

**PENERAPAN METODE *K-MEANS* UNTUK *CLUSTERING* TINGKAT
KEJAHATAN DI INDONESIA**

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Disusun Oleh
YUVRICO BAGUS FERDIANSYAH
H92219063

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA**

2023

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : YUVRICO BAGUS FERDIANSYAH

NIM : H92219063

Program Studi : Matematika

Angkatan : 2019

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "PENERAPAN METODE *K-MEANS* UNTUK *CLUSTERING* TINGKAT KEJAHATAN DI INDONESIA". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 11 Juli 2023

Yang menyatakan,



YUVRICO BAGUS FERDIANSYAH
NIM. H92219063

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

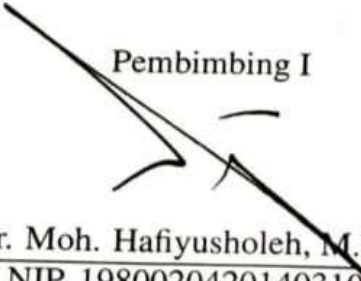
Nama : YUVRICO BAGUS FERDIANSYAH

NIM : H92219063


Judul skripsi : PENERAPAN METODE *K-MEANS* UNTUK
CLUSTERING TINGKAT KEJAHATAN DI INDONESIA

telah diperiksa dan disetujui untuk diujikan.

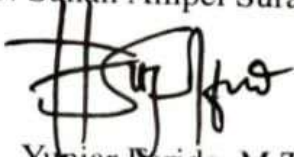
Pembimbing I


Dr. Moh. Hafiyusholeh, M.Si., M.Pmat.
NIP. 198002042014031001

Pembimbing II


Lutfi Hakim, M.Ag.
NIP. 197312252006041001

Mengetahui,
Ketua Program Studi Matematika
UIN Sunan Ampel Surabaya


Yuniar Parida, M.T
NIP. 197905272014032002

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh

Nama : YUVRICO BAGUS FERDIANSYAH
NIM : H92219063
Judul Skripsi : PENERAPAN METODE *K-MEANS* UNTUK
CLUSTERING TINGKAT KEJAHATAN DI INDONESIA

Telah dipertahankan di depan Tim Penguji
pada tanggal 11 Juli 2023

Mengesahkan,
Tim Penguji

Penguji I



Aris Fanani, M.Kom
NIP. 198701272014031002

Penguji II



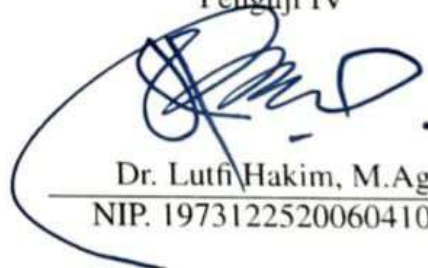
Wika Dianita Utami, M.Sc
NIP. 199206102018012003

Penguji III



Dr. Moh. Hafiyusholeh, M.Si., M.Pmat.
NIP. 198002042014031001

Penguji IV



Dr. Lutfi Hakim, M.Ag.
NIP. 197312252006041001

Mengetahui,

Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya



Dr. A. Saepul Hamdani, M.Pd
NIP. 196507312000031002



UIN SUNAN AMPEL
SURABAYA

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax. 031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Yuvrico Bagus Ferdiansyah
NIM : 1192219063
Fakultas/Jurusan : Sains dan Teknologi/Matematika
E-mail address : h92219063@student.uinsby.ac.id

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

Skripsi Tesis Disertasi Lain-lain (.....)
yang berjudul :

PENERAPAN METODE K-MEANS UNTUK CLUSTERING TINGKAT KEJAHATAN DI
INDONESIA

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara **fulltext** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 14 Juli 2023

Penulis

(YUVRICO BAGUS FERDIANSYAH)

ABSTRAK

PENERAPAN METODE *K-MEANS* UNTUK *CLUSTERING* TINGKAT KEJAHATAN DI INDONESIA

Jumlah kejahatan di Indonesia sebesar 247.218 kasus pada tahun 2020. Hal ini mengakibatkan kejahatan di Indonesia masih tergolong tinggi. Kasus kejahatan tersebut memiliki beragam motif seperti kejahatan terhadap nyawa, kejahatan terhadap fisik, kejahatan terhadap korupsi, dll. Tingginya kasus kejahatan di Indonesia berakibat perlunya *pengclustering* tingkat kejahatan guna mengelompokkan provinsi di Indonesia yang rawan terhadap kasus kejahatan. Tujuan penelitian ini adalah untuk *mengcluster* tingkat kejahatan di Indonesia. Metode *K-Means* adalah metode clustering yang dapat diterapkan dalam *clustering* tingkat kejahatan di Indonesia. Keunggulan dari metode *K-Means* diantaranya keakuratan dalam *clustering* cukup baik, mampu mengelompokkan data dengan kuantitas besar dalam waktu singkat, dan relatif mudah karena jumlah *cluster* dan pusat *cluster* dapat diinisialisasi secara acak. Data yang digunakan dalam penelitian ini adalah data klasifikasi kejahatan di Indonesia pada tahun 2020. Berdasarkan perhitungan, diperoleh *cluster* optimal sebesar $k = 3$ dengan kategori *cluster* aman sebanyak 22 anggota, *cluster* rawan sebanyak 11 anggota, dan *cluster* sangat rawan sebanyak 1 anggota. Dimana *cluster* $k = 3$ didapat nilai *silhouette coefficient* (SC) sebesar 0,77. Hal ini mengartikan bahwa hasil *cluster* memiliki struktur yang kuat.

Kata kunci: *Kejahatan, Clustering, K-Means*

ABSTRACT

APPLICATION OF THE K-MEANS METHOD FOR CLUSTERING CRIME LEVELS IN INDONESIA

The number of crimes in Indonesia is 247,218 cases in 2020. This has resulted in crimes in Indonesia still being relatively high. These criminal cases have various motives such as crimes against life, crimes against physical bodies, crimes against corruption, etc. The high number of crime cases in Indonesia results in the need for clustering crime rates in order to classify provinces in Indonesia that are prone to crime cases. The purpose of this study is to cluster crime rates in Indonesia. The K-Means method is a clustering method that can be applied in crime rate clustering in Indonesia. The advantages of the K-Means method include fairly good accuracy in clustering, being able to group large quantities of data in a short time, and being relatively easy because the number of clusters and cluster centers can be initialized randomly. The data used in this study is data on the classification of crimes in Indonesia in 2020. Based on calculations, an optimal cluster of $k = 3$ is obtained with a safe cluster category of 22 members, a vulnerable cluster of 11 members, and a very vulnerable cluster of 1 member. Where the cluster $k = 3$ obtained a silhouette coefficient (SC) value of 0.77. This means that the cluster results have a strong structure.

Keywords: *Crime, Clustering, K-Means*

DAFTAR ISI

| | |
|---|------|
| HALAMAN JUDUL | i |
| LEMBAR PERSETUJUAN PEMBIMBING | ii |
| PENGESAHAN TIM PENGUJI SKRIPSI | iii |
| HALAMAN PERNYATAAN KEASLIAN | iv |
| MOTTO | v |
| HALAMAN PERSEMBAHAN | vi |
| KATA PENGANTAR | vii |
| DAFTAR ISI | ix |
| DAFTAR TABEL | xi |
| DAFTAR GAMBAR | xii |
| ABSTRAK | xiii |
| ABSTRACT | xiv |
| I PENDAHULUAN | 1 |
| 1.1. Latar Belakang Masalah | 1 |
| 1.2. Rumusan Masalah | 6 |
| 1.3. Tujuan Penelitian | 7 |
| 1.4. Manfaat Penelitian | 7 |
| 1.5. Batasan Masalah | 7 |
| 1.6. Sistematika Penulisan | 8 |
| II TINJAUAN PUSTAKA | 9 |
| 2.1. Kejahatan di Indonesia | 9 |
| 2.2. <i>Clustering</i> | 12 |
| 2.3. <i>Uji Kaiser Meyer Olkin Measur of Sampling (KMO)</i> | 14 |
| 2.4. <i>Principal Component Analysis (PCA)</i> | 16 |
| 2.5. <i>K-Means</i> | 17 |
| 2.6. <i>Shilhouette Coefficient</i> | 19 |
| 2.7. Integrasi Keilmuan | 21 |

| | |
|---|-----------|
| III METODE PENELITIAN | 28 |
| 3.1. Jenis Penelitian | 28 |
| 3.2. Sumber Data | 28 |
| 3.3. Teknik Analisis Data | 29 |
| IV HASIL DAN PEMBAHASAN | 31 |
| 4.1. Analisis Deskriptif Data | 31 |
| 4.2. <i>Pre Processing Data</i> | 33 |
| 4.3. <i>K-Means Clustering</i> | 35 |
| 4.4. Perbandingan Hasil $k = 2$ dan $k = 3$ | 47 |
| 4.5. Integrasi Keilmuan | 47 |
| V PENUTUP | 50 |
| 5.1. Kesimpulan | 50 |
| 5.2. Saran | 51 |
| DAFTAR PUSTAKA | 51 |



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR TABEL

| | | |
|------|--|----|
| 2.1 | Kriteria nilai uji KMO | 15 |
| 2.2 | <i>Silhouette Coefficient</i> | 21 |
| 4.1 | Data Kejahatan di Indonesia | 31 |
| 4.2 | Deskripsi Data Kejahatan di Indonesia | 32 |
| 4.3 | Hasil uji KMO | 33 |
| 4.4 | Matriks Kovarian | 33 |
| 4.5 | Nilai Eigen | 34 |
| 4.6 | <i>Rotasi Faktor</i> | 35 |
| 4.7 | Data Kejahatan di Indonesia yang Telah di Reduksi dengan PCA | 36 |
| 4.8 | <i>Centroid Awal</i> | 37 |
| 4.9 | Jarak <i>Euclidean</i> Iterasi ke-1 | 37 |
| 4.10 | <i>Centroid Baru</i> Iterasi Ke-2 | 39 |
| 4.11 | Jarak <i>Euclidean</i> Iterasi ke-2 | 39 |
| 4.12 | <i>Centroid Baru</i> Iterasi Ke-3 | 40 |
| 4.13 | Jarak <i>Euclidean</i> Iterasi ke-3 | 40 |
| 4.14 | <i>Centroid Baru</i> Iterasi Ke-7 | 42 |
| 4.15 | Jarak <i>Euclidean</i> Iterasi ke-7 | 42 |
| 4.16 | Hasil Pembulatan <i>Centroid</i> Iterasi ke-7 | 43 |
| 4.17 | Tabel Keanggotaan Hasil 2 <i>Cluster</i> | 44 |
| 4.18 | Hasil Pembulatan <i>Centroid</i> Iterasi ke-2 | 45 |
| 4.19 | Tabel Keanggotaan Hasil 3 <i>Cluster</i> | 46 |
| 4.20 | Perbandingan Keakuratan Hasil 2 dan 3 <i>Cluster</i> | 47 |

DAFTAR GAMBAR

| | |
|---|----|
| 3.1 Diagram Alur Penelitian | 30 |
| 4.1 Pemetaan <i>Clustering</i> Tingkat Kejahatan di Indonesia dengan 2 <i>Cluster</i> | 44 |
| 4.2 Pemetaan <i>Clustering</i> Tingkat Kejahatan di Indonesia dengan 3 <i>Cluster</i> | 46 |



UIN SUNAN AMPEL
S U R A B A Y A

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Indonesia merupakan negara yang mempunyai banyak agama, ras, suku, dan beragam budaya, sehingga tingkat populasi di Indonesia sangat tinggi. Pada tahun 2022 tercatat Indonesia menempati ranking 4 dunia dengan tingginya populasi penduduk yaitu sebesar 3.39% dari penduduk dunia atau sejumlah 271 juta jiwa. Provinsi Jawa Barat memiliki jumlah penduduk tertinggi yaitu sebanyak 48 juta jiwa (Antara and Yogantari, 2018).

Tingginya jumlah penduduk di Indonesia menyebabkan munculnya masalah sosial yang ada di masyarakat seperti: pengangguran, kemiskinan, dan kesenjangan sosial. Masalah sosial ini bisa menjadi pemicu beberapa orang untuk melakukan tindak kejahatan. Beberapa contoh klasifikasi tindak kejahatan adalah kejahatan terhadap fisik, kesusilaan, nyawa, kemerdekaan orang, hak milik, narkoba, dan korupsi. Selain disebabkan oleh masalah sosial, kejahatan juga dapat terjadi dikarenakan ada kesempatan atau kelengahan dari korban ataupun memiliki masalah pribadi seperti dendam, iri, dan lainnya. Hal ini tentu sangat mengganggu ketertiban umum (Astuti et al., 2016).

Islam merupakan agama mayoritas penduduk Indonesia memberikan petunjuk untuk selalu berbuat kebaikan dan menghindari perbuatan kejahatan. Allah berfirman di dalam Al-Qur'an surat Al-Isra' ayat 7 yang berbunyi:

إِنْ أَحْسَنْتُمْ أَحْسَنْتُمْ لِأَنْفُسِكُمْ وَإِنْ أَسَأْتُمْ فَلَهَا

Artinya: "Jika kamu berbuat baik (berarti) kamu berbuat baik untuk dirimu sendiri. Dan jika kamu berbuat jahat, maka (kerugian kejahatan) itu untuk dirimu sendiri."

Begitu jelas petunjuk Allah SWT untuk selalu berbuat kebaikan. Petunjuk tersebut sekaligus memperingatkan tentang larangan bagi orang yang melakukan tindakan kejahatan. Selain petunjuk dan peringatan dari Allah, Rasulullah SAW juga telah mengingatkan melalui hadits berikut:

حَدَّثَنَا عَبْدُ اللَّهِ بْنُ مَسْلَمَةَ بْنِ قَعْنَبٍ: حَدَّثَنَا دَاوُدُ، يَعْنِي ابْنَ قَيْسٍ، عَنْ أَبِي سَعِيدٍ، مَوْلَى عَامِرِ بْنِ كُرَيْزٍ، عَنْ أَبِي هُرَيْرَةَ قَالَ: قَالَ رَسُولُ اللَّهِ ﷺ: (لَا تَحَاسَدُوا، وَلَا تَتَّاجِسُوا، وَلَا تَبَاغَضُوا، وَلَا تَدَابَرُوا، وَلَا يَبِعْ بَعْضُكُمْ عَلَى بَيْعِ بَعْضٍ، وَكُونُوا عِبَادَ اللَّهِ إِخْوَانًا). الْمُسْلِمُ أَخُو الْمُسْلِمِ، لَا يَظْلِمُهُ، وَلَا يَحْذَلُهُ، وَلَا يَحْقِرُهُ. التَّقْوَى هَا هُنَا (وَيُشِيرُ إِلَى صَدْرِهِ ثَلَاثَ مَرَّاتٍ): بِحَسَبِ أَمْرٍ مِنَ الشَّرِّ أَنْ يَحْقِرَ أَخَاهُ الْمُسْلِمَ. كُلُّ الْمُسْلِمِ عَلَى الْمُسْلِمِ حَرَامٌ: دَمُهُ

Artinya: "Abdullah bin Maslamah bin Qa'nab telah menceritakan kepada kami: Dawud bin Qais menceritakan kepada kami, dari Abu Sa'id maula 'Amir bin Kuraiz, dari Abu Hurairah. Beliau berkata: Rasulullah shallallahu 'alaihi wa sallam bersabda, "Janganlah kalian saling dengki, janganlah kalian melakukan tanajusy (menawar dengan harga yang lebih tinggi oleh orang yang tidak hendak membelinya untuk menaikkan harganya), janganlah kalian saling membenci, janganlah kalian saling bermusuhan, janganlah sebagian kalian membeli barang yang hendak dibeli orang lain, dan jadilah kalian hamba-hamba Allah yang bersaudara. Seorang muslim adalah saudara bagi muslim lainnya. Dia tidak menzaliminya, tidak mengabaikannya, dan dia tidak menghina. Takwa itu ada di sini." Beliau menunjuk ke dada sebanyak tiga kali. "Cukup orang itu dikatakan jelek dengan sebab ia menghina saudaranya yang muslim. Setiap muslim terhadap muslim lainnya haram darahnya, hartanya, dan kehormatannya." (Sahih Muslim,

no. hadits: 2564)

Untuk mencegah terjadinya tindakan kejahatan negara Indonesia sudah mengatur undang-undang terkait kejahatan. Salah satu contoh pada pasal 351 Kitab Undang-Undang Hukum Pidana (KUHP), yang menyatakan bahwa penganiayaan diancam dengan pidana penjara paling lama dua tahun delapan bulan atau pidana denda paling banyak empat ribu lima ratus rupiah (Surbakti & Natangsa, 2017).

Menurut data dari Badan Pusat Statistik (BPS), terdapat 294.281 total kejadian kejahatan pada tahun 2018, 269.324 kejadian kejahatan pada tahun 2019, dan 247.218 kejadian kejahatan pada tahun 2020. Frekuensi kejahatan di Indonesia cenderung menurun dari tahun 2018-2020. Meskipun frekuensi kejahatan setiap tahun mengalami penurunan bukan berarti kasus kejahatan terselesaikan, karena angka kejahatannya masih tinggi. Pada tahun 2020 provinsi dengan kasus kejahatan tertinggi yaitu provinsi Sumatera Utara dengan total 30.320 kasus kejahatan.

Contoh kasus kejahatan yang terjadi pada tahun 2020 yaitu Seorang oknum pembina pramuka yang berasal dari Sumatera Selatan melakukan tindakan pelecehan seksual dan pembunuhan terhadap siswi SMP. Tersangka tertarik kepada korban karena memiliki wajah yang cantik. Korban mendapatkan pesan dari tersangka untuk melakukan kegiatan pramuka di sekolah. Bukannya melakukan kegiatan pramuka, namun tersangka justru melakukan tindakan asusila dan pembunuhan. Hal tersebut tergolong ke dalam kategori kejahatan terhadap nyawa dan kesusilaan (KOMPAS, 2020).

Kasus penculikan di Medan Sumatera Utara yang digagalkan oleh masyarakat setempat. Awalnya korban sedang bermain di depan rumahnya,

tiba-tiba ada seorang wanita yang tak dikenalnya menghampiri dan memaksa korban untuk ikut pergi bersamanya. Korban berteriak dengan kencang yang mengakibatkan masyarakat setempat menghampiri suara teriakan korban. Akhirnya tersangka pun diamankan oleh Babinsa dan dibawa ke polsek setempat. Kasus tersebut termasuk ke dalam kategori kejahatan terhadap kemerdekaan orang ([ANTARA SUMUT, 2020](#)).

Pada tahun 2023 kejahatan masih banyak terjadi di Indonesia contoh kasus kejahatan yang sedang ramai diperbincangkan saat ini yaitu liarnya gangster di daerah Sidoarjo. Peristiwa tersebut diawali dengan tantangan tawuran antar kelompok yaitu kelompok Sidoarjo *Brawl*, RGS 21 dan T3Heroes. Dari tantangan tersebut akhirnya ketiga kelompok gangster ini tawuran di Desa Sepande, Kecamatan Candi, Kabupaten Sidoarjo. Akibatnya 10 orang tewas dengan bersimbah darah dalam aksi tawuran tersebut. Hal ini termasuk ke dalam kategori kejahatan terhadap nyawa ([Rinanda, 2023](#)).

Kasus Mario Dandy Satrio yang berasal dari Sumatera Utara dan merupakan anak dari pejabat yang melakukan penganiayaan terhadap David Ozora. Awalnya Mario Dandy mendapatkan informasi bahwa David melecehkan AG. AG sendiri adalah pacar dari Mario. Oleh karena itu Mario emosi dan melakukan penganiayaan terhadap David. Kasus ini tergolong kedalam kategori kejahatan terhadap fisik ([detikcom, 2023](#)).

Selanjutnya kasus korupsi oleh Surya Damadi Pemilik PT Darmex Group didakwa melakukan tindak pidana pencucian uang (TPPU) dan korupsi dalam kegiatan usaha perkebunan kelapa sawit yang mengakibatkan kerugian keuangan negara mencapai Rp 80 triliun. Surya didakwa melanggar Pasal 2 ayat 1 atau Pasal 3 Jo Pasal 18 Undang-undang Pemberantasan Tindak Pidana Korupsi (UU Tipikor)

Jo Pasal 55 ayat 1 ke 1 KUHP. Kemudian Pasal 3 ayat 1 huruf c UU Tindak Pidana Pencucian Uang serta Pasal 3 atau Pasal 4 UU Pencegahan dan Pemberantasan TPPU. Kasus ini termasuk kedalam kategori kejahatan penipuan, penggelapan dana, dan korupsi (CNN Indonesia, 2022).

Peringatan, larangan, undang-undang telah dibuat sedemikian rupa untuk mencegah terjadinya tindak kejahatan, tetapi dari tahun ke tahun selalu ada tindakan kejahatan, oleh karena itu perlu dilakukan *clustering* untuk melihat tingkat kejahatan dari setiap provinsi di Indonesia. Dengan diketahuinya tingkat kejahatan disetiap provinsi, tentu pemerintah dapat langsung bergegas untuk melakukan tindakan preventif disegala sektor. Khususnya mewaspadai daerah-daerah yang masih tinggi tingkat kejahatannya (Pradana, 2019).

Terdapat beberapa metode dalam melakukan clustering diantaranya *Probabilistic Distance Clustering* (PDC), *K-Means Clustering*, dan *Fuzzy C-Means* (FCM). Tujuan dari *clustering* yaitu membagi data yang tidak mempunyai label menjadi beberapa *cluster*, sehingga dapat memudahkan untuk menganalisis data jika sudah dikelompokkan (Atma et al., 2021). Hal tersebut tentu sangat membantu peran pemerintahan dan juga aparat keamanan untuk melakukan analisis kejadian kejahatan dan menyusun strategi dalam upaya penertiban dan mencegah kejadian kejahatan di Indonesia.

Pada penelitian ini, peneliti menggunakan metode *K-Means clustering*. Penggunaan metode *K-Means clustering* didasari pada keakuratan dalam *clustering* yang cukup baik, mampu mengelompokkan data dengan kuantitas besar dalam waktu relatif singkat, dan relatif lebih mudah digunakan karena jumlah *cluster* dan pusat *cluster* dapat diinisialisasi secara acak (Farida et al., 2022).

Beberapa penelitian yang berkaitan dengan *clustering* dilakukan oleh

Oktaviana (2022) menggunakan metode *K-Means clustering* didapatkan hasil yang baik karena memiliki nilai *shilhouette index* = 0.501 dengan *cluster* optimal sebanyak 3 yaitu rawan, sedang, dan tidak rawan. Selanjutnya penelitian oleh Sholihah (2021) menggunakan metode *K-Means clustering* didapatkan hasil yang baik, karena *shilhouette index* = 0,74 dengan jumlah klaster sebanyak 2 yaitu Aman dan Rawan. Selain itu, penelitian oleh Suryadi and Supriatna (2019) dengan judul “Sistem *Clustering* Tindak Kejahatan Pencurian di Wilayah Jawa Barat Menggunakan Algoritma *K-Means clustering*” *Clustering* dilakukan dengan 6 kali iterasi didapatkan hasil 8 wilayah dengan tingkat pencurian tinggi, 10 wilayah dengan tingkat kajahatan sedang, dan 9 wilayah dengan tingkat kejahatan rendah. Dari beberapa contoh penelitian terdahulu hanya meneliti *clustering* tindak kejahatan di wilayah kota maupun satu provinsi tertentu saja, oleh karena itu peneliti melakukan penelitian *clustering* tindak kejahatan di Indonesia menurut provinsi untuk mencari informasi terkait provinsi mana yang tingkat kejahatannya tinggi, sedang, ataupun rendah.

Berdasarkan uraian diatas, maka peneliti tertarik untuk melakukan *clustering* tingkat kejahatan di Indonesia menggunakan metode *K-Means clustering*. Sehingga dalam penelitian ini dapat diketahui tingkat kerawanan kejahatan di Indonesia. Variabel pada penelitian ini yaitu klasifikasi jenis kejahatan dan objek penelitian yang digunakan adalah 34 provinsi di Indonesia.

1.2. Rumusan Masalah

Berdasarkan yang sudah dipaparkan pada latar belakang diatas, maka rumusan masalah yang akan dibahas yaitu:

1. Bagaimana *clustering* tingkat kejahatan di Indonesia menggunakan metode

K-Means?

2. Bagaimana keakuratan hasil *clustering* tingkat kejahatan di Indonesia menggunakan metode *K-Means*?

1.3. Tujuan Penelitian

Tujuan dari penelitian ini yaitu :

1. *Mengclustering* tingkat kejahatan di Indonesia menurut provinsi menggunakan metode *K-Means*.
2. Mengetahui keakuratan hasil *clustering* tingkat kejahatan di Indonesia menggunakan metode *K-Means*.

1.4. Manfaat Penelitian

Manfaat yang diperoleh dengan adanya penelitian ini antara lain:

1. Bagi Penulis: Sebagai penerapan ilmu statistika khususnya metode *K-Means clustering* dalam kehidupan sehari-hari.
2. Bagi Pemerintah: Sebagai gambaran pengelompokan tingkat kejahatan di Indonesia menurut provinsi untuk meningkatkan tingkat keamanan pada daerah yang memiliki tingkat kejahatan tinggi.
3. Bagi Masyarakat: Himbauan untuk selalu berhati-hati dan peringatan untuk tidak melakukan tindak kejahatan.

1.5. Batasan Masalah

Batasan masalah dalam penelitian ini yaitu:

1. Data yang digunakan adalah data jumlah kasus kejahatan di 34 provinsi di Indonesia tahun 2020.
2. Percobaan inialisasi *cluster k* sebanyak 2 dan 3 *cluster*.

1.6. Sistematika Penulisan

Sistematika Penulisan dalam penelitian ini yaitu:

1. BAB I PENDAHULUAN

Bab ini berisikan latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

2. BAB II TINJAUAN PUSTAKA

Bab ini berisikan tentang landasan teori disertai teori pendukung yang terkait dengan topik penelitian yaitu *clustering* daerah rawan kejahatan di Indonesia menggunakan metode *K-Means clustering*.

3. BAB III METODE PENELITIAN

Bab ini menjelaskan tentang deskripsi data penelitian yaitu, sumber data, variabel yang digunakan, kerangka penelitian dan alur penelitian.

4. BAB IV HASIL DAN PEMBAHASAN

Berisi hasil dari penerapan metode *K-Means* untuk *clustering* tingkat kejahatan di Indonesia, serta menganalisis hasil dari penerapan tersebut.

5. PENUTUP

Berisi kesimpulan dari hasil penelitian yang telah dilakukan dan saran untuk penelitian selanjutnya.

BAB II

TINJAUAN PUSTAKA

2.1. Kejahatan di Indonesia

Kejahatan adalah sebuah tindakan merugikan yang terus terjadi di masyarakat. Tindak kejahatan dapat diartikan sebagai perbuatan yang bertentangan dengan hukum. Pelaku kejahatan disebut sebagai kriminal. Ciri-ciri kriminal adalah pencuri, pembunuh, teroris, dan perampok. (Ali Hasymi et al., 2021). Definisi kejahatan secara yuridis adalah perilaku yang bertentangan dengan aturan atau hukum negara tertentu. Kemudian kejahatan dapat dilihat secara sosiologis sebagai suatu perbuatan yang merugikan masyarakat dan pasti akan menimbulkan tanggapan negatif dari masyarakat (Maharani, 2016).

Kitab Undang-Undang Hukum Pidana (KUHP) berisi mengenai definisi untuk setiap tindak kriminal kejahatan. Tindak pidana mengenai narkoba diatur dalam pasal 127 KUHP, lalu untuk tindak pidana pemerkosaan diatur dalam pasal 281-303 KUHP. Sedangkan tindak pidana mengenai penculikan diatur dalam pasal 324-334 KUHP. Untuk mengenai tindak kriminal pencurian dengan kekerasan diatur dalam pasal 365 KUHP. Menurut Muhson and UNODC (2013) Klasifikasi jenis kejahatan memiliki 8 unsur yaitu:

1. Kejahatan Terhadap Nyawa

Kejahatan jenis ini adalah kejahatan yang mempunyai tingkat paling tinggi hierarkinya dalam lingkup nasional maupun internasional. Contoh kejahatan jenis ini adalah pembunuhan. Didalam KUHP Indonesia sudah diatur dalam

pasal 338 yang berbunyi, "Barang siapa dengan sengaja menghilangkan nyawa orang lain, diancam karena pembunuhan akan dipidana penjara paling lama lima belas tahun."

2. Kejahatan terhadap Fisik/Badan

Beberapa contoh jenis kejahatan ini adalah penganiayaan ringan, penganiayaan berat, dan kekerasan dalam rumah tangga (KDRT). Tindakan ini diatur dalam pasal 351 KUHP yang berbunyi:

- (a) Penganiayaan diancam dengan pidana denda paling banyak empat ribu lima ratus rupiah atau pidana penjara paling lama dua tahun delapan bulan.
- (b) Jika perbuatan mengakibatkan luka-luka berat, yang bersalah diancam dengan pidana penjara paling lama lima tahun.

3. Kejahatan Terhadap Kesusilaan

Percabulan dan pemerkosaan merupakan jenis kejahatan terhadap kesusilaan. Di dalam pasal 281 hingga pasal 302 KUHP. Menurut pasal 281 KUHP menyatakan barang siapa yang melanggar kesusilaan akan mendapat ancaman pidana penjara paling lama dua tahun delapan bulan atau denda paling banyak lima ratus rupiah.

4. Kejahatan Terhadap Kemerdekaan Orang Diantara jenis kejahatan ini yaitu penculikan dan memperkejakan anak dibawah umur. Dalam perspektif hukum, penculikan diatur dalam pasal 328 dan pasal 333 KUHP. Jika korban anak, aturan yang digunakan adalah Undang-Undang Perlindungan Anak (UU No. 23 Tahun 2022).

5. Kejahatan Terhadap Hak Milik Dengan Kekerasan

Contoh kejahatan ini yaitu pencurian dengan kekerasan menggunakan senjata api dan pencurian dengan kekerasan menggunakan senjata tajam. Hal ini juga sudah diatur dalam pasal 335 KUHP yang berbunyi, "Penganiayaan berat dilakukan dengan rencana terlebih dahulu diancam dengan pidana penjara paling lama dua belas tahun."

6. Kejahatan Terhadap Hak Milik Tanpa Kekerasan

Di dalam KUHP pasal 362 menyatakan bahwa setiap perbuatan mengambil barang milik orang lain dapat dianggap sebagai "Melawan Hukum" jika perbuatan tersebut dilakukan dengan niat. Contoh kejahatan ini yaitu pencurian dan pembakaran dengan sengaja.

7. Kejahatan Terhadap Narkotika

Kejahatan narkotika memiliki banyak bentuk seperti bandar, pengedar, maupun pengonsumsi narkotika. Dalam UU Narkotika sudah diatur dalam pasal 112 dan 114. Misalnya pada pasal 112 ayat 1 menyatakan bahwa setiap orang yang tanpa hak atau melawan hukum memiliki, menyimpan, menguasai, atau menyediakan Narkotika golongan I bukan tanaman, dipidana dengan penjara paling singkat empat tahun dan paling lama dua belas tahun dan pidana denda paling sedikit Rp 800 juta dan paling banyak Rp 8 miliar.

8. Kejahatan Terhadap Penipuan, Penggelapan, Korupsi

Kejahatan jenis ini diantaranya penipuan, perbuatan curang, penggelapan dana, dan korupsi. Hal ini sudah diatur dalam KUHP pasal 603 yang berbunyi, "Setiap orang yang melawan hukum melakukan perbuatan

memperkaya diri sendiri, orang lain, atau korporasi yang merugikan keuangan negara atau perekonomian negara, dipidana dengan pidana penjara seumur hidup atau pidana penjara paling singkat dua tahun dan paling lama dua puluh tahun.”

Di Indonesia garda terdepan untuk mengatasi tindak kejahatan adalah polisi. Meskipun tidak mudah bagi polisi untuk menyelidiki dan mengevaluasi kejahatan. Pengembangan teknologi dan analisis kejahatan untuk memerangi kejahatan yang signifikan secara efektif dan efisien dengan memprediksi wilayah mana yang aman dan wilayah mana yang rawan kejahatan (Ali Hasymi et al., 2021). Dalam upaya membentuk citra yang baik, polisi melakukan pemetaan kejahatan untuk mengelola dan menginterpretasikan informasi lokal tentang kejahatan. Pemetaan daerah rawan kejahatan berguna untuk menemukan kejahatan atau mempelajari insiden yang dilaporkan. (Octavianti et al., 2017).

2.2. Clustering

Clustering adalah proses membagi data dari kelompok yang diberikan menjadi beberapa kelompok yang mempunyai karakteristik data yang serupa dengan karakteristik data dari kelompok lainnya. Kategori kelompok yang termasuk dalam proses *clustering* disebut *cluster*. *Cluster* adalah satu kelompok data yang memiliki karakteristik yang sama dan berbeda dengan *cluster* lainnya. (Febrianti et al., 2016).

Secara umum *clustering* berfungsi sebagai penentu sifat pola data dan memiliki karakteristik *unsupervised* (tanpa arah). *Clustering* dapat dikembangkan melalui partisi. Data dikumpulkan ke dalam satu *cluster* menggunakan teknik partisi sehingga setiap *cluster* memiliki setidaknya satu item dari kelompok data

yang sama. (Cahyati et al., 2017).

Untuk menemukan data di berbagai kelompok data yang memiliki sifat yang serupa, teknik *clustering* juga sering digunakan. memberikan dasar untuk mengklasifikasikan data yang tidak terklasifikasi. Landasan konsep *clustering* adalah pengelompokan beberapa data ke dalam *cluster* yang sesuai, yaitu *cluster* yang berbeda secara signifikan dari item *cluster* lainnya. (Nurjoko et al., 2020). Oleh karena itu, *clustering* adalah teknik yang digunakan untuk menandai data yang kategori kelompoknya belum diketahui secara pasti.

Untuk melakukan *clustering* harus memenuhi beberapa syarat berikut (Rizal Muzaky, 2020):

1. Skalabilitas

Clustering harus dapat diskalakan untuk menangani kumpulan data yang besar. Volume data yang besar saat ini sering dipakai di banyak sektor, seperti database. Basis data yang cukup besar berisi jutaan objek selain ratusan ribu objek.

2. Mampu menganalisa data dari berbagai macam bentuk

Algoritma pengelompokan ini mampu menganalisis data dalam berbagai format. Data berupa data nominal, ordinal atau gabungan dari data keduanya dapat dianalisis lebih lanjut.

3. Mampu mengolah data yang hilang atau rusak

Data tidak selalu dalam bentuk yang sempurna. Data terkadang bisa hilang atau rusak. Oleh karena itu metode *clustering* harus mampu menggunakan algoritma pengelompokan untuk menangani data yang hilang atau rusak.

4. Mampu mengolah data yang mempunyai dimensi tinggi

Kumpulan data bisa memiliki banyak dimensi, oleh karena itu metode *clustering* harus memiliki algoritma yang mampu menghitung data dengan dimensi yang tinggi.

2.3. Uji Kaiser Meyer Olkin Measur of Sampling (KMO)

Sebelum melakukan proses *clustering* menggunakan metode *K-Means*, perlu dilakukan *pre-processing* data yaitu penyederhanaan data melalui reduksi variabel yang tidak signifikan. Untuk mereduksi data tersebut digunakan uji KMO dan PCA. Uji KMO berfungsi untuk melakukan pengecekan awal data dapat dilakukan proses *factoring* atau tidak. *Factoring* adalah proses reduksi data dari variabel yang tidak signifikan. Skala ujinya yaitu berkisar antara 0 sampai 1. Jika nilai dari KMO dibawah 0.5, maka data tidak dapat dilakukan proses *factoring*. Persamaan untuk menghitung nilai uji KMO yaitu (Hartono et al., 2022):

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p \alpha_{ij}^2} \quad (2.1)$$

Keterangan:

i : 1,2,3, ... ,p, dimana $i \neq j$.

j : 1,2,3, ... ,p, dimana $i \neq j$.

r_{ij} : Nilai koefisien korelasi variabel ke- i dengan variabel ke- j .

α_{ij} : Nilai koefisien korelasi parsial variabel ke- i dengan variabel ke- j .

Untuk mencari nilai koefisien korelasi (r) persamaannya adalah sebagai berikut:

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}} \quad (2.2)$$

Keterangan:

r : Nilai koefisien korelasi.

n : Jumlah data.

x : Nilai Variabel bebas.

y : Nilai variabel respon.

Lalu untuk mencari nilai koefisien korelasi parsial α dapat menggunakan persamaan berikut:

$$\alpha = \frac{r_{yx_i} - r_y r_{x_i}}{\sqrt{1 - r_y^2} \sqrt{1 - r_{x_i}^2}} \quad (2.3)$$

Keterangan:

α : Nilai koefisien korelasi parsial.

r : Nilai koefisien korelasi.

x_i : Nilai variabel bebas. $i = 1, 2, \dots, n$

y : Nilai variabel respon.

Uji KMO memiliki kriteria yaitu:

Tabel 2.1 Kriteria nilai uji KMO

| Nilai KMO | Kriteria |
|----------------|-------------|
| $KMO \geq 0,9$ | Sangat Baik |
| $KMO \geq 0,8$ | Baik |
| $KMO \geq 0,7$ | Menengah |
| $KMO \geq 0,6$ | Cukup |
| $KMO \geq 0,5$ | Kurang Baik |
| $KMO < 0,5$ | Buruk |

2.4. *Principal Component Analysis (PCA)*

Principal Component Analysis (PCA) merupakan metode yang digunakan untuk melakukan proses *factoring* atau penyederhanaan data tanpa mengurangi karakteristik data (Putra et al., 2016). PCA berfungsi untuk mengetahui nilai loading setiap variabel untuk melihat signifikansi variabel yang digunakan terhadap kriminalitas. Jika nilai loading kurang dari 0,5 maka variabel tersebut akan direduksi. Langkah-langkah *factoring* menggunakan PCA adalah sebagai berikut (Anwar et al., 2022):

1. Menghitung matriks kovarian dengan persamaan sebagai berikut:

$$Cov(xy) = \frac{\sum xy}{n} - (\bar{x})(\bar{y}) \quad (2.4)$$

Keterangan:

$Cov(xy)$: Nilai matriks kovarian.

x : Nilai data ke- x .

y : Nilai data ke- y .

\bar{x} : Rata-rata data ke- x .

\bar{y} : Rata-rata data ke- y .

2. Menghitung nilai eigen dengan persamaan sebagai berikut:

$$[Cov - \lambda I] = 0 \quad (2.5)$$

Keterangan:

Cov : Matriks kovarian.

λ : Nilai eigen.

I : Matriks *identitas*.

3. Mengitung vektor eigen dengan persamaan sebagai berikut:

$$[Cov - \lambda I][V] = 0 \quad (2.6)$$

Dimana V adalah vektor eigen.

4. Mengurutkan nilai vektor eigen dari yang terbesar sampai yang terkecil.

Untuk menentukan jumlah komponen utama dalam PCA dapat diketahui dengan nilai eigen ≥ 1 . Jika nilai eigennya < 1 , maka nilai tersebut tidak masuk kedalam komponen utama.

5. Rotasi faktor

Rotasi faktor digunakan untuk mengetahui nilai *loading* dari setiap variabel.

Nilai *loading* merupakan nilai yang berfungsi untuk mengetahui hubungan antar variabel. Variabel yang memiliki nilai *loading* $> 0,5$ dapat diartikan sebagai variabel yang signifikan mempengaruhi kriminalitas. Sedangkan variabel yang memiliki nilai *loading* kurang dari 0,5 tidak signifikan mempengaruhi kriminalitas, oleh karena itu variabel tersebut dapat di reduksi.

2.5. K-Means

Algoritma *K-Means* membagi data set menjadi sejumlah k *cluster* tertentu sebagai bagian dari proses *clustering*. Algoritma *K-Means* sangat mudah

digunakan, populer, dan efisien. *K-Means* secara historis menjadi teknik penambangan data yang paling signifikan (Nabila et al., 2021).

K-Means merupakan strategi *clustering* non-hierarchical yang memisahkan data penelitian menjadi beberapa *cluster* atau kelompok, sehingga beberapa data memiliki kesamaan ciri dan beberapa data memiliki karakteristik yang berbeda (Annur, 2019).

Analisis *cluster* menggunakan metode *K-Means* mempunyai algoritma sebagai berikut (Martiano et al., 2019):

1. Menginisialisasi *cluster* sebanyak k .
2. Menentukan *centroid* awal secara random dari objek data sebanyak k .
3. Menghitung jarak terdekat dari setiap data dengan *centroid* yang sudah ditentukan menggunakan jarak *euclidean* dengan rumus:

$$d(x_i, \mu_i) = \sqrt{\sum (x_i - \mu_i)^2} \quad (2.7)$$

Keterangan

$d(x_i, \mu_i)$: Jarak antara data ke- i dan pusat cluster.

x_i : Nilai data ke- i .

μ_i : Pusat pada cluster i .

4. Menentukan jarak terdekat data dengan *centroid*, yaitu dari nilai terkecil perhitungan jarak *euclidean*.
5. Menentukan (*centroid*) baru dengan menggunakan persamaan:

$$c_k = \frac{1}{n_k} \sum_{i=1}^n d_i \quad (2.8)$$

Keterangan

c_k : Nilai *centroid* baru.

n_k : Jumlah data di dalam *cluster*

d_i : Jumlah dari nilai jarak di setiap *cluster*.

6. Lakukan pengulangan dari langkah 2 sampai 5 sampai anggota tiap *cluster* tidak ada yang berubah.
7. Jika langkah 6 terpenuhi, maka nilai dari *centroid* iterasi terakhir menjadi *centroid* pada *cluster* yang akan diklasifikasi.

2.6. *Shilhouette Coefficient*

Shilhouette Coefficient adalah metode untuk mengukur kualitas klaster yang menekankan pada validasi dan interpretasi kumpulan data. Selain itu, derajat kepemilikan masing-masing komponen dalam *cluster* dapat ditampilkan dengan menggunakan *shilhouette coefficient*. Langkah pertama dalam menghitung nilai *shilhouette coefficient* yaitu mencari jarak rata-rata data ke- i dengan semua data di *cluster* yang sama. Misalkan data ke- i berada di *cluster* A. Rumus dari $\alpha(x_i)$ yaitu: (Struyf et al., 1996)

$$\alpha(x_i) = \frac{1}{|x_\alpha| - 1} \sum_{j \in A, j \neq i} d(x_i, x_j) \quad (2.9)$$

Dimana $\alpha(x_i)$ adalah jarak rata-rata data ke- i dengan semua data di *cluster* yang sama, x_i merupakan data ke- i , x_α merupakan jumlah anggota *cluster* A. $|x_\alpha| - 1$ karena untuk mencari nilai perbandingan jarak data didalam 1 *cluster* yang sama, misalkan ada 3 data di dalam 1 *cluster*, maka perbandingan jarak data ke 1 dengan data ke 2, perbandingan jarak data ke 1 dengan jarak ke 3, diperoleh 2 nilai jarak data atau dengan kata lain jumlah anggota di dalam *cluster* harus dikurangi 1.

Sedangkan $d(x_i, x_j)$ adalah fungsi jarak data ke- i pada *cluster* j .

Lalu, langkah selanjutnya adalah mencari nilai $b(x_i)$ yang merupakan nilai minimum dari jarak rata-rata data ke- i dengan semua data di *cluster* yang berbeda.

Rumus dari $b(x_i)$ yaitu:

$$b(x_i) = \min_{x \neq x_\alpha} \frac{1}{|x_b|} \sum_{j \in x_b} d(x_i, x_j) \quad (2.10)$$

Dimana, $b(x_i)$ adalah nilai minimum dari jarak rata-rata data ke- i dengan semua data di *cluster* yang berbeda. x_i merupakan data ke- i , x_b adalah jumlah anggota *cluster* b , dan $d(x_i, x_j)$ adalah fungsi jarak data ke- i pada *cluster* j .

Langkah terakhir, yaitu menghitung nilai *silhouette coefficient* (SC) menggunakan rumus (Struyf et al., 1996):

$$S(C) = \frac{b(x_i) - \alpha(x_i)}{\max(\alpha(x_i) - b(x_i))} \quad (2.11)$$

Keterangan:

$S(C) \approx 1$, maka data ke- i dikelompokkan dengan baik dalam *cluster* A.

$S(C) \approx 0$, maka data ke- i berada ditengah *cluster* A dan B.

$S(C) \approx -1$, maka data ke- i tergolong lemah atau dengan kata lain dekat dengan *cluster* B.

Perhitungan koefisien *silhouette* menghasilkan nilai antara -1 dan 1. Nilai *clustering* baik, jika nilai koefisien *silhouette* bernilai positif, yaitu jika $\alpha_i < b_i$ dan α_i mendekati 0, sehingga diperoleh nilai koefisien *silhouette* terbesar adalah 1. Jika $S(C) \approx 1$, data i berada di *cluster* yang sesuai. Selain itu, jika $S(C) \approx 0$, maka data berada di antara kedua klaster, menunjukkan struktur objek yang belum jelas. Jika $S(C) \approx -1$, maka objek mengalami *overlapping* nilai yang mengakibatkan

dimasukkan ke dalam *cluster* lain (Rousseeuw, 1987).

Setiap objek data dalam *cluster* memiliki faktor *silhouette* rata-rata, yang menentukan seberapa baik data harus masuk ke dalam kluster. Berikut tabel nilai *silhouette coefficient*:

Tabel 2.2 Silhouette Coefficient

| Nilai <i>Silhouette Coefficient</i> | Struktur |
|-------------------------------------|-------------------------|
| $0.7 < SC \leq 1$ | Kuat |
| $0.5 < SC \leq 0.7$ | Standar |
| $0.25 < SC \leq 0.5$ | Lemah |
| $SC \leq 0.25$ | Tidak memiliki Struktur |

2.7. Integrasi Keilmuan

1. Kejahatan dalam Perspektif Islam

Sebagai manusia harus saling berbuat kebaikan kepada siapapun dan dimanapun kita berada. Orang yang berbuat baik akan mendapatkan kebaikan juga, sedangkan orang yang berbuat jahat akan mendapatkan kejahatan pula. Begitu juga dalam agama Islam sangat melarang untuk melakukan perbuatan jahat misalnya pada surat Al-Isra' ayat 33 yang berbunyi:

وَلَا تَقْتُلُوا النَّفْسَ الَّتِي حَرَّمَ اللَّهُ إِلَّا بِالْحَقِّ وَمَنْ قُتِلَ مَظْلُومًا فَقَدْ جَعَلْنَا لَوْلِيهِ سُلْطٰنًا فَلَا يُسْرِفُ فِي الْقَتْلِ إِنَّهُ كَانَ مَنْصُورًا

Artinya: "Dan janganlah kamu membunuh orang yang diharamkan Allah

(membunuhnya), kecuali dengan suatu (alasan) yang benar. Dan barang siapa dibunuh secara zalim, maka sungguh, Kami telah memberi kekuasaan kepada walinya, tetapi janganlah walinya itu melampaui batas dalam pembunuhan. Sesungguhnya dia adalah orang yang mendapat pertolongan.”

Allah dengan jelas memperingatkan kita untuk tidak melakukan kejahatan khususnya pembunuhan, kecuali dengan alasan yang benar. Misalnya dalam peperangan melawan orang-orang kafir dengan tujuan memperjuangkan agama Islam, hal ini diperbolehkan untuk melakukan pembunuhan, karena tujuannya benar dan jelas. Begitu juga dalam kasus kejahatan lainnya seperti pencurian, Allah telah memperingatkan kita dalam surat Al-Maidah ayat 38 yang berbunyi:

وَالسَّارِقُ وَالسَّارِقَةُ فَاقْطَعُوا أَيْدِيَهُمَا جَزَاءً بِمَا كَسَبَا نَكَالًا مِّن
اللَّهِ ۗ وَاللَّهُ عَزِيزٌ حَكِيمٌ

Artinya: ”Adapun orang laki-laki maupun perempuan yang mencuri, potonglah tangan keduanya (sebagai) balasan atas perbuatan yang mereka lakukan dan sebagai siksaan dari Allah. Dan Allah Mahaperkasa, Mahabijaksana.”

Dari ayat diatas sudah jelas bahwa hukum dari melakukan tindakan kejahatan pencurian sangatlah berat. Hukum di dunia dengan dipotongnya kedua tangan si pelaku pencurian dan nanti diakhirat akan mendapat siksaan yang berat juga dari Allah. Rasulullah juga memperingatkan kita melalui hadis yang diriwayatkan oleh Imam Bukhari yang berbunyi:

حَدَّثَنَا عُمَرُ بْنُ حَفْصِ بْنِ غِيَاثٍ حَدَّثَنِي أَبِي حَدَّثَنَا الْأَعْمَشُ قَالَ سَمِعْتُ أَبَا
صَالِحٍ عَنْ أَبِي هُرَيْرَةَ عَنِ النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ لَعَنَ اللَّهُ السَّارِقَ
يَسْرِقُ الْبَيْضَةَ فَتُقَطَّعُ يَدُهُ وَيَسْرِقُ الْحَبْلَ فَتُقَطَّعُ يَدُهُ قَالَ الْأَعْمَشُ كَانُوا يَرَوْنَ
أَنَّهُ بَيِّضُ الْحَدِيدِ وَالْحَبْلُ كَانُوا يَرَوْنَ أَنَّهُ مِنْهَا مَا يَسْوَى دَرَاهِمٍ

Artinya: "Telah menceritakan kepada kami Umar bin Hafsh bin Ghiyats telah menceritakan kepadaku ayahku Telah menceritakan kepada kami Al A'masy mengatakan; aku mendengar Abu Shalih dari Abu Hurairah dari Nabi shallallahu 'alaihi wasallam bersabda: "Allah melaknat si pencuri telur sehingga tangannya dipotong, dan Allah melaknat si pencuri tali hingga dipotong tangannya." Al A'masy mengatakan, para sahabat berpendapat bahwa yang dimaksud telur disini adalah besi dan yang dimaksud tali adalah beberapa dirham." (HR. Bukhari No. 6285)

Selain itu contoh tindak kejahatan lainnya yang sering terjadi bahkan dikalangan para pejabat yaitu adalah korupsi atau suap menyuap. dalam surat Al-Baqarah ayat 188 Allah berfirman :

UIN SUNAN AMPEL

وَلَا تَأْكُلُوا أَمْوَالَكُمْ بَيْنَكُمْ بِالْبَاطِلِ وَتُدْلُوا بِهَا إِلَى الْحُكَّامِ لِتَأْكُلُوا
فَرِيقًا مِّنْ أَمْوَالِ النَّاسِ بِالْإِثْمِ وَأَنْتُمْ تَعْلَمُونَ

Artinya: "Dan janganlah kamu makan harta di antara kamu dengan jalan yang batil, dan (janganlah) kamu menyuap dengan harta itu kepada para hakim, dengan maksud agar kamu dapat memakan sebagian harta orang lain itu dengan jalan dosa, padahal kamu mengetahui."

2. Faktor-Faktor yang Mempengaruhi Kejahatan

Faktor yang mempengaruhi tindakan kejahatan salah satunya yaitu pengangguran atau orang yang tidak memiliki kesibukan. Efek dari pengangguran membuat ekonomi akan semakin sulit yang mengakibatkan kemiskinan (Nugroho and Harmadi, 2015). Oleh karena itu hendaknya orang yang pengangguran segera mencari pekerjaan untuk menghidupi dan memenuhi kebutuhannya sendiri, supaya tidak merasa kekurangan dan tidak ada dorongan untuk melakukan kejahatan. Jika ada orang yang ekonominya sulit hendaknya sesama muslim berbagi dan membantu semampunya. Allah berfirman melalui surat Az-Zariyat ayat 19 yang berbunyi:

وَفِي أَمْوَالِهِمْ حَقٌّ لِّلسَّائِلِ وَالْمَحْرُومِ

Artinya: "Dan pada harta benda mereka ada hak untuk orang miskin yang meminta dan orang miskin yang tidak meminta."

Selain itu faktor yang mempengaruhi seseorang untuk melakukan tindakan kejahatan adalah minuman keras. Karena efek dari minuman keras dapat membuat otak tidak bekerja dengan baik dan sering kali lepas kendali. Allah berfirman dalam surat Al-Maidah ayat 90 yang berbunyi:

يَا أَيُّهَا الَّذِينَ آمَنُوا إِنَّمَا الْخَمْرُ وَالْمَيْسِرُ وَالْأَنْصَابُ وَالْأَزْلَامُ
رِجْسٌ مِّنْ عَمَلِ الشَّيْطَانِ فَاجْتَنِبُوهُ لَعَلَّكُمْ تُفْلِحُونَ

Artinya: "Wahai orang-orang yang beriman! Sesungguhnya minuman keras, berjudi, (berkurban untuk) berhala, dan mengundi nasib dengan anak panah,

adalah perbuatan keji dan termasuk perbuatan setan. Maka jauhilah (perbuatan-perbuatan) itu agar kamu beruntung.”

3. Menjauhi Perbuatan Kejahatan

Sebagai manusia senantiasa harus menjauhi perbuatan kejahatan. Janganlah menghalalkan semua cara untuk mendapatkan rezeki, jika makan dari harta yang diperoleh dengan cara yang haram hal itu merupakan jalan yang buruk dan dosa besar. Efek dari orang yang sering makan makanan yang haram akan dicampakkan ke neraka dengan penuh hina dan siksaan. Hal ini juga sama seperti sabda Rasulullah SAW yang berbunyi ”Tidaklah tumbuh daging dari makanan haram, kecuali neraka lebih utama untuknya.” (HR. Tirmidzi).

Manusia memanglah tempatnya salah dan lupa. Oleh karena itu sering kali manusia lalai dan khilaf. Kita harus bersabar dengan memohon kepada Allah dengan mengharap keridhoannya. Dalam surat Al-Kahfi ayat 28 Allah berfirman:

وَاصْبِرْ نَفْسَكَ مَعَ الَّذِينَ يَدْعُونَ رَبَّهُمْ بِالْغَدْوَةِ وَالْعَشِيِّ يُرِيدُونَ
وَجْهَهُ وَلَا تَعْدُ عَيْنِكَ عَنْهُمْ تُرِيدُ زِينَةَ الْحَيَاةِ الدُّنْيَا وَلَا تُطِعْ مَنْ
أَغْفَلْنَا قَلْبَهُ عَن ذِكْرِنَا وَاتَّبَعَ هَوَاهُ وَكَانَ أَمْرُهُ فُرُطًا

Artinya: ”Dan bersabarlah engkau (Muhammad) bersama orang yang menyeru Tuhannya pada pagi dan senja hari dengan mengharap keridhaan-Nya; dan janganlah kedua matamu berpaling dari mereka (karena) mengharapkan perhiasan kehidupan dunia; dan janganlah engkau mengikuti

orang yang hatinya telah Kami lalaikan dari mengingat Kami, serta menurut keinginan dan keadaannya sudah melewati batas.”

Oleh karena itu, meskipun manusia tempatnya salah dan lupa, kita harus bersabar dan terus menyeru kepada Allah dengan mengharapkan ridhonya. Dunia hanyalah sementara, jangan terlalu mengharapkan keindahan dunia semata, tapi ingat tujuan kita nanti adalah kehidupan kekal abadi di akhirat, karena sebaik baik manusia adalah pendosa yang bertaubat. Rasulullah mengatakan dalam hadis yang diriwayatkan oleh At-Tirmidzi yang berbunyi:

كُلُّ بَنِي آدَمَ خَطَاءٌ ، وَخَيْرُ الْخَطَائِينَ التَّوَّابُونَ

Artinya:”Setiap anak Adam pasti berbuat salah dan sebaik-baik orang yang berbuat kesalahan adalah yang bertaubat”.(HR Tirmidzi No. 2499)

Berdasarkan uraian diatas dapat disimpulkan bahwa kita harus menjadi orang yang baik dan menjauhi tindak kejahatan. Karena kalau berbuat baik, maka kebaikan itu akan kembali kepada kita. Dan sebaliknya jika berbuat kejahatan, maka kita akan mendapat hukuman pidana oleh negara dan juga mendapat hukuman dari Allah SWT. Peringatan Allah dan Rasulullah SAW sangat cukup menyadarkan kita untuk menjauhi tindakan kejahatan, Meskipun begitu jumlah kasus kejahatan di Indonesia masih tergolong tinggi. Jadi selalu waspada dan berhati-hati, karena bagaimanapun manusia tempatnya salah dan lupa.

Pada penelitian ini melakukan *clustering* untuk pemetaan tingkat daerah rawan kejahatan di Indonesia. Hasil pemetaan sangat berguna untuk membantu aparat keamanan negara dalam mereduksi tindak kejahatan. Selain itu, dengan diketahuinya daerah yang rawan kejahatan, aparat keamanan dapat meningkatkan keamanan sebagai upaya *preventif* untuk mencegah tindak kejahatan. Dengan

begitu perlahan-lahan tingkat kejahatan di masing-masing daerah akan mengalami penurunan.



UIN SUNAN AMPEL
S U R A B A Y A

BAB III

METODE PENELITIAN

3.1. Jenis Penelitian

Jenis penelitian yang dipakai pada penelitian ini yaitu kuantitatif deskriptif. Penelitian kuantitatif deskriptif menggunakan angka rumus, dan grafik untuk hasil perhitungannya. Selain itu hasil perhitungannya terstruktur, sistematis, dan akurat. Penelitian diawali dengan mengumpulkan literatur yang berkaitan dengan permasalahan dalam penelitian. Data yang didapatkan kemudian disusun secara sistematis lalu dilakukan perhitungan untuk mendapatkan hasil dan kesimpulan (Prasetyo and Abduh, 2021).

3.2. Sumber Data

Penelitian menggunakan jenis data sekunder yang didapat dari website resmi BPS (Badan Pusat Statistika) Indonesia yaitu data statistik kejahatan di Indonesia pada tahun 2020 yang diperoleh melalui <https://www.bps.go.id/publication/statistik-kriminal-2021>.

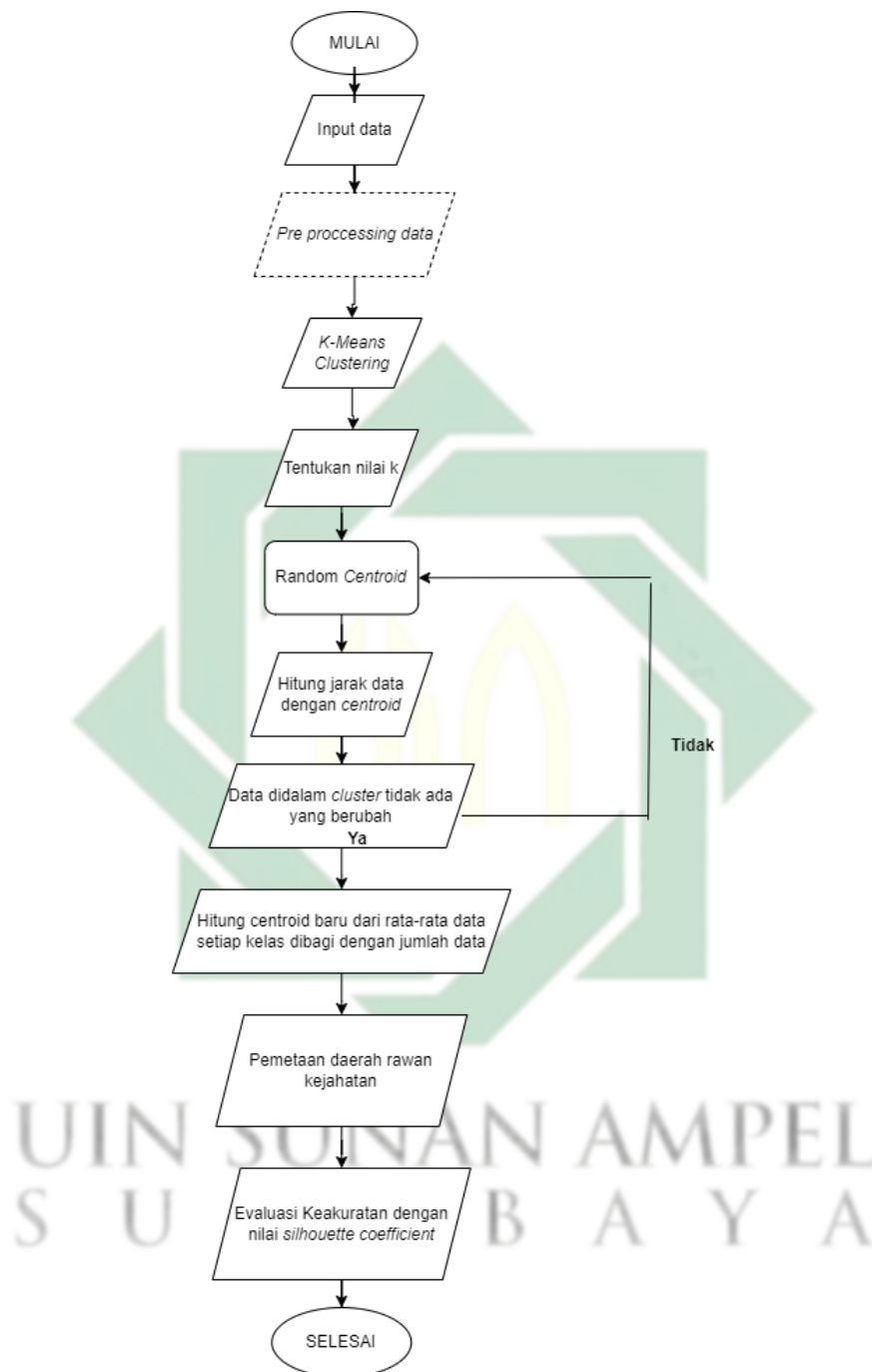
Variabel yang digunakan terdiri dari klasifikasi tingkat kejahatan di Indonesia yaitu kejahatan terhadap nyawa (X1), kejahatan terhadap fisik/badan (X2), kejahatan terhadap kesusilaan (X3), kejahatan terhadap kemerdekaan orang (X4), kejahatan terhadap hak milik dengan kekerasan (X5), kejahatan terhadap hak milik tanpa kekerasan (X6), kejahatan terhadap narkoba (X7), dan kejahatan terhadap penipuan, penggelapan dana, dan korupsi (X8).

3.3. Teknik Analisis Data

Teknik analisis data dalam penelitian ini yaitu:

1. Mengumpulkan data klasifikasi kejahatan di Indonesia tahun 2020 menurut provinsi.
2. *Pre porcessing* data yang terdiri dari uji *Kaiser Meyer Olkin Measur of Sampling* (KMO) dan uji *Principal Component Analysis* (PCA)
3. Inisialisasi jumlah *cluster* = k.
4. Mencari nilai *centroid* awal secara acak.
5. Hitung jarak data ke *centroid* dengan menggunakan rumus *euclidean distance* pada persamaan (2.7).
6. Tentukan pusat *centroid* baru menggunakan persamaan (2.8).
7. Lakukan evaluasi nilai anggota setiap *cluster* dengan *shilhouette coefficient*.
8. Pemetaan daerah kerawanan kejahatan.
9. Evaluasi keakuratan hasil *clustering* dengan *silhouette coefficient*

Jika tahapan penelitian tersebut digambarkan melalui flowchart, maka akan seperti gambar berikut:



Gambar 3.1 Diagram Alur Penelitian

BAB IV

HASIL DAN PEMBAHASAN

4.1. Analisis Deskriptif Data

Deskripsi data merupakan penjelasan data yang akan digunakan dalam penelitian. Data yang digunakan berbentuk *cross section*. Data *cross section* adalah sekumpulan data dalam satu waktu tertentu. Selanjutnya data tersebut dijabarkan dengan deskripsi statistik seperti Tabel 4.1 berikut.

Tabel 4.1 Data Kejahatan di Indonesia

| No | Provinsi | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 |
|----|----------------------|----|------|-----|-----|-----|-------|------|------|
| 1 | Sumatera Utara | 99 | 6207 | 774 | 55 | 780 | 10916 | 5932 | 5562 |
| 2 | Sumatera Selatan | 77 | 1551 | 172 | 141 | 563 | 2911 | 2554 | 1766 |
| 3 | Sumatera Barat | 20 | 1724 | 294 | 97 | 185 | 3732 | 913 | 989 |
| 4 | Bengkulu | 33 | 479 | 108 | 1 | 106 | 963 | 340 | 556 |
| 5 | Kep. Bangka Belitung | 10 | 150 | 52 | 47 | 42 | 690 | 332 | 130 |
| 6 | Jambi | 11 | 531 | 69 | 9 | 87 | 17770 | 619 | 767 |
| 7 | Lampung | 21 | 647 | 247 | 41 | 399 | 2532 | 1521 | 1373 |
| 8 | Kalimantan Barat | 35 | 184 | 137 | 112 | 54 | 1254 | 755 | 431 |
| 9 | Kalimantan Timur | 12 | 352 | 113 | 1 | 151 | 862 | 1354 | 356 |
| 10 | Kalimantan Tengah | 17 | 305 | 115 | 1 | 86 | 837 | 590 | 315 |
| 11 | Kalimantan Selatan | 38 | 462 | 51 | 0 | 367 | 1208 | 1582 | 539 |
| 12 | Kalimantan Utara | 10 | 91 | 32 | 10 | 19 | 340 | 245 | 72 |
| 13 | Banten | 8 | 521 | 81 | 151 | 94 | 1657 | 634 | 718 |
| 14 | Jawa Barat | 20 | 1445 | 371 | 11 | 257 | 2950 | 1141 | 2792 |
| 15 | Jawa Tengah | 30 | 382 | 349 | 11 | 236 | 3883 | 1601 | 1814 |
| 16 | Jawa Timur | 55 | 1404 | 256 | 177 | 341 | 4976 | 2629 | 2817 |
| 17 | DI Yogyakarta | 6 | 573 | 98 | 27 | 126 | 1769 | 460 | 1225 |
| 18 | Bali | 6 | 414 | 65 | 4 | 54 | 821 | 668 | 408 |
| 19 | Nusa Tenggara Timur | 40 | 1726 | 160 | 14 | 42 | 1494 | 41 | 574 |
| 20 | Nusa Tenggara Barat | 11 | 646 | 95 | 46 | 284 | 2120 | 571 | 876 |

| | | | | | | | | | |
|----|-------------------|----|------|------|-----|-----|------|------|------|
| 21 | Gorontalo | 13 | 1043 | 129 | 73 | 4 | 696 | 101 | 467 |
| 22 | Sulawesi Barat | 20 | 505 | 64 | 97 | 15 | 607 | 175 | 989 |
| 23 | Sulawesi Tengah | 18 | 1391 | 175 | 22 | 140 | 2121 | 340 | 677 |
| 24 | Sulawesi Utara | 74 | 1976 | 242 | 88 | 72 | 1394 | 100 | 945 |
| 25 | Sulawesi Tenggara | 26 | 734 | 173 | 9 | 42 | 632 | 192 | 122 |
| 26 | Sulawesi Selatan | 71 | 3848 | 208 | 311 | 348 | 2911 | 1013 | 1786 |
| 27 | Maluku | 10 | 993 | 1398 | 119 | 23 | 894 | 142 | 307 |
| 28 | Maluku Utara | 1 | 255 | 100 | 13 | 14 | 112 | 120 | 49 |
| 29 | Papua Barat | 10 | 655 | 47 | 124 | 155 | 1301 | 120 | 356 |
| 30 | Papua | 27 | 1296 | 100 | 176 | 532 | 3678 | 241 | 701 |
| 31 | DKI Jakarta | 32 | 1449 | 113 | 31 | 259 | 3153 | 5981 | 4595 |
| 32 | Aceh | 15 | 1206 | 232 | 29 | 145 | 2408 | 1596 | 1130 |
| 33 | Riau | 21 | 1013 | 143 | 11 | 413 | 3152 | 1646 | 1095 |
| 34 | Kepulauan Riau | 13 | 514 | 109 | 1 | 103 | 1003 | 362 | 543 |

Selanjutnya hasil deskripsi data tersebut dideskripsikan pada Tabel 4.2 berikut.

Tabel 4.2 Deskripsi Data Kejahatan di Indonesia

| Variabel | Min | Max | Mean |
|----------|-----|-------|---------|
| X1 | 1 | 99 | 26,76 |
| X2 | 91 | 6207 | 1078,60 |
| X3 | 32 | 1398 | 202,12 |
| X4 | 0 | 311 | 60,59 |
| X5 | 4 | 780 | 192,30 |
| X6 | 112 | 10916 | 2110 |
| X7 | 41 | 5981 | 1076,80 |
| X8 | 49 | 5562 | 1113 |

Tabel 4.2 menjelaskan deskripsi data untuk masing-masing variabel yang digunakan dalam penelitian ini. Pada variabel X1 yaitu kejahatan terhadap nyawa memiliki nilai min 1, max 99, dan mean 26,76, dan seterusnya hingga variabel X8 memiliki nilai min 49., max 5562, dan mean 1113. Deskripsi data tersebut dapat diartikan pada tahun 2020 variabel X1 minimal ada 1 kejahatan terhadap nyawa, maksimal ada 99 kejahatan terhadap nyawa, dan rata-rata kejahatan terhadap

nyawa terjadi sebesar 26,76 di Indonesia. hingga pada variabel X8 yaitu kejahatan terhadap penipuan, penggelapan dana, dan korupsi minimal ada 49 kejahatan korupsi, maksimal 5562 kejahatan korupsi, dan rata-rata kejahatan korupsi sebesar 1113 di Indonesia.

4.2. Pre Processing Data

Pre-processing data dilakukan terlebih dahulu untuk mengetahui keterkaitan satu variabel dengan variabel lainnya. Untuk mengetahui jumlah *cluster* optimum. Jika ada data yang tidak memiliki keterkaitan dengan variabel lainnya atau tidak memenuhi dalam uji signifikansi, data tersebut dapat direduksi dalam perhitungan. *Pre-processing* data yang pertama yaitu uji KMO, untuk mengetahui hasil uji KMO dapat dilihat dilihat pada Tabel 4.3

Tabel 4.3 Hasil uji KMO

| | |
|--|---------|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | 0.780 |
| Bartlett's Test of Sphericity | 206.383 |
| df | 28 |
| Sig. | <0.001 |

Hasil nilai uji KMO yang didapatkan yaitu 0,780. Nilai tersebut lebih besar dari 0,5 dan masuk dalam kategori baik, Artinya data dapat dianalisis lebih lanjut.

Langkah selanjutnya yaitu menghitung nilai kovarian dari setiap variabel dan nilai tersebut dimasukkan kedalam matriks kovarian seperti pada Tabel 4.4

Tabel 4.4 Matriks Kovarian

| | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 |
|----|-----------|-------------|------------|-----------|------------|-------------|------------|-------------|
| X1 | 538.913 | 21003.143 | 1372.877 | 677.021 | 2729.738 | 29471.747 | 17732.344 | 17191.818 |
| X2 | 21003.143 | 1353960.310 | 123463.504 | 31515.189 | 136398.852 | 1844148.390 | 942820.398 | 1012094.152 |
| X3 | 1372.877 | 123463.504 | 63433.865 | 2823.262 | 8418.964 | 185233.854 | 77374.207 | 91708.061 |

| | | | | | | | | |
|----|-----------|-------------|------------|-----------|------------|-------------|-------------|-------------|
| X4 | 677.021 | 31515.189 | 2823.262 | 4967.401 | 3620.094 | 30602.178 | 1395.337 | 11983.061 |
| X5 | 2729.738 | 136398.852 | 8418.964 | 3620.094 | 34040.396 | 294591.786 | 173892.426 | 149945.636 |
| X6 | 29471.747 | 1844140.390 | 185233.854 | 30602.178 | 294591.796 | 3830170.775 | 2070790.619 | 2003731.121 |
| X7 | 17732.344 | 942820.398 | 77374.207 | 1395.337 | 173892.426 | 2070790.619 | 2000602.471 | 1562198.788 |
| X8 | 17191.818 | 1012094.152 | 91708.061 | 11983.061 | 149945.636 | 2003731.121 | 1562198.788 | 1487093.939 |

Setelah didapatkan nilai matriks kovarian, langkah selanjutnya adalah mencari nilai eigen dan vektor eigen. Vektor eigen diurutkan dari nilai yang terbesar sampai terkecil seperti pada Tabel 4.5

Tabel 4.5 Nilai Eigen

| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | |
|-----------|---------------------|---------------|--------------|-------------------------------------|---------------|--------------|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 4.775 | 59.682 | 59.682 | 4.775 | 59.682 | 59.682 |
| 2 | 1.159 | 14.483 | 74.165 | 1.159 | 14.483 | 74.165 |
| 3 | 0.897 | 11.218 | 85.383 | | | |
| 4 | 0.413 | 5.157 | 90.540 | | | |
| 5 | 0.370 | 4.623 | 95.164 | | | |
| 6 | 0.242 | 3.020 | 98.184 | | | |
| 7 | 0.89 | 1.116 | 99.300 | | | |
| 8 | 0.56 | 0.700 | 100.000 | | | |

Extraction Method: Principal Component Analysis

Terdapat 2 komponen yang memiliki nilai eigen ≥ 1 . Komponen 1 mempunyai nilai 4,775 dan komponen 2 bernilai 1,159. Untuk mengetahui variabel yang digunakan dalam proses perhitungan, ada atau tidaknya variabel yang direduksi maka dilakukan *rotasi faktor*. Sehingga diperoleh hasil seperti pada Tabel 4.6

Tabel 4.6 Rotasi Faktor

| Component | 1 | 2 |
|-----------|-------|-------|
| X1 | 0,612 | 0,584 |
| X2 | 0,696 | 0,580 |
| X3 | 0,253 | 0,419 |
| X4 | -0,65 | 0,902 |
| X5 | 0,775 | 0,319 |
| X6 | 0,883 | 0,322 |
| X7 | 0,932 | -0,22 |
| X8 | 0,929 | 0,146 |

Dari hasil diatas diketahui variabel yang berpengaruh secara signifikan terhadap kriminalitas ada 7 variabel yaitu kejahatan terhadap nyawa (X1) dengan nilai *loading* sebesar 0,612. Kejahatan terhadap fisik/badan (X2) dengan nilai *loading* sebesar 0,696. Kejahatan terhadap kemerdekaan orang (X4) dengan nilai *loading* sebesar 0,902. Kejahatan terhadap hak milik dengan kekerasan (X5) dengan nilai *loading* sebesar 0,775. Kejahatan terhadap hak milik tanpa kekerasan (X6) dengan nilai *loading* sebesar 0,883. Kejahatan terhadap narkoba (X7) dengan nilai *loading* sebesar 0,932. Kejahatan terhadap penipuan, penggelapan, korupsi (X8) dengan nilai *loading* sebesar 0,929. Sedangkan untuk kejahatan terhadap kesusilaan (X3) mengalami reduksi data karena memiliki nilai *loading* < 0,5 yaitu sebesar 0,419.

4.3. K-Means Clustering

1. Untuk Cluster $k = 2$

K-Means clustering dilakukan pada data klasifikasi kejahatan tahun 2020 pada 34 provinsi yang ada di Indonesia. Berikut data yang akan digunakan untuk perhitungan *clustering*.

Tabel 4.7 Data Kejahatan di Indonesia yang Telah di Reduksi dengan PCA

| No | Provinsi | X1 | X2 | X4 | X5 | X6 | X7 | X8 |
|----|----------------------|----|------|-----|-----|-------|------|------|
| 1 | Sumatera Utara | 99 | 6207 | 55 | 780 | 10916 | 5932 | 5562 |
| 2 | Sumatera Selatan | 77 | 1551 | 141 | 563 | 2911 | 2554 | 1766 |
| 3 | Sumatera Barat | 20 | 1724 | 97 | 185 | 3732 | 913 | 989 |
| 4 | Bengkulu | 33 | 479 | 1 | 106 | 963 | 340 | 556 |
| 5 | Kep. Bangka Belitung | 10 | 150 | 47 | 42 | 690 | 332 | 130 |
| 6 | Jambi | 11 | 531 | 9 | 87 | 17770 | 619 | 767 |
| 7 | Lampung | 21 | 647 | 41 | 399 | 2532 | 1521 | 1373 |
| 8 | Kalimantan Barat | 35 | 184 | 112 | 54 | 1254 | 755 | 431 |
| 9 | Kalimantan Timur | 12 | 352 | 1 | 151 | 862 | 1354 | 356 |
| 10 | Kalimantan Tengah | 17 | 305 | 1 | 86 | 837 | 590 | 315 |
| 11 | Kalimantan Selatan | 38 | 462 | 0 | 367 | 1208 | 1582 | 539 |
| 12 | Kalimantan Utara | 10 | 91 | 10 | 19 | 340 | 245 | 72 |
| 13 | Banten | 8 | 521 | 151 | 94 | 1657 | 634 | 718 |
| 14 | Jawa Barat | 20 | 1445 | 11 | 257 | 2950 | 1141 | 2792 |
| 15 | Jawa Tengah | 30 | 382 | 11 | 236 | 3883 | 1601 | 1814 |
| 16 | Jawa Timur | 55 | 1404 | 177 | 341 | 4976 | 2629 | 2817 |
| 17 | DI Yogyakarta | 6 | 573 | 27 | 126 | 1769 | 460 | 1225 |
| 18 | Bali | 6 | 414 | 4 | 54 | 821 | 668 | 408 |
| 19 | Nusa Tenggara Timur | 40 | 1726 | 14 | 42 | 1494 | 41 | 574 |
| 20 | Nusa Tenggara Barat | 11 | 646 | 46 | 284 | 2120 | 571 | 876 |
| 21 | Gorontalo | 13 | 1043 | 73 | 4 | 696 | 101 | 467 |
| 22 | Sulawesi Barat | 20 | 505 | 97 | 15 | 607 | 175 | 989 |
| 23 | Sulawesi Tengah | 18 | 1391 | 22 | 140 | 2121 | 340 | 677 |
| 24 | Sulawesi Utara | 74 | 1976 | 88 | 72 | 1394 | 100 | 945 |
| 25 | Sulawesi Tenggara | 26 | 734 | 9 | 42 | 632 | 192 | 122 |
| 26 | Sulawesi Selatan | 71 | 3848 | 311 | 348 | 2911 | 1013 | 1786 |
| 27 | Maluku | 10 | 993 | 119 | 23 | 894 | 142 | 307 |
| 28 | Maluku Utara | 1 | 255 | 13 | 14 | 112 | 120 | 49 |
| 29 | Papua Barat | 10 | 655 | 124 | 155 | 1301 | 120 | 356 |
| 30 | Papua | 27 | 1296 | 176 | 532 | 3678 | 241 | 701 |
| 31 | DKI Jakarta | 32 | 1449 | 31 | 259 | 3153 | 5981 | 4595 |
| 32 | Aceh | 15 | 1206 | 29 | 145 | 2408 | 1596 | 1130 |
| 33 | Riau | 21 | 1013 | 11 | 413 | 3152 | 1646 | 1095 |
| 34 | Kepulauan Riau | 13 | 514 | 1 | 103 | 1003 | 362 | 543 |

Selanjutnya data diatas akan dilakukan perhitungan *K-Means Clustering*

dengan inialisasi jumlah k cluster sebanyak 2 cluster. Setelah didapatkan jumlah cluster, tentukan *centroid* awal secara random seperti pada Tabel 4.8

Tabel 4.8 Centroid Awal

| No | Centroid | X1 | X2 | X4 | X5 | X6 | X7 | X8 |
|----|----------|----|-----|----|-----|------|-----|-----|
| 1 | C1 | 11 | 531 | 9 | 87 | 1770 | 619 | 767 |
| 2 | C2 | 11 | 646 | 46 | 284 | 2120 | 571 | 876 |

Setelah *centroid* awal diketahui, lakukan perhitungan jarak masing-masing data dengan kedua *centroid* menggunakan rumus jarak *euclidean* seperti pada Persamaan (2.1)

$$\begin{aligned}(X_1, C_1) &= \sqrt{(99 - 11)^2 + (6207 - 531)^2 + (55 - 9)^2 + (780 - 87)^2 +} \\ &\quad \sqrt{(10916 - 1770)^2 + (5932 - 619)^2 + (5562 - 767)^2} \\ &= 12945,13\end{aligned}$$

$$\begin{aligned}(X_1, C_2) &= \sqrt{(99 - 11)^2 + (6207 - 646)^2 + (55 - 46)^2 + (780 - 284)^2 +} \\ &\quad \sqrt{(10916 - 2120)^2 + (5932 - 571)^2 + (5562 - 876)^2} \\ &= 12619,31\end{aligned}$$

Perhitungan dilakukan sampai data ke-34. Hasil perhitungan semua data yaitu:

Tabel 4.9 Jarak Euclidean Iterasi ke-1

| No | Provinsi | C1 | C2 | Cluster |
|----|----------------------|----------------|-----------------|---------|
| 1 | Sumatera Utara | 12945,13 | 12619,31 | 2 |
| 2 | Sumatera Selatan | 2707,93 | 2502,06 | 2 |
| 3 | Sumatera Barat | 2329,34 | 1975,56 | 2 |
| 4 | Bengkulu | 881,60 | 1247,59 | 1 |
| 5 | Kep. Bangka Belitung | 1242,82 | 1721,37 | 1 |
| 6 | Jambi | 0,00 | 435,99 | 1 |
| 7 | Lampung | 1368,72 | 1154,38 | 2 |
| 8 | Kalimantan Barat | 728,24 | 1119,43 | 1 |

| | | | | |
|----|---------------------|----------------|----------------|---|
| 9 | Kalimantan Timur | 1252,92 | 1603,81 | 1 |
| 10 | Kalimantan Tengah | 1061,51 | 1455,58 | 1 |
| 11 | Kalimantan Selatan | 1174,38 | 1418,11 | 1 |
| 12 | Kalimantan Utara | 1692,93 | 2073,80 | 1 |
| 13 | Banten | 188,99 | 553,23 | 1 |
| 14 | Jawa Barat | 2574,86 | 2307,66 | 2 |
| 15 | Jawa Tengah | 2563,21 | 2263,29 | 2 |
| 16 | Jawa Timur | 4402,03 | 4093,47 | 2 |
| 17 | DI Yogyakarta | 488,55 | 536,66 | 1 |
| 18 | Bali | 1023,09 | 1422,80 | 1 |
| 19 | Nusa Tenggara Timur | 1370,56 | 1410,96 | 1 |
| 20 | Nusa Tenggara Barat | 435,99 | 0,00 | 2 |
| 21 | Gorontalo | 1336,01 | 1628,71 | 1 |
| 22 | Sulawesi Barat | 1269,91 | 1598,02 | 1 |
| 23 | Sulawesi Tengah | 975,59 | 818,14 | 2 |
| 24 | Sulawesi Utara | 1594,01 | 1604,14 | 1 |
| 25 | Sulawesi Tenggara | 1391,71 | 1730,38 | 1 |
| 26 | Sulawesi Selatan | 3696,06 | 3461,19 | 2 |
| 27 | Maluku | 1198,39 | 1484,84 | 1 |
| 28 | Maluku Utara | 1896,08 | 2268,56 | 1 |
| 29 | Papua Barat | 819,21 | 1080,45 | 1 |
| 30 | Papua | 2144,55 | 1751,58 | 2 |
| 31 | DKI Jakarta | 6796,36 | 6694,20 | 2 |
| 32 | Aceh | 1846,87 | 1556,04 | 2 |
| 33 | Riau | 1846,87 | 1556,04 | 2 |
| 34 | Kepulauan Riau | 839,72 | 1206,02 | 1 |

Setelah itu buat *cluster* baru dengan rata-rata dari setiap variabel data didalam *cluster* yang sama.

$$C_{1X_1} = \frac{33 + 10 + 11 + 35 + 12 + 17 + 38 + 10 + 8 + 55 + 6 + 6 + 13 + 20 + 74 + 26 + 10 + 1 + 10 + 13}{20} = 19,65$$

$$C_{2X_1} = \frac{99 + 77 + 20 + 21 + 20 + 30 + 55 + 11 + 18 + 71 + 27 + 32 + 15 + 21}{14} = 36,93$$

Perhitungan tersebut dilakukan sampai di dapat *centroid* baru dari setiap variabel seperti pada Tabel 4.10

Tabel 4.10 *Centroid* Baru Iterasi Ke-2

| No | <i>Centroid</i> | X1 | X2 | X4 | X5 | X6 | X7 | X8 |
|----|-----------------|-------|---------|-------|--------|---------|---------|---------|
| 1 | C1 | 19,65 | 623,15 | 45,05 | 82,80 | 1015,20 | 446,60 | 493,45 |
| 2 | C2 | 36,93 | 1729,21 | 82,79 | 348,71 | 3674,50 | 1977,07 | 1998,07 |

Setelah didapatkan nilai *centroid* baru pada iterasi ke dua, langkah selanjutnya yaitu menghitung jarak data dengan *centroid* dan menentukan jarak terdekatnya. Perhitungan dilakukan dengan cara yang sama pada iterasi sebelumnya, Jika keanggotaan setiap *cluster* tidak mengalami perubahan, maka iterasi dihentikan. Hasil perhitungan jarak data dengan *centroid* ada pada Tabel 4.11

Tabel 4.11 Jarak *Euclidean* Iterasi ke-2

| No | Provinsi | C1 | C2 | Cluster Iterasi Ke-1 | Cluster Iterasi Ke-2 |
|----|----------------------|----------------|-----------------|----------------------|----------------------|
| 1 | Sumatera Utara | 13619,01 | 10051,05 | 2 | 2 |
| 2 | Sumatera Selatan | 3280,03 | 1025,82 | 2 | 2 |
| 3 | Sumatera Barat | 3011,50 | 1476,85 | 2 | 2 |
| 4 | Bengkulu | 203,56 | 3706,81 | 1 | 1 |
| 5 | Kep. Bangka Belitung | 690,37 | 4206,33 | 1 | 1 |
| 6 | Jambi | 827,14 | 2915,02 | 1 | 1 |
| 7 | Lampung | 2080,67 | 1754,92 | 2 | 2 |
| 8 | Kalimantan Barat | 595,35 | 3504,77 | 1 | 1 |
| 9 | Kalimantan Timur | 972,58 | 3596,93 | 1 | 1 |
| 10 | Kalimantan Tengah | 432,83 | 3861,68 | 1 | 1 |
| 11 | Kalimantan Selatan | 1198,95 | 3159,37 | 1 | 1 |
| 12 | Kalimantan Utara | 981,19 | 4541,72 | 1 | 1 |
| 13 | Banten | 720,67 | 3007,14 | 1 | 1 |
| 14 | Jawa Barat | 3196,23 | 1396,02 | 2 | 2 |
| 15 | Jawa Tengah | 23373,97 | 1432,37 | 2 | 2 |
| 16 | Jawa Timur | 5152,16 | 1704,28 | 2 | 2 |
| 17 | DI Yogyakarta | 1052,83 | 2814,35 | 1 | 1 |
| 18 | Bali | 374,80 | 3769,39 | 1 | 1 |
| 19 | Nusa Tenggara Timur | 1272,62 | 3260,33 | 1 | 1 |
| 20 | Nusa Tenggara Barat | 1193,10 | 2613,83 | 2 | 1 |
| 21 | Gorontalo | 636,66 | 3914,80 | 1 | 1 |

| | | | | | |
|----|------------------------|----------------|----------------|----------|----------|
| 22 | Sulawesi Barat | 712,19 | 3909,71 | 1 | 1 |
| 23 | Sulawesi Tengah | 1364,28 | 2645,86 | 2 | 1 |
| 24 | Sulawesi Utara | 1517,44 | 3157,83 | 1 | 1 |
| 25 | Sulawesi Tenggara | 604,10 | 4129,53 | 1 | 1 |
| 26 | Sulawesi Selatan | 4016,08 | 2469,78 | 2 | 2 |
| 27 | Maluku | 536,81 | 3822,11 | 1 | 1 |
| 28 | Maluku Utara | 1123,22 | 4714,92 | 1 | 1 |
| 29 | Papua Barat | 468,82 | 3601,74 | 1 | 1 |
| 30 | Papua | 2801,35 | 2219,55 | 2 | 2 |
| 31 | DKI Jakarta | 7261,95 | 4810,06 | 2 | 2 |
| 32 | Aceh | 2002,52 | 1679,71 | 2 | 2 |
| 33 | Riau | 2574,59 | 1311,74 | 2 | 2 |
| 34 | Kepulauan Riau | 155,14 | 3661,55 | 1 | 1 |

Dari hasil tersebut, masih ada anggota yang mengalami perpindahan *cluster*, yaitu pada data yang berwarna merah. Pada provinsi Nusa Tenggara Barat dan Sulawesi Tengah yang awalnya pada iterasi pertama berada di *cluster* 2, di iterasi ke-2 berpindah ke *cluster* 1. Karena masih ada perpindahan anggota, maka iterasi dilanjutkan dengan membuat lagi *centroid* baru untuk iterasi ke-3 seperti Tabel 4.12

Tabel 4.12 Centroid Baru Iterasi Ke-3

| No | Centroid | X1 | X2 | X4 | X5 | X6 | X7 | X8 |
|----|----------|-------|---------|-------|--------|---------|---------|---------|
| 1 | C1 | 19,18 | 659,09 | 44,05 | 94,55 | 1115,68 | 447,41 | 519,18 |
| 2 | C2 | 40,67 | 1847,67 | 90,92 | 371,50 | 3933,50 | 2230,67 | 2201,67 |

Setelah *centroid* baru iterasi ke-3 didapatkan, selanjutnya hitung jarak euclidean data dengan *centroid*.

Tabel 4.13 Jarak Euclidean Iterasi ke-3

| No | Provinsi | C1 | C2 | Cluster Iterasi Ke-2 | Cluster Iterasi Ke-3 |
|----|------------------|----------|----------------|----------------------|----------------------|
| 1 | Sumatera Utara | 13520,82 | 9639,62 | 2 | 2 |
| 2 | Sumatera Selatan | 3200,49 | 1211,76 | 2 | 2 |
| 3 | Sumatera Barat | 2903,04 | 1816,03 | 2 | 2 |

| | | | | | |
|----|----------------------|----------------|----------------|----------|----------|
| 4 | Bengkulu | 266,10 | 4130,20 | 1 | 1 |
| 5 | Kep. Bangka Belitung | 779,75 | 4627,15 | 1 | 1 |
| 6 | Jambi | 732,63 | 3340,45 | 1 | 1 |
| 7 | Lampung | 1995,10 | 2144,70 | 2 | 1 |
| 8 | Kalimantan Barat | 594,76 | 3919,39 | 1 | 1 |
| 9 | Kalimantan Timur | 1006,12 | 3987,94 | 1 | 1 |
| 10 | Kalimantan Tengah | 516,72 | 4278,97 | 1 | 1 |
| 11 | Kalimantan Selatan | 1188,10 | 3541,46 | 1 | 1 |
| 12 | Kalimantan Utara | 1082,73 | 4960,77 | 1 | 1 |
| 13 | Banten | 630,88 | 3431,44 | 1 | 1 |
| 14 | Jawa Barat | 3107,51 | 1638,66 | 2 | 2 |
| 15 | Jawa Tengah | 3280,76 | 1649,96 | 2 | 2 |
| 16 | Jawa Timur | 5057,27 | 1352,59 | 2 | 2 |
| 17 | DI Yogyakarta | 966,45 | 3234,89 | 1 | 1 |
| 18 | Bali | 459,73 | 4184,71 | 1 | 1 |
| 19 | Nusa Tenggara Timur | 1205,69 | 3677,55 | 1 | 1 |
| 20 | Nusa Tenggara Barat | 1089,67 | 3042,24 | 1 | 1 |
| 21 | Gorontalo | 674,77 | 4337,00 | 1 | 1 |
| 22 | Sulawesi Barat | 765,88 | 4323,45 | 1 | 1 |
| 23 | Sulawesi Tengah | 1259,11 | 3074,38 | 1 | 1 |
| 24 | Sulawesi Utara | 1455,74 | 3560,24 | 1 | 1 |
| 25 | Sulawesi Tenggara | 683,07 | 4553,77 | 1 | 1 |
| 26 | Sulawesi Selatan | 3931,31 | 2598,51 | 2 | 2 |
| 27 | Maluku | 556,56 | 4247,88 | 1 | 1 |
| 28 | Maluku Utara | 1227,48 | 5134,67 | 1 | 1 |
| 29 | Papua Barat | 422,28 | 4032,72 | 1 | 1 |
| 30 | Papua | 2693,63 | 2571,69 | 2 | 2 |
| 31 | DKI Jakarta | 7213,52 | 4536,24 | 2 | 2 |
| 32 | Aceh | 1914,25 | 2084,69 | 2 | 1 |
| 33 | Riau | 2478,41 | 1697,78 | 2 | 2 |
| 34 | Kepulauan Riau | 208,74 | 4085,40 | 1 | 1 |

Pada iterasi ke-3 pun masih ada data yang mengalami perpindahan *cluster*, maka iterasi akan dilanjutkan sampai tidak ada data yang mengalami perpindahan *cluster*. Setelah dilakukan beberapa proses iterasi, ternyata pada iterasi ke-7 data sudah tidak ada yang mengalami perpindahan *cluster*,

Sehingga iterasi dihentikan.

Tabel 4.14 Centroid Baru Iterasi Ke-7

| No | Centroid | X1 | X2 | X4 | X5 | X6 | X7 | X8 |
|----|----------|-------|---------|--------|--------|---------|---------|---------|
| 1 | C1 | 19,48 | 755,04 | 49,00 | 139,04 | 1483,22 | 583,70 | 618,89 |
| 2 | C2 | 54,86 | 2326,57 | 105,29 | 397,71 | 4528,57 | 2978,71 | 3018,86 |

Setelah *centroid* baru iterasi ke-7 didapatkan, selanjutnya hitung jarak *euclidean* iterasi ke-7.

Tabel 4.15 Jarak Euclidean Iterasi ke-7

| No | Provinsi | C1 | C2 | Cluster Iterasi Ke-6 | Cluster Iterasi Ke-7 |
|----|----------------------|----------------|----------------|----------------------|----------------------|
| 1 | Sumatera Utara | 13120,86 | 8437,84 | 2 | 2 |
| 2 | Sumatera Selatan | 2839,29 | 2235,43 | 2 | 2 |
| 3 | Sumatera Barat | 2499,15 | 3071,10 | 1 | 1 |
| 4 | Bengkulu | 643,24 | 5408,49 | 1 | 1 |
| 5 | Kep. Bangka Belitung | 1143,31 | 5912,24 | 1 | 1 |
| 6 | Jambi | 400,00 | 4645,50 | 1 | 1 |
| 7 | Lampung | 1620,64 | 3412,67 | 1 | 1 |
| 8 | Kalimantan Barat | 674,32 | 5203,26 | 1 | 1 |
| 9 | Kalimantan Timur | 1101,51 | 5210,26 | 1 | 1 |
| 10 | Kalimantan Tengah | 847,14 | 5553,44 | 1 | 1 |
| 11 | Kalimantan Selatan | 1104,22 | 4755,59 | 1 | 1 |
| 12 | Kalimantan Utara | 1475,71 | 6233,26 | 1 | 1 |
| 13 | Banten | 331,50 | 4732,25 | 1 | 1 |
| 14 | Jawa Barat | 2770,53 | 2593,77 | 2 | 2 |
| 15 | Jawa Tengah | 2893,48 | 2753,83 | 2 | 2 |
| 16 | Jawa Timur | 4657,68 | 1105,77 | 2 | 2 |
| 17 | DI Yogyakarta | 705,91 | 4509,37 | 1 | 1 |
| 18 | Bali | 784,77 | 5448,92 | 1 | 1 |
| 19 | Nusa Tenggara Timur | 1118,25 | 4930,71 | 1 | 1 |
| 20 | Nusa Tenggara Barat | 710,45 | 4362,66 | 1 | 1 |
| 21 | Gorontalo | 988,72 | 5593,49 | 1 | 1 |
| 22 | Sulawesi Barat | 1073,31 | 5552,09 | 1 | 1 |
| 23 | Sulawesi Tengah | 935,26 | 4381,04 | 1 | 1 |
| 24 | Sulawesi Utara | 1359,42 | 4758,47 | 1 | 1 |
| 25 | Sulawesi Tenggara | 1066,03 | 5832,13 | 1 | 1 |
| 26 | Sulawesi Selatan | 3642,30 | 3218,74 | 2 | 2 |

| | | | | | |
|----|----------------|----------------|----------------|---|---|
| 27 | Maluku | 845,36 | 5525,60 | 1 | 1 |
| 28 | Maluku Utara | 1639,32 | 6399,10 | 1 | 1 |
| 29 | Papua Barat | 577,33 | 5341,88 | 1 | 1 |
| 30 | Papua | 2324,75 | 3831,04 | 1 | 1 |
| 31 | DKI Jakarta | 6944,41 | 3766,37 | 2 | 2 |
| 32 | Aceh | 1531,34 | 3362,06 | 1 | 1 |
| 33 | Riau | 2069,55 | 3017,95 | 1 | 1 |
| 34 | Kepulauan Riau | 589,29 | 5365,78 | 1 | 1 |

Dari iterasi ke-7 sudah tidak ada data yang mengalami perpindahan *cluster*, oleh karena itu iterasi dihentikan. Jumlah data pada *cluster* 1 sebanyak 27 anggota dan *cluster* 2 memiliki 7 anggota. Dari hasil *centroid* pada Tabel 4.12 dilakukan pembulatan untuk mengetahui karakteristik *cluster*. Hasil pembulatan *centroid* iterasi ke-7 seperti pada Tabel 4.16

Tabel 4.16 Hasil Pembulatan *Centroid* Iterasi ke-7

| No | <i>Centroid</i> | X1 | X2 | X4 | X5 | X6 | X7 | X8 |
|----|-----------------|----|------|-----|-----|------|------|------|
| 1 | C1 | 19 | 755 | 49 | 139 | 1483 | 584 | 619 |
| 2 | C2 | 55 | 2327 | 105 | 398 | 4529 | 2979 | 3019 |

Dari hasil pembulatan tersebut pada *cluster* 1 (Rawan) terdapat 19 kasus kejahatan terhadap nyawa, 755 kasus kejahatan terhadap fisik/badan, 49 kasus kejahatan terhadap kemerdekaan orang, 139 kasus kejahatan terhadap hak milik dengan kekerasan, 1483 kasus kejahatan terhadap hak milik tanpa kekerasan, 584 kejahatan terhadap narkoba, dan 619 kasus kejahatan terhadap penipuan, penggelapan, korupsi.

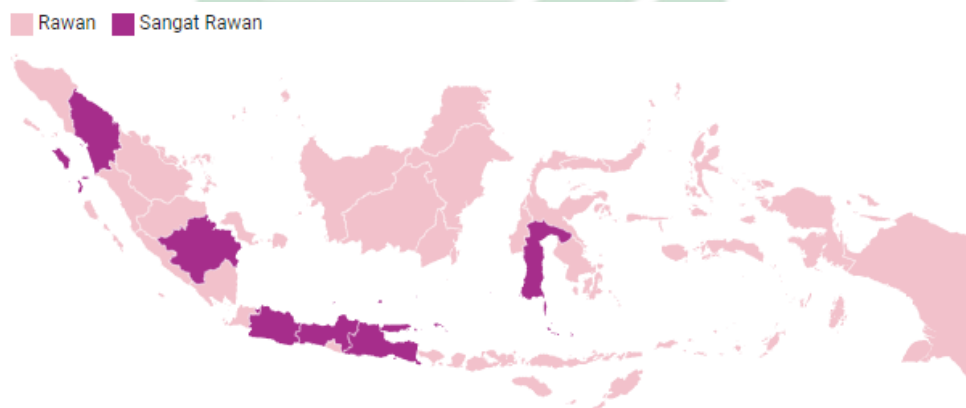
Sedangkan pada *cluster* 2 (Sangat Rawan) terdapat 55 kasus kejahatan terhadap nyawa, 2327 kasus kejahatan terhadap fisik/badan, 105 kasus kejahatan terhadap kemerdekaan orang, 398 kasus kejahatan terhadap hak milik dengan kekerasan, 4529 kasus kejahatan terhadap hak milik tanpa

kekerasan, 2979 kejahatan terhadap narkoba, dan 3019 kasus kejahatan terhadap penipuan, penggelapan, korupsi. Anggota dari setiap *cluster* dapat dilihat pada Tabel 4.17

Tabel 4.17 Tabel Keanggotaan Hasil 2 Cluster

| <i>Cluster</i> | Jumlah Anggota | Anggota |
|--------------------------|----------------|--|
| Rawan (C1) | 27 | Sumatera Barat, Bengkulu, Kep. Bangka Belitung, Jambi, Lampung, Kalimantan Barat, Kalimantan Timur, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Utara, Banten, DI Yogyakarta, Bali, Nusa Tenggara Timur, Nusa Tenggara Barat, Gorontalo, Sulawesi Barat, Sulawesi Tengah, Sulawesi Utara, Sulawesi Tenggara, Maluku, Maluku Utara, Papua Barat, Papua, Aceh, Riau, Kep. Riau |
| Sangat Rawan (C2) | 7 | Sumatera Utara, Sumatera Selatan, Jawa Barat, Jawa Tengah, Jawa Timur, Sulawesi Selatan, DKI Jakarta |

Hasil *cluster* divisualisasikan dalam bentuk peta pada Gambar 4.1



Gambar 4.1 Pemetaan Clustering Tingkat Kejahatan di Indonesia dengan 2 Cluster

Dari persamaan (2.11) didapatkan nilai *Silhouette Coefficient* untuk hasil 2 *cluster* yaitu sebesar 0,51 yang artinya keakuratan hasil *clustering*nya termasuk kedalam kategori standar.

2. Untuk Cluster $k = 3$

Dengan langkah-langkah yang sama seperti percobaan dengan menggunakan 2 *cluster*. Hasil pada percobaan 3 *cluster* berhenti pada iterasi ke-2 dengan hasil pembulatan *centroid* seperti pada Tabel. 4.18

Tabel 4.18 Hasil Pembulatan *Centroid* Iterasi ke-2

| No | <i>Centroid</i> | X1 | X2 | X4 | X5 | X6 | X7 | X8 |
|----|-----------------|----|------|----|-----|-------|------|------|
| 1 | C1 | 19 | 659 | 44 | 95 | 1116 | 447 | 519 |
| 2 | C2 | 35 | 1451 | 94 | 334 | 3299 | 1894 | 1896 |
| 2 | C3 | 99 | 6207 | 55 | 780 | 10916 | 5932 | 5562 |

Dari hasil pembulatan *centroid* tersebut pada *cluster* 1 Aman terdapat 19 kasus kejahatan terhadap nyawa, 659 kasus kejahatan terhadap fisik/badan, 44 kasus kejahatan terhadap kemerdekaan orang, 95 kasus kejahatan terhadap hak milik dengan kekerasan, 1116 kasus kejahatan terhadap hak milik tanpa kekerasan, 447 kejahatan terhadap narkoba, dan 519 kasus kejahatan terhadap penipuan, penggelapan, korupsi.

Pada *cluster* 2 Rawan terdapat 35 kasus kejahatan terhadap nyawa, 1451 kasus kejahatan terhadap fisik/badan, 94 kasus kejahatan terhadap kemerdekaan orang, 334 kasus kejahatan terhadap hak milik dengan kekerasan, 3299 kasus kejahatan terhadap hak milik tanpa kekerasan, 1894 kejahatan terhadap narkoba, dan 1896 kasus kejahatan terhadap penipuan, penggelapan, korupsi.

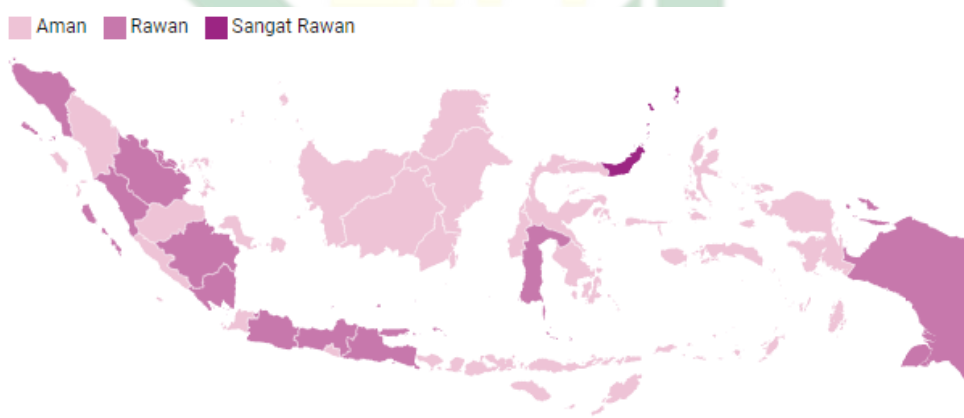
Pada *cluster* 3 Sangat Rawan terdapat 99 kasus kejahatan terhadap nyawa, 6207 kasus kejahatan terhadap fisik/badan, 55 kasus kejahatan terhadap kemerdekaan orang, 780 kasus kejahatan terhadap hak milik dengan kekerasan, 10916 kasus kejahatan terhadap hak milik tanpa kekerasan, 5932 kejahatan terhadap narkoba, dan 5562 kasus kejahatan terhadap penipuan, penggelapan, korupsi. Keanggotaan dari setiap cluster dapat dilihat pada

Tabel 4.19

Tabel 4.19 Tabel Keanggotaan Hasil 3 Cluster

| <i>Cluster</i> | <i>Jumlah Anggota</i> | <i>Anggota</i> |
|--------------------------|-----------------------|--|
| Aman (C1) | 22 | Bengkulu, Kep. Bangka Belitung, Jambi, Banten, Kalimantan Barat, Kalimantan Timur, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Utara, DI Yogyakarta, Bali, Nusa Tenggara Timur, Nusa Tenggara Barat, Gorontalo, Sulawesi Barat, Sulawesi Tengah, Sulawesi Utara, Sulawesi Tenggara, Maluku, Maluku Utara, Papua Barat, Kep. Riau |
| Rawan (C2) | 11 | Sumatera Selatan, Sumatera Barat, Lampung, Jawa Barat, Jawa Tengah, Jawa Timur, Sulawesi Selatan, DKI Jakarta, Riau, Aceh, Papua |
| Sangat Rawan (C3) | 1 | Sumatera Utara |

Jika hasil tersebut dipetakan, maka akan seperti pada Gambar 4.2

**Gambar 4.2 Pemetaan Clustering Tingkat Kejahatan di Indonesia dengan 3 Cluster**

Dari persamaan (2.11) didapatkan nilai *Silhouette Coefficient* untuk hasil 3 cluster yaitu sebesar 0,77 yang artinya keakuratan hasil *clusteringnya* termasuk kedalam kategori kuat.

4.4. Perbandingan Hasil $k = 2$ dan $k = 3$

Perbandingan antara hasil 2 dan 3 *cluster* untuk *clustering* tingkat kerawanan kriminalitas di Indonesia dengan menggunakan metode *K-Means* dapat dilihat mana yang lebih akurat dengan melihat nilai *silhouette coefficient*nya seperti pada Tabel 4.20

Tabel 4.20 Perbandingan Keakuratan Hasil 2 dan 3 Cluster

| Jumlah Cluster | Silhouette Coefficient | Kategori Keakuratan |
|----------------|------------------------|---------------------|
| 2 | 0,51 | Standar |
| 3 | 0,77 | Kuat |

Nilai *silhouette coefficient* hasil 3 cluster lebih tinggi yaitu sebesar 0,77 yang termasuk ke dalam kategori kuat daripada hasil dari percobaan 2 cluster dengan nilai 0,51 yang termasuk ke dalam kategori standar. Artinya *clustering* tingkat kerawanan terhadap kriminalitas di Indonesia menggunakan metode *K-Means clustering* lebih baik menggunakan 3 cluster. Dimana cluster 1 aman merupakan wilayah yang berpotensi aman dari tindak kejahatan karena minimnya jumlah kasus kejahatan yang terjadi, cluster 2 rawan merupakan wilayah yang berpotensi rawan akan kejahatan karena banyaknya kasus kejahatan yang terjadi, dan cluster 3 adalah wilayah yang memiliki potensi paling tinggi terjadinya kejahatan karena sangat banyak kasus kejahatan terjadi di daerah tersebut.

4.5. Integrasi Keilmuan

Dari hasil penelitian ini dapat memberikan kita peringatan bahwa masih banyak daerah yang sangat rawan terjadi tindak kejahatan. Selain itu penelitian ini juga sebagai himbauan, jauhilah faktor-faktor penyebab kejahatan misalnya minum

minuman keras dan berjudi. Allah berfirman dalam surat Al-Maidah Ayat 91 yang berbunyi:

إِنَّمَا يُرِيدُ الشَّيْطَانُ أَنْ يُوقِعَ بَيْنَكُمُ الْعَدَاوَةَ وَالْبَغْضَاءَ فِي
الْخَمْرِ وَالْمَيْسِرِ وَيَصُدَّكُمْ عَنْ ذِكْرِ اللَّهِ وَعَنِ الصَّلَاةِ فَهَلْ أَنْتُمْ
مُنْتَهُونَ

Artinya: "Dengan minuman keras dan judi itu, setan hanyalah bermaksud menimbulkan permusuhan dan kebencian di antara kamu, dan menghalang-halangi kamu dari mengingat Allah dan melaksanakan salat, maka tidakkah kamu mau berhenti?"

Dari ayat tersebut, Allah telah memperingatkan untuk menjauhi perbuatan tercela yang menyebabkan dampak yang lebih parah seperti dikarenakan minum minuman keras yang membuat seseorang tidak bisa berfikir jernih, lupa terhadap Allah, melalaikan salat dan juga timbul rasa ingin melakukan tindak kejahatan. Selain itu selalu berhati-hatilah disetiap waktu, karena kejahatan bisa terjadi kapan saja dan dimana saja. Selain waspada, taat lah kepada Allah SWT dan Rasulullah SAW agar selalu mendapat perlindungan. Allah berfirman dalam lanjutan surat Al-Maidah ayat 91 yaitu surat Al-Maidah ayat 92 yang berbunyi:

وَاطِيعُوا اللَّهَ وَاطِيعُوا الرَّسُولَ وَاحْذَرُوا فَإِن تَوَلَّيْتُمْ فَأَعْلَمُوا
أَنَّمَا عَلَى رَسُولِنَا الْبَلْغُ الْمُبِينُ

Artinya : "Dan taatlah kamu kepada Allah dan taatlah kamu kepada Rasul serta berhati-hatilah. Jika kamu berpaling, maka ketahuilah bahwa kewajiban Rasul Kami hanyalah menyampaikan (amanat) dengan jelas."

Dari dua ayat pada surat Al-Maidah tersebut memperingatkan selalu berhati-hati untuk menjauhi perbuatan-perbuatan tercela yang menyebabkan pemicu terjadinya tindak kejahatan dan senantiasa taat kepada Allah SWT dan Rasulullah SAW agar selalu mendapat perlindungan.



UIN SUNAN AMPEL
S U R A B A Y A

BAB V

PENUTUP

5.1. Kesimpulan

Hasil penelitian penerapan metode *K-Means* untuk *clustering* tingkat kejahatan di Indonesia dapat diambil kesimpulan yaitu:

1. Hasil *clustering* tingkat kejahatan di Indonesia menggunakan metode *K-Means* didapatkan cluster optimalnya yaitu 3 *cluster*. Didapatkan 22 anggota pada *cluster* 1 Aman (C1), 11 anggota pada *cluster* 2 Rawan (C2). Daerah yang termasuk ke dalam cluster 1 Aman yaitu Bengkulu, Kep. Bangka Belitung, Jambi, Kalimantan Barat, Kalimantan Timur, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Utara, Banten, DI Yogyakarta, Bali, Nusa Tenggara Timur, Nusa Tenggara Barat, Gorontalo, Sulawesi Barat, Sulawesi Tengah, Sulawesi Utara, Sulawesi Tenggara, Maluku, Maluku Utara, Papua Barat, dan Kep. Riau. Lalu anggota pada *cluster* 2 Sangat Rawan yaitu Sumatera Utara. Sedangkan daerah yang termasuk kedalam anggota cluster 3 Rawan yaitu Sumatera Selatan, Sumatera Barat, Jawa Barat, Jawa Tengah, Jawa Timur, Sulawesi Selatan, dan DKI Jakarta, Riau, Aceh, dan Papua, dan 1 anggota pada *cluster* 3 Sangat Rawan (C3)
2. Keakuratan hasil *clustering* dengan 3 *cluster* yaitu masuk dalam kategori kuat dengan nilai *silhouette coefficient* sebesar 0,77.

5.2. Saran

Saran yang disampaikan peneliti yang ingin melanjutkan ke penelitian selanjutnya adalah sebagai berikut.

1. Sebagai pembandingan hasil percobaan yang lebih variatif pada beberapa *k cluster* misalnya percobaan perhitungan pada 2, 3, 4, dan 5 *cluster*.
2. Faktor penyebab kejahatan bisa berubah sewaktu-waktu, oleh karena itu perlu dianalisis ulang disetiap tahunnya.
3. Untuk pengembangan penelitian selanjutnya, pemetaan tindak kejahatan disetiap provinsi dibuat sedetail mungkin supaya terlihat jelas daerah mana yang rawan dan sangat rawan tindak kejahatan.

Saran untuk kepolisian di Indonesia agar memperketat keamanan disetiap daerah, khususnya daerah yang sangat rawan terjadi tindak kejahatan dan melakukan tindakan tegas kepada pelaku kejahatan supaya menimbulkan efek jera. Dengan begitu perlahan-lahan kasus kejahatan di Indonesia akan menurun.

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR PUSTAKA

- Ali Hasymi, M., Faisol, A., and Ariwibisono, F. (2021). Sistem Informasi Geografis Pemetaan Warga Kurang Mampu Di Kelurahan Karang Besuki Menggunakan Metode K-Means Clustering. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 5(1):284–290.
- Annur, H. (2019). Penerapan Data Mining Menentukan Strategi Penjualan Variasi Mobil Menggunakan Metode K-Means Clustering. *Jurnal Informatika Upgris*, 5(1).
- Antara, M. and Yogantari, M. V. (2018). Keragaman Budaya Indonesia Sumber Inspirasi Inovasi Industri Kreatif. *Senada*, 1:292–301.
- ANTARA SUMUT (2020). Babinsa Gagalkan Penculikan Anak di Medan.
- Anwar, K., Goejantoro, R., and Prangga, S. (2022). Pengelompokan Kabupaten/Kota Di Pulau Kalimantan Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2020 Menggunakan Optimasi K-Means Cluster Dengan Principle Component Analysis (PCA) Grouping Of Regencies/Cities in Kalimantan Island Based On Human Development Index Indicators In 2020 Using K-Means Cluster Optimization Using Principle Component Analysis (PCA). 13:131–140.
- Astuti, W., Widodo, A., Elektro, J. T., Teknik, F., and Semarang, U. N. (2016). Pemetaan Tindak Kejahatan Jalanan di Kota Semarang Menggunakan Algoritma K-Means Clustering. *Jurnal Teknik Elektro*, 8(1):5–7.

- Atma, H., Putra, A. R., and Enri, Ultach, N. (2021). Clustering Clustering Data Eskspor Buah-Buahan Berdasarkan Negara Tujuan Menggunakan Algoritma K-Means. *Bina Insani Ict Journal*, 8(1):73.
- Cahyati, E. D., Herawatie, D., and Wuryanto, E. (2017). Implementasi K-Means Clustering Untuk Pemetaan Desa Dan Kelurahan Di Kabupaten Bangkalan Berdasarkan Contraceptive Prevalence Rate Dan Tingkat Pendidikan. *Seminar Nasional Matematika dan Aplikasinya*, pages 341–348.
- CNN Indonesia (2022). Surya Darmadi Didakwa Rugikan Negara Hampir Rp80 Triliun.
- detikcom, T. (2023). Hakim Ungkap Awal Mula AG Bersama Mario Dandy Aniaya David dengan Sadis.
- Farida, Y., Fahmi Khariri, A., Yuliati, D., Khaulasari, H., Sunan Ampel, U., Farida, Y., Khariri, A., Yuliati, D., and Khaulasari, H. (2022). Clustering Couples of Childbearing Age to Get Family Planning Counseling Using K-Means Method. *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 22(1):189–200.
- Febrianti, F., Hafiyusholeh, M., and Asyhar, A. H. (2016). Perbandingan Pengklusteran Data Iris Menggunakan Metode K-Means Dan Fuzzy C-Means. *Jurnal Matematika "MANTIK"*, 2(1):7.
- Hartono, W., Hadi, S., Rosnawati, R., and Retnawati, H. (2022). Uji Kecocokan Model Parameter Logistik Soal Diagnosa Kemampuan Matematika Dasar. *JNPM (Jurnal Nasional Pendidikan Matematika)*, 6(1):125.

- KOMPAS (2020). Kaleidoskop 2020: 5 Kasus Kriminal di Sejumlah Daerah yang Menyita Perhatian Publik.
- Maharani, S. D. (2016). Manusia Sebagai Homo Economicus: Refleksi Atas Kasus-Kasus Kejahatan Di Indonesia. *Jurnal Filsafat*, 26(1):30.
- Martiano, M., Zarlis, M., and Wage, S. (2019). Modification K-Means Model With Local Deviation Method To Improve The Accuracy In Forming Clusters.
- Muhson, A. and UNODC (2013). International Classification of Crime for Statistical Purposes (Iccs). *Makalah Teknik Analisis II*, (03):1–7.
- Nabila, Z., Rahman Isnain, A., and Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi dan Sistem Informasi (JTSI)*, 2(2):100.
- Nugroho, A. H. and Harmadi, S. H. B. (2015). Analisis Spasial Kriminalitas Harta Benda di Wilayah Jabetabek Spatial Analysis on Property Crime in Jabetabek. *Jurnal Ekonomi dan Pembangunan Indonesia*, 15(2):158–172.
- Nurjoko, Dwirohayati, D., and Sudiby, N. H. (2020). Sistem Informasi Pemetaan Wilayah Rawan Kriminalitas Polresta Bandar Lampung Menggunakan K-Means Clustering. *Teknika*, 14(2):127–135.
- Octavianti, R. D., Santoso, N., and Romlah, S. (2017). Pemetaan Data Kriminalitas Di Kota Malang Berbasis Webgis. *Jurnal Informatika Polinema*, 1(3):41.
- Oktaviana, E. (2022). *Clustering bencana alam di Indonesia menggunakan algoritma K-Means*.
- Pradana, A. (2019). Pemetaan daerah kerawanan kriminalitas menggunakan.

- Prasetyo, A. D. and Abduh, M. (2021). Peningkatan Keaktifan Belajar Siswa Melalui Model Discovery Learning Di Sekolah Dasar. *Jurnal Basicedu*, 5(4):1717–1724.
- Putra, S. F., Pradina, R., and Hafidz, I. (2016). 16035-42097-1-Pb. *Jurnal Teknik*, 5(2):5–9.
- Rinanda, H. M. (2023). 7 Fakta Sadis Gangster di Sidoarjo Keroyok Remaja hingga Tewas.
- Rizal Muzaky, D. (2020). Prediksi Penjualan Barang Dagangan di Toko ”Hadi Damis” Dengan Metode Fuzzy C-Means. *Gastronomía ecuatoriana y turismo local.*, 1(69):5–24.
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20(C):53–65.
- Sholihah, S. A. (2021). Analisis Cluster Untuk Pemetaan Data Kasus Covid-19 Indonesia Menggunakan K-Means.
- Struyf, A., Hubert, M., and Rousseeuw, P. J. (1996). Clustering in an object-oriented environment. *Journal of Statistical Software*, 1:1–30.
- Surbakti & Natangsa, S. (2017). *Hukum Pidana Dasar-Dasar Hukum Pidana Berdasarkan KUHP dan RUU KUHP*, volume 39.
- Suryadi, U. T. and Supriatna, Y. (2019). Sistem Clustering Tindak Kejahatan Pencurian Di Wilayah Jawa Barat Menggunakan Algoritma K-Means. *Jurnal Teknologi dan Komunikasi STMIK Subang*, 12(1):15–27.