

## PENERAPAN ALGORITMA *K-MEANS* PADA DATASET PENGANGGURAN TERBUKA BERDASARKAN PENDIDIKAN DI PROVINSI JAWA BARAT

Sintia Kusuma Arum, Rini Astuti, Fadhil Muhammad Basysyar

Teknik Informatika, STMIK IKMI CIREBON

Jln, Perjuangan No.10 B Karyamulya Kesambi Kota Cirebon

Kusumasintia6@gmail.com

### ABSTRAK

Pengangguran merupakan istilah untuk orang yang tidak bekerja sama sekali, bekerja kurang dari dua hari selama satu minggu, atau seseorang yang berusaha mencari pekerjaan. Jumlah kasus Pengangguran Terbuka berdasarkan Tingkat Pendidikan Akhir adalah salah satu indikator penting dalam mengukur tingkat pengangguran di suatu Daerah. Dataset yang digunakan dalam penelitian ini merupakan dataset yang bersumber dari website resmi pemerintahan Open Data Jabar periode tahun 2019-2022. Algoritma *K-Means Clustering* telah diterapkan pada dataset pengangguran terbuka berdasarkan Pendidikan untuk mengelompokkan data berdasarkan tingkat Pendidikan akhir serta jumlah pengangguran provinsi yang terletak di Kota maupun dan di Kabupaten. Penelitian ini dilakukan bertujuan untuk mendapatkan beberapa hasil pengelompokan data Pengangguran di Provinsi Jawa Barat berdasarkan Tingkat Pendidikan Akhir. Pengelompokan tersebut, dibagi menjadi 3 Bagian yaitu tingkat pengangguran tinggi, tingkat pengangguran menengah dan tingkat pengangguran rendah, Tingkat Pengangguran tinggi diperoleh sebanyak 389 cluster, Tingkat Pengangguran Menengah dengan 162 cluster, dan Tingkat Pengangguran rendah dengan hasil 29 Cluster. Hasil dari penelitian ini diharapkan dapat memberikan wawasan penting untuk kebijakan pemerintah terkait masalah peningkatan angka pengangguran di berbagai daerah, serta menunjukan bahwa kemampuan Algoritma *K-Means* dalam mengelompokkan data pengangguran terbuka berdasarkan tingkat Pendidikan telah mendapatkan hasil K terbaik.

**Kata kunci :** Pengangguran terbuka, Jawa Barat, *K-Means*, Data Mining

### 1. PENDAHULUAN

Pengangguran dan Kemiskinan sering menjadi masalah utama yang umum di alami oleh Negara Berkembang termasuk di Indonesia, yang Sebagian wilayahnya masih tercatat dengan tingkat pengangguran tinggi. Masalah Pengangguran Terbuka memiliki pengaruh yang besar terhadap tingkat pertumbuhan ekonomi di Masyarakat, hal ini yang menghambat Pembangunan Daerah karena akan menimbulkan masalah-masalah sosial seperti, dapat mengurangi produktifitas, munculnya taraf hidup yang tidak merata, meningkatnya kemiskinan, kriminalitas, dan lain sebagainya. Masalah Pengangguran terbuka menjadi perhatian bagi pemerintah dan masyarakat untuk mengambil tindakan dan kebijakan yang tepat dalam menangani masalah pengangguran di daerah tersebut.

Metode *K-Means clustering* telah terbukti efektif dalam mengelompokkan data Tingkat Pengangguran Terbuka di daerah-daerah di Indonesia seperti Jawa barat, Bali dan Banten. Penggunaan Metode Algoritma *K-Means Clustering* pada Dataset Tingkat Pengangguran Terbuka (TPT) Berdasarkan Tingkat Pendidikan Akhir dapat membantu mengidentifikasi kelompok atau Cluster Tingkat Pendidikan yang lebih rentan mengalami pengangguran. Kelompok yang rentan merupakan Kelompok Pendidikan rendah atau tidak memiliki kualifikasi pendidikan yang memadai untuk memenuhi kebutuhan pasar kerja, Penduduk yang tinggal di daerah dengan lapangan pekerjaan yang terbatas atau kurang berkembang.

### 2. TINJAUAN PUSTAKA

Penelitian dilakukan sebelumnya oleh Mochamad Wahyudi, Lise Pujiastuti, dan Solikhun tahun 2020 dalam judul penelitiannya “Penerapan Data Mining Dalam Pengelompokan Data Pengangguran Terbuka Berdasarkan Provinsi Menggunakan Algoritma *K-Means*” dalam penelitiannya menyebutkan bahwa Pengangguran Terbuka merupakan Angkatan kerja yang sama sekali tidak mempunyai pekerjaan. Pengangguran ini Terjadi karena Angkatan kerja belum mendapatkan pekerjaan padahal sudah berusaha secara Maksimal, dan Sebagian lagi tergolong Angkatan kerja malas bekerja, serta pekerja dengan upah harian rendah. Dalam hal ini Pemerintah mengambil peran penting dalam Upaya mengatasi masalah tersebut. Penelitian ini bertujuan untuk mengklasifikasikan data pengangguran terbuka berdasarkan Provinsi tahun 2015. [1]

Penelitian yang kedua dilakukan oleh Fawaidul badri & Anang Habibi) Pada penelitian ini kita akan menganalisis hasil sebaran seluruh provinsi di Indonesia berdasarkan data tersebut. Mendiskusikan apakah hal ini dapat ditentukan. Dari data Tingkat Pengangguran Terbuka menggunakan analisis cluster. Algoritma yang digunakan *K-Means* menunjukan bahwa seluruh 34 provinsi di Indonesia terbagi menjadi 2 kelompok pengangguran, yaitu kelompok pertama dengan tingkat pengangguran tinggi, dan yang kedua adalah tingkat pengangguran rendah. Kelompok Tingkat pengangguran tinggi terdiri dari 10 kelompok, kelompok tingkat pengangguran menengah, dan

tingkat pengangguran rendah terdiri dari 24 hasil dari *K-Mean Clustering* yaitu nilai *Mean Square* kelompok 1 dengan standar pengangguran tinggi dan 2 dengan standar pengangguran rendah kurang dari 0,05. Sehingga *cluster* 1 dan 2 sangat baik dan terbukti dan hasil yang baik bisa diperoleh.. [2].

Penelitian berikutnya adalah Penelitian yang dilakukan oleh Anggia arfiani (2021) Penelitian ini membahas permasalahan pengangguran dengan menggunakan *indicator* Tingkat Pengangguran khususnya jumlah penduduk dan upah minimum Kabupaten/Kota dengan menggunakan metode *K-Means Clustering*. Tujuan penelitian ini adalah untuk memperoleh hasil pengelompokan Kabupaten/Kota di Provinsi Jawa Barat berdasarkan *indicator* tingkat pengangguran. [3]

Penelitian (N P Darshinni 2023) Penelitian ini menyebutkan bahwa pengangguran merupakan salah satu masalah yang dihadapi negara berkembang dan maju. Ketika tingkat pengangguran suatu negara mencapai tingkat yang tinggi, maka hal tersebut dapat menjadi masalah bagi negara tersebut. Pembayaran pajak pemerintah akan berkurang sehingga berdampak pada penurunan pendapatan negara. Di sisi lain, kualitas pekerjaan yang tidak memadai juga mempengaruhi tingkat pengangguran. Tujuan penelitian ini adalah untuk mendeskripsikan pola dan dampak pengangguran di Provinsi Jawa Barat, Indonesia pada tahun 2011 hingga 2022 dengan menggunakan metode *K-Means Clustering*. Hasil dari penelitian ini telah berhasil menerapkan Algoritma *K-Means Clustering* pada data Pengangguran Terbuka di Provinsi Jawa Barat dengan mengklasifikasikan data berdasarkan tingkat Pendidikan, tahun, dan kabupaten. [4]

Tujuan dari Penelitian ini adalah untuk Mengembangkan model pengelompokan data serta Menganalisis pola dan tren Pengangguran Terbuka di Provinsi Jawa Barat berdasarkan tingkat Pendidikan akhir dengan mempertimbangkan *measure type* berupa nilai *silhouette score* untuk mengevaluasi kluster, serta menentukan jumlah *maxruns* yang memungkinkan untuk mencapai konvergensi model dengan tingkat nilai yang baik dan optimal.

Metode Penelitian dalam penelitian ini penambangan data dilakukan dengan Teknik *Clustering* dengan metode *K-Means* Metode *K-Means* digunakan dalam penelitian ini merupakan suatu Teknik pengelompokan data yang berupaya membagi data menjadi dua kelompok atau lebih. Algoritma *k-means* salah satu dari algoritma *clustering* yang populer digunakan untuk mengelompokkan data berdasarkan karakteristik yang serupa. Kelompok data yang dihasilkan disebut *Cluster*.

Mampu Mengidentifikasi kelompok pengangguran terbuka berdasarkan tingkat pendidikan di Provinsi Jawa barat menggunakan *K-Means*, dengan *measure type* berupa tingkat kesesuaian (*goodness of fit*) antara data pengangguran terbuka dan hasil klusterisasi, serta menentukan jumlah *maxruns* yang

optimal untuk mencapai konvergensi dalam proses klusterisasi. *K-Means* juga merupakan Teknik pengelompokan data yang *nonhierarki* (terbatas). Sehingga data yang berkarakteristik sama akan masuk kedalam satu kelompok yang sama.

## 2.1. Data Mining

Data Mining adalah proses pengumpulan atau pemrosesan data dengan tujuan untuk mengekstraksi informasi darinya. Saat mengumpulkan dan mengekstrak informasi, kita dapat menggunakan perangkat lunak yang menggunakan perhitungan statistik dan teknologi kecerdasan buatan (AI). Istilah lain dari data mining adalah *Knowledge Discovery in Database* (KDD). [5]

## 2.2. Clustering

Clustering adalah proses membagi objek data menjadi beberapa kelompok yang biasa disebut *cluster*. Metode ini bermanfaat untuk mengidentifikasi kelompok yang tidak diketahui dalam data, seperti dalam bisnis *intelligence* yang dimana dapat mengelompokkan data pengangguran terbuka ke berbagai segmen serta mengelompokkan tingkat pengangguran di tiap daerah. [6]

## 2.3. K-Means

K-means merupakan suatu metode yang menganalisis data atau data mining tanpa supervisi yang memodelkan data melalui tahapan partisi untuk mengelompokkan data ke dalam kelompok – kelompok dimana setiap kelompok memiliki karakteristik yang serupa dan berbeda dengan kelompok lainnya. Tujuannya adalah meminimalkan variasi antar data dalam *cluster* dan memaksimalkan variasi antar *cluster*. [7]

## 2.4. KDD

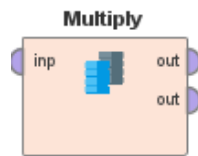
*Knowledge Discovery in Database* (KDD) adalah istilah ilmu komputer yang merujuk dalam proses penemuan informasi berharga dari kumpulan data digital besar, seperti database atau dataset. Proses KDD melibatkan beberapa tahapan, termasuk persiapan data, pemilihan data, pembersihan data, implementasi informasi pada dataset, dan interpretasi hasil. [8]

## 3. METODE PENELITIAN

Metode penelitian yang digunakan dapat dikategorikan sebagai kuantitatif. Kuantitatif ini merujuk pada pendekatan penelitian yang berfokus pada pengumpulan data analisis data berupa angka atau variabel kuantitatif untuk mendapatkan pemahaman tentang fenomena yang diteliti. Dalam hal ini, penggunaan metode *K-means clustering* menunjukkan pendekatan analisis statistik untuk mengelompokkan Data Pengangguran Terbuka di Jawa Barat, di pecah berdasarkan tingkat Pendidikan dan Kabupaten.



Dataset sebelumnya memiliki data yang hilang atau bernilai 0, Operator Filter Example diterapkan dalam menghilangkan data yang kosong agar dapat dilakukan proses pembagian data dengan operator *Multiply*.



Gambar 4. Operator *Multiply*.

Digunakan untuk membuat Salinan objek rapidminer. Operator ini mengambil objek Rapidminer dari port input dan mengirim Salinan objek ke port output. Setiap Port output yang terhubung akan membuat Salinan yang independen. Pada penelitian ini multiply akan digunakan untuk membuat cabang pada data yang nantinya akan digunakan operator correlation matrix. Operator Multiply membantu untuk membangun struktur proses Rapidminer yang lebih baik. Contoh penggunaannya adalah dengan mengalikan objek ExampleSet Titanic menjadi dua Salinan.

#### 4.3. Transformation

Data transformation digunakan untuk mengubah data dalam bentuk yang sesuai dalam proses data mining. Yang akan dipilih kedalam bentuk yang cocok untuk digunakan dalam algoritma yang digunakan, ialah algoritma *K-means*. Proses ini menggunakan operator “Nominal to Numeric” untuk mengubah atribut yang bersifat non-numerik menjadikan tipe numerik.

Row No	Survived	Sex	Age	Embarked	Class	Fare	Survived	Sex	Age	Embarked	Class	Fare	Survived	Sex	Age	Embarked	Class	Fare
1	0	Male	22	S	3	5.44	0	Male	22	S	3	5.44	0	Male	22	S	3	5.44
2	0	Male	38	S	3	5.13	0	Male	38	S	3	5.13	0	Male	38	S	3	5.13
3	0	Male	26	S	3	5.31	0	Male	26	S	3	5.31	0	Male	26	S	3	5.31
4	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10
5	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10
6	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10
7	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10
8	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10
9	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10
10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10
11	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10
12	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10
13	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10
14	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10
15	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10	0	Male	35	S	3	5.10

Gambar 5 . Hasil Operator Nominal to Numerik

Pada transformasi di atas menggunakan operator “Nominal to Numerical” karena algoritma *k-means* hanya bekerja pada data numerik, operator ini dapat mengubah atribut non – numerik kedalam representasi numerik, contohnya seperti teknik yang pengodean seperti pengubahan label menjadi nilai numerik sesuai dengan yang ditentukan. Melakukan transformasi ini, mengubah atribut yang semula non – numerik menjadi representasi numerik yang bisa digunakan oleh algoritma *k-means*, dengan menganalisis klusterisasi pada data yang diubah kedalam format yang sesuai.

#### 4.4. Pemodelan proses dirapidmaner

RapidMiner, sebagai platform analisis data yang canggih, melibatkan serangkaian tahapan esensial dalam proses pemodelan. Tahapan-tahapan ini membentuk kerangka kerja yang komprehensif, memungkinkan pengguna untuk mengelola dan menganalisis data dengan tujuan membangun model yang dapat memberikan wawasan berharga. Dapat dilihat pemodelan yang digunakan pada penelitian ini pada gambar 6 berikut.



Gambar 6. Pemodelan Rapidminer

Pada Tahapan yang tertera di atas merupakan keseluruhan dari pemodelan *Algoritma K-Means Clustering* dengan Rapidminer Studio, yang menggunakan beberapa tahapan atribut Pembaca Data (*Read Excel*), *Data Cleaning* membersihkan data (*Filter Example*), Membagi Data (*Multiply*), *Set Rule*, Mengubah menjadi *Numeric (Nominal to Numeric)*, Menyeleksi atribut yang di butuhkan (*Select Attribute*), *Clustering*. Sementara dalam pembagian data untuk menentukan Korelasi hanya digunakan 3 operator yaitu Nominal to Numeric, Correlation matrix dan Terakhir untuk menampilkan Hasil menggunakan *Performance* dengan perolehan percobaan parameter *K*

#### 4.5. Data Mining

Pada tahap ini menggunakan operator Clustering yang akan menentukan cluster atau pengelompokan data. Dalam penelitian ini memakai algoritma *k-means* serta menggunakan 3 cluster dengan maxrun 6 perulangan. Supaya menghasilkan nilai *k* yang optimal, dengan melakukan 6 kali perulangan. Dari 6 kali perulangan, nilai *k* yang optimal dihasilkan pada *k* = 3 dengan hasil DBI nya yaitu 0.487

Tabel 1. Hasil DBI

No	Cluster	Davies Bouldi Index
1	K2	0.555
2	K3	0.487
3	K4	0.489
4	K5	0.500
5	K6	0.512
6	K7	0.501
7	K8	0.504
8	K9	0.493
9	K10	0.533

#### 4.6. Evaluation / Interpretation



Gambar 7. Bagan Cluster

Pada hasil pengelompokan dalam penelitian ini untuk menganalisis Tingkat Pengangguran Terbuka Berdasarkan Tingkat Pendidikan Akhir.

- Cluster 0 adalah tingkat pengangguran tinggi berdasarkan pada Pendidikan Tingkat akhir yang di dominasi oleh SD(Kebawah),DIPLOMA I-III dan SMA UMUM, dengan Mayoritas wilayah Kabupaten Bogor,Kabupaten Sukabumi,Kabupaten Bandung.Sedangkan wilayah Cianjur, Kuningan, Cirebon Pengangguran terbuka berdasarkan tingkat Pendidikan SD Kebawah,(SMP) dan SMA UMUM dan KEJURUAN .
- Cluster 1 adalah Tingkat pengangguran terbuka rendah pada Tingkat Pendidikan Akhir yang di dominasi oleh tingkat Pendidikan akhir SMP,SMA UMUM,SD (KEBAWAH), dan SMA kejuruan pada wilayah BANJAR,PANGANDARAN dan INDRAMAYU. Sedangkan CIANJUR dan CIMAHI meliputi lulusan Pendidikan akhir SMA kejuruan,SMA umum,SD (KEBAWAH),DIPLOMA.
- Cluster 2 adalah dengan tingkat pengangguran terbuka Menengah tinggi berdasarkan Tingkat Pendidikan Akhir yang terdiri dari SMA umum,SMA Kejuruan,SD (KEBAWAH),Diploma I II III,SMA UMUM,KEJURUAN,SMP SMP yang berada di Kabupaten Bogor, Kabupaten Sukabumi, Kabupaten Cianjur, Kabupaten Bandung, Kabupaten Garut, Kabupaten Tasikmalaya, Kabupaten Ciamis, Kabupaten Kuningan, Kabupaten Cirebon, Kabupaten Majalengka, Kabupaten Sumedang, Kabupaten Indramayu.

Berdasarkan cluster. Diketahui jumlah Pengangguran Terbuka Berdasarkan Pendidikan akhir paling tinggi pada cluster\_0 dengan jumlah 389 item cluster\_1 menengah dengan 162 item,dan cluster 2 merupakan cluster terendah yaitu 29 item.Dan terdapat hasil dari centroid table berikut.

	cluster_0	cluster_1	cluster_2
center_x	11.000	2.070	11.000
center_y	3.000	3.000	3.000
center_x2	11.000	2.070	11.000
center_y2	3.000	3.000	3.000
center_x3	11.000	2.070	11.000
center_y3	3.000	3.000	3.000

Gambar 8 . Hasil Centroid Table

Pada penelitian ini membahas mengenai Pengelompokan Dataset Pengangguran Terbuka di Jawa Barat Berdasarkan tingkat Pendidikan akhir beserta daerah yang terdampak,muali dari tingkat tertinggi,menengah dan terendah. Proses penelitian dilakukan untuk mengetahui cluster mana yang terdapat di paling tinggi dan rendah dalam menentukan jumlah pengangguran tertinggi.proses agar dapat mengetahui hasil tersebut dengan mengelompokkan data dalam metode clustering menggunakan k-means.



Gambar 9. Hasil Davies Bouldin Index

Dari banyak nya pengujian parameter K, Penulis dapat mengetahui cluster berapa yang digunakan dalam penelitian ini ialah dengan menghitung nilai DBI dan dapat diketahui hasil nilai DBI terdapat pada K3 (klaster 3).

## 5. KESIMPULAN DAN SARAN

Penerapan metode clustering dalam data mining efektif digunakan dalam penelitian dataset pengangguran terbuka untuk mengidentifikasi pola dan hubungan antara variabel-variabel yang diteliti. Algoritma *K-Means* dapat digunakan untuk mengelompokkan wilayah berdasarkan indeks Pendidikan di provinsi Jawa Barat.

Implementasi algoritma *K-Means* dapat memetakan jumlah pengangguran di provinsi Jawa Barat berdasarkan tingkat Pendidikan akhir.Penelitian ini menunjukkan hubungan antara tingkat Pendidikan,dan angka pengangguran,dimana semakin tinggi tingkat Pendidikan,semakin rendah angka pengangguran.

Hasil penelitian menggambarkan data pengangguran terbuka Provinsi Jawa Barat tahun 2013 hingga 2022 menggunakan algoritma *K-Means* data pengangguran terbuka tahun 2013-2022 menurut provinsi Jawa barat menggunakan Agoritma *K-Means*.Nilai K = 3 memberikan cluster terbaik.Hasil clustering divalidasi menggunakan Davies-Bouldin Index (DBI) dengan nilai 0,487.Eksperimen K-Means dengan nilai K = 3 dibagi 3 cluster,0 sebanyak 389 ,cluster 1 dengan 29 item, dan cluster 3 adalah 162 anggota.

Saran untuk penelitian berikutnya tentang tingkat pengangguran terbuka berdasarkan pendidikan di Jawa Barat dengan K-Means clustering  $k=3$  dapat mencakup pengumpulan data yang lebih lengkap dan terbaru, penggunaan metode validasi yang lebih baik seperti Davies Bouldin Index (DBI), dan analisis yang lebih mendalam terhadap hasil pengelompokkan. Selain itu, penelitian selanjutnya dapat mempertimbangkan faktor-faktor lain yang mempengaruhi tingkat pengangguran, seperti kondisi ekonomi, demografi, dan geografis. Studi sebelumnya telah menggunakan algoritma K-Means clustering untuk memetakan jumlah pengangguran di Provinsi Jawa Barat berdasarkan tingkat pendidikan dengan nilai  $k=3$ . Oleh karena itu, penelitian selanjutnya dapat memperluas penggunaan metode ini untuk memetakan tingkat pengangguran terbuka berdasarkan pendidikan di Jawa Barat dengan  $k=3$  dan membandingkan hasilnya dengan studi sebelumnya.

#### DAFTAR PUSTAKA

- [1] Azzahro, Z. (2022). Implementasi Metode K-Means Clustering Untuk Pengelompokan Data Kemiskinan Di Jawa Timur (Doctoral dissertation, Universitas Muhammadiyah Malang).
- [2] Fajriani, F 2020, '*K-Means Clustering* analysis pada Persebaran Tingkat pengangguran kabupaten/kota di Sulawesi selatan', *Jurnal Varian*, [journal.universitasbumigora.ac.id](http://journal.universitasbumigora.ac.id), <https://journal.universitasbumigora.ac.id/index.php/Varian/article/view/652>
- [3] Harits, D, Puji, AA, Andivas, M, Dermawan, D, & ... 2022, 'Analisis Klaster Tingkat Pengangguran di Provinsi Kalimantan Timur Menggunakan Algoritma K-Means', *Jurnal Surya* ..., [ejurnal.umri.ac.id](http://ejurnal.umri.ac.id), <https://ejurnal.umri.ac.id/index.php/JST/article/view/4430>
- [4] Maliqi, R, Falgenti, K, Priani, S, Fithri, F, & ... 2022, 'Penerapan Metode *K-Means Clustering* pada Data Tingkat Pengangguran Terbuka Tahun 2016-2018 dan 2019-2021', *Computer Science* ..., [eprints.bsi.ac.id](http://eprints.bsi.ac.id), <http://eprints.bsi.ac.id/index.php/com-science/article/view/1151>
- [5] Megawaty, M. (2022). Algoritma partitional clustering untuk pengklasifikasian tingkat pengangguran di kabupaten lahut. *Algoritma partitional clustering untuk pengklasifikasian tingkat pengangguran di kabupaten lahut*.
- [6] Muharni, S, & Andriyanto, S 2022, 'Ditarik: Penerapan Metode *K-Means Clustering* Pada Data Tingkat Pengangguran Terbuka Tahun 2016-2018 Dan 2019-2021', *Jurnal Informatika*, [jurnal.darmajaya.ac.id](http://jurnal.darmajaya.ac.id), <https://jurnal.darmajaya.ac.id/index.php/JurnalInformatika/article/view/3172>
- [7] Muharni, S, Andriyanto, S, No, JK, & Metro, M 2022, 'Penerapan Metode *K-Means Clustering* pada Data Tingkat Pengangguran Terbuka Tahun 2016-2018 dan 2019-2021', *J. Inform*, [academia.edu](http://academia.edu), <https://www.academia.edu/download/90650381/pdf.pdf>
- [8] Naya, C, & Siswandi, A 2023, 'Implementasi Data Mining Untuk Pengelompokan Pengangguran Terbuka di Indonesia Dengan Metode Clustering', *Jurnal SIGMA*, [jurnal.pelitabangsa.ac.id](http://jurnal.pelitabangsa.ac.id), <http://jurnal.pelitabangsa.ac.id/index.php/sigma/article/view/2577>
- [9] Pastia, NI, & Dikananda, FN 2023, 'Pengelompokan Data Pengangguran Terbuka Menggunakan Algoritma K-Means Berdasarkan Provinsi Jawa Barat', *Jurnal Dinamika Informatika*, [jdi.upy.ac.id](http://jdi.upy.ac.id), <https://jdi.upy.ac.id/index.php/jdi/article/view/223>
- [10] Putrie, A. A., & Sanjaya, R. (2021). Pengelompokan Kabupaten/Kota Berdasarkan Indikator Tingkat Pengangguran Menggunakan Algoritma K-Means Clustering (Studi Kasus: Provinsi Jawa Barat). *eProsiding Sistem Informasi (POTENSI)*, 2(2), 111-121.
- [11] Rais, M, Goejantoro, R, & Prangga, S 2021, '... K-Means Cluster dengan Principal Component Analysis pada Pengelompokan Kabupaten/Kota di Pulau Kalimantan Berdasarkan Indikator Tingkat Pengangguran ...', *EKSPONENSIAL*, [jurnal.fmipa.unmul.ac.id](http://jurnal.fmipa.unmul.ac.id), <http://jurnal.fmipa.unmul.ac.id/index.php/exponensial/article/view/805>
- [12] Sarbaini, S, Saputri, W, & Muttakin, F 2022, 'Cluster Analysis Menggunakan Algoritma Fuzzy K-Means Untuk Tingkat Pengangguran Di Provinsi Riau', *Jurnal Teknologi Dan ...*, [jurnal-tmit.com](http://jurnal-tmit.com), <http://jurnal-tmit.com/index.php/home/article/view/30>
- [13] Tanjung, FA, Windarto, AP, & ... 2021, 'Penerapan Metode K-Means Pada Pengelompokan Pengangguran Di Indonesia', *Jurasik (Jurnal Riset ...)*, [tunasbangsa.ac.id](http://tunasbangsa.ac.id), <http://tunasbangsa.ac.id/ejurnal/index.php/jurasik/article/view/271>
- [14] Wahyudi, M, Pujiastuti, L, & Solikhun, S 2020, 'Penerapan Data Mining Dalam Mengelompokkan Data Pengangguran Terbuka Menurut Provinsi Menggunakan Algoritma K-Means', ... *Seminar Nasional Riset ...*, [tunasbangsa.ac.id](http://tunasbangsa.ac.id), <http://tunasbangsa.ac.id/seminar/index.php/senaris/article/view/192>