2023. "IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKAN PENGANGGURAN TERBUKA DI INDONESIA DENGAN METODE CLUSTERING." *Jurnal Teknologi Pelita Bangsa* Vol. 14 No 2 : 99-104.
- [4] Sinta Kusuma Arum, dkk. 2024. "PENERAPAN ALGORITMA K-MEANS PADA DATASET PENGANGGURAN TERBUKA BERDASARKAN PENDIDIKAN DI PROVINSI JAWA BARAT." *JATI (Jurnal Mahasiswa Teknik Informatika)* Vol. 8 No. 2: 2221-2226. DOI : <https://doi.org/10.36040/Jati.V8i2.9440>.
- [5] Muhammad Nurfathullah, Intan Purnamasari. 2024. "Implementasi K-Means Untuk Mengelompokkan Provinsi di Indonesia Berdasarkan Indeks Jumlah Pengangguran Terbuka." *JATI (Jurnal Mahasiswa Teknik Informatika)* Vol. 8, No.2: 2277-2282. DOI : <https://doi.org/10.36040/jati.v8i2.9466>.
- [6] Nok Imas Pastia, Fatiha Nursari Dikananda. 2023. "Pengelompokan Data Pengangguran Terbuka Menggunakan Algoritma K-Means Berdasarkan Provinsi Jawa Barat." *Jurnal Dinamika Informatika* Vol. 12, No. 1: 59-69.
- [7] Rayhan Maliki, dkk. 2022. "Perbandingan Tingkat Pengangguran Terbuka Provinsi di Indonesia Berbasis Metode K-Means Clustering." *Computer Science (CO-SCIENCE)* Vol. 2, No. 2: 109-116. DOI: <https://doi.org/10.31294/coscience.v2i2.1151>
- [8] Fawaidul Badri, Anang Habibi. 2022. "Implementasi Metode K-Mean Clustering Analysis pada Pengelompokan Pengangguran di Indonesia sebagai Dampak dari Pandemi Covid-19." *Journal of Computer Science and Applied Informatics* Vol. 4, No. 2: 171-179. DOI : <https://doi.org/10.28926/ilkomnika.v4i2.471>
- [9] Aan Ansen Andryadi, Nisa. 2023. "PENERAPAN METODE ALGORITMA K-MEANS UNTUK DATA PENGANGGURAN DI JAWA BARAT." *Jurnal Wahana Informatika (JWI)* Vol. 2, No. 2: 277-283.
- [10] Falentino Sembiring, dkk. 2017. "ANALISIS PEMETAAN TINGKAT PENGANGGURAN DI PULAU JAWA DAN BALI DENGAN METODE K-MEANS." *Jurnal Rekayasa Teknologi Nusa Putra* Vol. 4, No.1: 13-18. DOI: <https://doi.org/10.52005/rekayasa.v4i1.142>

- [11] Akramunnisa, Fajriani. 2020. "K-MEANS Clustering Analysis Pada Persebaran Tingkat Pengangguran Kabupaten/Kota Di Sulawesi Selatan." *Jurnal Varian* Vol. 3, No. 2. doi:DOI 10.30812/varian.v3i2.652.
- [12] Fadhillah Azmi Tanjung, dkk. n.d. "Penerapan Metode K-Means Pada Pengelompokan Pengangguran di Indonesia." *Jurnal Riset Sistem Informasi dan Teknik Informatika (JURASIK)* Vol. 6, No. 1: 61-74. DOI : <http://dx.doi.org/10.30645/jurasik.v6i1.271>.
- [13] Muhammad Rais, dkk. 2021. "Optimalisasi K-Means Cluster dengan Principal Component Analysis pada Pengelompokan Kabupaten/Kota di Pulau Kalimantan Berdasarkan Indikator Tingkat Pengangguran Terbuka." *Jurnal EKSPONENSIAL* Vol. 12, No. 2: 129-136. DOI : <https://doi.org/10.30872/eksponensial.v12i2.805>.
- [14] Fatchi Arifatul Ulya, dkk. 2023. "Pengelompokan Tingkat Pengangguran Terbuka di Jawa Tengah Menggunakan Metode K-Means Clustering." *Journal of Data Insight* Vol. 1, No. 2: 71-80. DOI : <https://doi.org/10.26714/jodi.v1i2.148>.
- [15] Dewa Ayu Indah Cahya Dewi, Dewa Ayu Kadek Pramita. 2019. "Analisis Perbandingan Metode Elbow dan Sillhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali." *Jurnal Matrix* Vol. 9, No. 3: 102-109. doi:DOI : <https://dx.doi.org/10.31940/matrix.v9i3.1662>.



ZONAsi: Jurnal Sistem Informasi

is licensed under a [Creative Commons Attribution International \(CC BY-SA 4.0\)](https://creativecommons.org/licenses/by-sa/4.0/)