

CLUSTERING KASUS KEPOLISIAN DENGAN METODE *K-MEANS*

(Studi Kasus: Polsek Kota Gorontalo)

Oleh

NANCY M. HILIPITO

T3117071

SKRIPSI

Untuk memenuhi salah satu syarat ujian
guna memperoleh gelar Sarjana



**PROGRAM SARJANA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS ICHSAN GORONTALO
GORONTALO
2022**

LEMBAR PERSETUJUAN SKRIPSI
CLUSTERING KASUS KEPOLISIAN DENGAN
METODE K-MEANS
(Studi Kasus : Polsek Kota Gorontalo)

Oleh

NANCY M. HILIPITO

T3117071

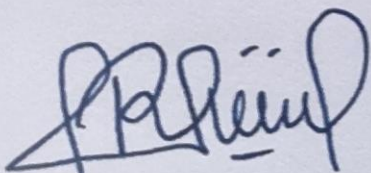
SKRIPSI

Untuk memenuhi salah satu syarat ujian
guna memperoleh gelar Sarjana Program Studi Teknik Informatika.

Ini telah disetujui oleh pembimbing

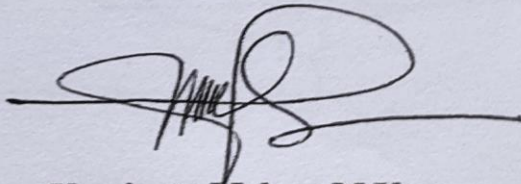
Gorontalo, Juni 2022

Pembimbing I



Rezwati Ishak M. Kom
NIDN: 0903087901

Pembimbing II



Yusrianto Malago M. Kom
NIDN: 0909108901

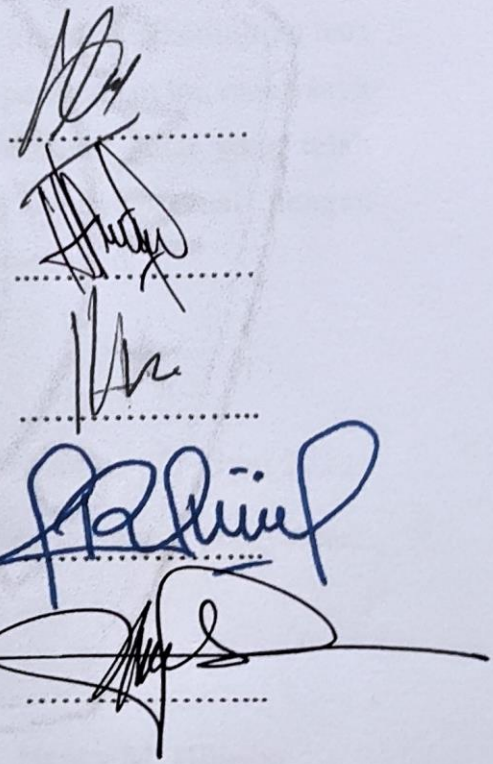
LEMBAR PENGESAHAN SKRIPSI
CLUSTERING KASUS KEPOLISIAN DENGAN
METODE *K-MEANS*

OLEH
NANCY M. HILIPITO

T3117071

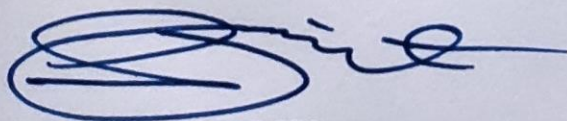
Di Periksa oleh Panitia Ujian Strata Satu (S1)
Universitas Ichsan Gorontalo
Gorontalo, Juni 2022

1. Ketua Penguji
Husdi, M. Kom
2. Anggota Penguji
Sudirman Melangi, M.Kom
3. Anggota Penguji
Mohamad Efendy Lasulika, M.Kom
4. Anggota Penguji
Rezqiwati Ishak, M.Kom
5. Anggota Penguji
Yusrianto Malago, M.Kom



Mengetahui :

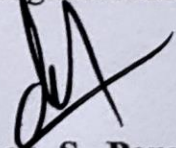
Dekan Fakultas Ilmu Komputer



Jorry Karim, M.Kom

NIDN.0918077302

Ketua Program Studi



Sudirman S. Panna

NIDN.0924038205

LEMBAR PERNYATAAN SKRIPSI

Dengan ini saya menyatakan bahwa :

1. Karya tulis (Skripsi) saya ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana) baik di Universitas Ichsan Gorontalo maupun di perguruan tinggi lainnya.
2. Karya tulis (Skripsi) saya ini adalah murni gagasan, rumusan, dan penelitian saya sendiri, tanpa bantuan pihak lain, kecuali dari arahan Tim Pembimbing.
3. Dalam karya tulis (Skripsi) saya ini tidak terdapat karya atau pendapat yang telah dipublikasi orang lain, kecuali serta tertulis dicantumkan sebagai acauan/ sitasi dalam naskah dan dicantumkan pula dalam daftar pustaka.
4. Pernyataan isi saya buat dengan sesungguhnya dan apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya tulis ini serta sanksi lainnya sesuai dengan normanorma yang berlaku di Universitas Ichsan Gorontalo.

Gorontalo, Juni 2022

Yang membuat pernyataan



Nancy M. Hilipito

ABSTRACT

NANCY M HILIPITO. T3117071. THE CLUSTERING OF CASES UNDER POLICE HANDLING USING THE K-MEANS METHOD

This study is aimed to group cases under police handling by the Gorontalo City Subregional Police. A 'case' here means the actual state of a case related to someone or something else. With the number of crime cases in Gorontalo City, it is necessary to have a computerized system of regional grouping to assist the police in making decisions on whether an area requires extra supervision. The optimum number of clusters to carry out clustering cases under police handling using the K-Means method is five clusters through the Elbow technique, namely areas for cases of theft and violence with a small number, areas for cases of kidnapping and raping with low numbers, areas for cases of kidnapping, raping, theft, moderate amounts of violence, and areas for cases of theft and violence cases with large numbers.

Keywords: Clustering, Cases, K-Means



ABSTRAK

NANCY M. HILIPITO. T3117071. CLUSTERING KASUS KEPOLISIAN DENGAN METODE *K-MEANS*.

Penelitian ini bertujuan untuk mengelompokkan kasus kepolisian di Polsek Kota Gorontalo. Kasus merupakan keadaan yang sebenarnya dari suatu perkara yang berhubungan dengan seseorang atau hal lain. Banyaknya kasus kejahatan di Kota Gorontalo, perlu adanya sebuah sistem komputerisasi pengelompokan daerah untuk membantu kepolisian dalam mengambil keputusan apakah suatu daerah memerlukan pengawasan ekstra atau tidak. Jumlah cluster yang optimum untuk melakukan Clustering Kasus Kepolisian Dengan Metode *K-Means* yaitu sebanyak 5 cluster dengan menggunakan teknik Elbow yaitu, daerah kasus pencurian dan kekerasan dengan jumlah kecil, daerah kasus penculikan dan pemerkosaan dengan jumlah rendah, daerah kasus penculikan, pemerkosaan, pencurian, kekerasan dengan jumlah sedang, dan daerah kasus pencurian dan kekerasan dengan jumlah banyak.

Kata Kunci : Clustering, Kasus, K-Means

KATA PENGANTAR

Assalamualaikum warahmatullahi wabarakatuh.

Alhamdulillah penulis dapat menyelesaikan Skripsi dengan judul **“CLUSTERING KASUS KEPOLISIAN DENGAN METODE *K-MEANS*”**. untuk memenuhi salah satu syarat penyusunan Skripsi Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.

Penulis menyadari sepenuhnya bahwa Skripsi ini tidak mungkin terwujud tanpa bantuan dan dorongan dari berbagai pihak, baik bantuan moril maupun materil. Untuk itu, dengan segala keikhlasan dan kerendahan hati, penulis mengucapkan banyak terima kasih dan penghargaan yang setinggi-tingginya kepada :

1. Ibu Dr. Dra. Juriko Abdussamad, M.Si, selaku Ketua Yayasan Pengembangan Ilmu Pengetahuan dan Teknologi (YPIPT) Ichsan Gorontalo;
2. Bapak Dr. Abdul Gaffar La Tjokke M.Si, selaku Rektor Universitas Ichsan Gorontalo;
3. Pak Jorry Karim, S.Kom, M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
4. Bapak Sudirman Melangi, M.Kom, selaku Wakil Dekan I Bidang Akademik Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
5. Ibu Irma Surya Kumala Idris, M.Kom, selaku Wakil Dekan II Bidang Administrasi Umum dan Keuangan Fakultas Ilmu Komputer Universitas Ichsan
6. Bapak Sudirman S. Panna, M.Kom, selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
7. Ibu Rezqiwati Ishak, M. Kom, selaku Pembimbing I yang telah banyak membimbing penulis selama ini;
8. Bapak Yusrianto Malago, M. Kom, selaku Pembimbing II yang telah banyak membimbing penulis selama ini;

9. Bapak dan Ibu Dosen Universitas Ichsan Gorontalo yang sudah mendidik dan mengajarkan berbagai disiplin ilmu kepada penulis;
10. Kedua Orang Tua saya yang tercinta, atas segala kasih sayang, jerih payah dan doa restunya dalam membesarkan dan mendidik penulis;
11. Rekan-rekan seperjuangan yang telah banyak memberikan bantuan dan moril yang sangat besar kepada penulis;
12. Kepada semua pihak yang ikut membantu dalam penyelesaian proposal ini yang tak sempat penulis sebutkan satu-persatu.

Semoga Allah SWT, melimpahkan balasan atas jasa-jasa mereka kepada kami. Penulis menyadari sepenuhnya bahwa dalam penyusunan ini masih jauh dari kesempurnaan dan masih banyak kekurangan. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran yang konstruktif. Akhirnya penulis berharap semoga hasil yang telah dicapai ini dapat bermanfaat bagi kita semua. Aamiinn.

Gorontalo, Juni 2022

Penulis

DAFTAR ISI

LEMBAR PERSETUJUAN SKRIPSI	ii
LEMBAR PENGESAHAN SKRIPSI	iii
LEMBAR PERNYATAAN SKRIPSI	iv
ABSTRACT	v
ABSTRAK	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.5.1 Manfaat Teoritis.....	4
1.5.2 Manfaat Praktis	5
BAB II LANDASAN TEORI	6
2.1 Tinjauan Studi	6
2.2 Tinjauan Pustaka	8
2.2.1 Kasus.....	8
2.2.2 Data Mining	8
2.2.3 Fungsi Data Mining	9
2.2.4 Clustering.....	10
2.2.5 Metode K-Means	10
2.2.6 Penerapan Metode K-Means.....	11
2.2.7 Perangkat Lunak	15
2.3 Kerangka Pikir.....	16

BAB III METODE PENELITIAN	17
3.1 Jenis, Metode, Subjek, Objek, Waktu dan Lokasi Penelitian.....	17
3.2 Pengumpulan Data	17
3.3 Tahapan <i>Clustering</i>	19
BAB IV HASIL PENELITIAN.....	21
4.1 Hasil Pengumpulan Data	21
4.2 Proses Pengolahan Data	22
4.2.1 Tahap Selection.....	22
4.2.2 Tahap Preprocessing	22
4.2.3 Tahap Transformasi	24
4.3 Hasil Implentasi K-Means Dengan Python	26
BAB V PEMBAHASAN PENELITIAN	32
5.1 Kinerja Model.....	32
5.2 Pembahasan Model.....	33
5.2.1 Pemodelan Metode K-Means.....	33
5.2.2 Visualisasi Hasil Pemodelan K-Means.....	34
5.2.3 Pusat Cluster dan Jumlah Data	35
5.2.4 Pelabelan Data Hasil Cluster	37
BAB VI KESIMPULAN DAN SARAN.....	40
6.1 Kesimpulan.....	40
6.2 Saran	40

DAFTAR GAMBAR

Gambar 2. 1 Proses <i>Konwledge Discoveryin Data Database</i> (KDD) [1]	8
Gambar 2. 2 Kerangka Pikir	17
Gambar 4. 1 Dataset	28
Gambar 4. 2 Hasil Konversi Dataset	29
Gambar 4. 3 Jumlah Data	30
Gambar 4. 4 Urutan Dataset	30
Gambar 4. 5 Hasil Konversi Data.....	31
Gambar 4. 6 Penentuan Kolom Dataset	32
Gambar 5.1 Visualisasi Teknik <i>Elbow</i>	34
Gambar 5.2 Visualisasi Hasil Clustering.....	36

DAFTAR TABEL

Tabel 1. 1 Data Laporan Kasus Polsek Kota Gorontalo.....	2
Tabel 2. 1 Penelitian Tentang Clustering	6
Tabel 2. 2 Data Siswa yang Mengajukan Beasiswa	11
Tabel 2. 3 Jarak Rata-Rata (E) Ke C1 dan C2.....	12
Tabel 2. 4 Jarak Rata-Rata (E) ke C1 Baru dan C2 Baru	13
Tabel 2. 5 Kelompok Siswa Berprestasi	14
Tabel 2. 6 Contoh <i>Confusion Matrix</i> [15]	14
Tabel 2. 7 Perangkat Lunak Pendukung.....	16
Tabel 3. 1 Variabel Data.....	19
Tabel 3. 2 Konversi Nilai Atribut.....	20
Tabel 4. 1 Data Set Kepolisian	22
Tabel 4. 2 Hasil Rekapitulasi Kasus.....	23
Tabel 4. 3 Hasil Rekapitulasi Kasus Akhir	24
Tabel 4. 4 Konversi Nilai Atribut Tempat Kejadian	25
Tabel 4. 5 Konversi Nilai Atribut Jenis Kasus	26
Tabel 4. 6 Hasil Dataset.....	26
Tabel 5. 1 Hasil Output	37
Tabel 5. 2 Hasil Pelabelan Cluster	39

DAFTAR LAMPIRAN

Data Operasional

Surat Keterangan Telah Melakukan Penelitian

Potongan Kode Program

Biodata

Surat Keterangan Bebas Pustaka

Hasil Turnitin

BAB I

PENDAHULUAN

1.1 Latar Belakang

Negara Indonesia adalah negara hukum sebagaimana yang disebutkan dalam Pasal 1 ayat (3) Undang-Undang Dasar Negara Republik Indonesia Tahun 1945 (UUD 1945). Hukum bertujuan untuk menjamin adanya kepastian hukum dalam masyarakat dan hukum itu harus pula bersendikan pada nilai keadilan yaitu asas-asas keadilan pada masyarakat itu.

Kasus kriminal di Indonesia sedang marak terjadi, dapat dengan mudah ditemukannya berita mengenai kriminalitas di berbagai media dengan beragam jenis kasus yang berbeda. Indonesia mencatat jumlah kejahatan pada tahun 2019 menurun 53.360 kasus atau 19,3% bila dibanding tahun 2018. Pada 2018, jumlah kejahatan dilaporkan berjumlah 275.903 kasus, sedangkan pada 2019 jumlah kejahatan dilaporkan 222.543 kasus.

Kejahatan paling dominan yang dilaporkan sepanjang 2019 adalah kejahatan konvensional yaitu 202.292 kasus atau 90,9 persen dari seluruh kejahatan, diikuti kejahatan transnasional yaitu 36.219 kasus atau 16,2 persen. Dari total 222.543 kasus kejahatan dilaporkan pada 2019, sebanyak 154.963 kasus berhasil diselesaikan perkaranya atau menurun 14,4% dibandingkan 2018 [2].

Tabel 1. 1 Data Laporan Kasus Polsek Kota Gorontalo

Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
1 jan 2021	Dumbo Raya	Kekerasan	2
	Kota Selatan	pencurian	1
2 jan 2021	Hulothalangi	Kekerasan	1
3 jan 2021	Kota Timur	Kekerasan	1
4 jan 2021	Dungingi	Pembunuhan	1
5 jan 2021	Kota Barat	Kekerasan	4
6 jan 2021	Kota Tengah	Kekerasan	2
	Sipatana	Kekerasan	2
8 jan 2021	Kota Selatan	Kekerasan	1
9 jan 2021	Kota Barat	Pemeriksaan	1
10 jan 2021	Sipatana	Pencurian	2
11 jan 2021	Dumbo Raya	Pencurian	2
12 jan 2021	Dumbo Raya	Kekerasan	1
14 jan 2021	Sipatana	Kekerasan	3
15 jan 2021	Sipatana	Penculikan	1
...
31 des 2021	Kota Tengah	Penculikan	1
	Kota Timur	Pembunuhan	1
	Sipatana	Kekerasan	2
	Kota Barat	Kekerasan	1

(Sumber: Polsek Kota Gorontalo 2021)

Dengan banyaknya kasus kejahatan di Kota Gorontalo, perlu adanya sebuah sistem komputerisasi pengelompokan daerah tindak kejahatan sebagai salah satu

usaha untuk membantu pihak kepolisian dalam mengambil keputusan apakah suatu daerah memerlukan pengawasan ekstra atau tidak, selain itu informasi tersebut dibutuhkan untuk mengetahui intensitas tindak kejahatan. Pengelompokan tindak kejahatan dapat dilakukan dengan menggunakan teknik data mining berupa analisis kelompok. Salah satu metode analisis kelompok yang dapat digunakan adalah metode non hierarki, yaitu K-Means Clustering.

Salah satu manfaat Clustering merupakan metode segmentasi data yang sangat berguna dalam prediksi dan analisa masalah bisnis tertentu.

Metode K-Means merupakan salah satu algoritma dalam data mining yang bisa digunakan untuk melakukan pengelompokan/clustering suatu data. Ada banyak pendekatan untuk membuat cluster, diantaranya adalah membuat aturan yang mendikte keanggotaan dalam group yang sama berdasarkan tingkat persamaan diantara anggota-anggotanya [3].

Adapula metode ini digunakan untuk Clustering status gizi balita, penelitian ini dilakukan oleh Windha Mega Pradnya Duhita menggunakan metode K-Means. Data yang digunakan yaitu data tinggi dan berat badan balita. Dengan membandingkan hasil pengelompokan menggunakan table Growth Chart dan algoritma K-Means didapat 17 data yang memiliki kelompok yang sama. Dari angka ini dapat disimpulkan bahwa algoritma K-Means hanya memiliki nilai akurasi 34% benar. Nilai ini bisa berubah seiring dengan penambahan data latih [3].

Metode K-Means juga digunakan pada pengelompokan mahasiswa berdasarkan nilai akademik, penelitian ini menerapkan metode K-Means untuk mengelompokkan mahasiswa. Dari data yang dilatih, didapatkan 4 kelompok dengan hasil sebagai berikut: 1. Mahasiswa dengan $IPK = 0.5167$ untuk cluster 0, sebanyak 9 Mahasiswa dari 124 Mahasiswa (7%) 2. Mahasiswa dengan $IPK = 3.4143$ untuk cluster 1, sebanyak 28 Mahasiswa dari 124 Mahasiswa (23%) 3. Mahasiswa dengan $IPK = 3.3092$ untuk cluster 2, sebanyak 40 Mahasiswa dari 124 Mahasiswa (32%) 4. Mahasiswa dengan $IPK = 3.8991$ untuk cluster 3, sebanyak 47 Mahasiswa dari 124 Mahasiswa (38%) [4].

Berdasarkan penelitian diatas, maka peneliti bermaksud untuk menulis penelitian yang berjudul “*Clustering Kasus Kepolisian Dengan Metode K-Means*”,

1.2 Identifikasi Masalah

Dengan banyaknya kasus kejahatan di Kota Gorontalo, perlu adanya sebuah sistem komputerisasi pengelompokan daerah tindak kejahatan sebagai salah satu usaha untuk membantu pihak kepolisian dalam mengambil keputusan apakah suatu daerah memerlukan pengawasan ekstra atau tidak, selain itu informasi tersebut dibutuhkan untuk mengetahui intensitas tindak kejahatan.

1.3 Rumusan Masalah

Berdasarkan identifikasi masalah diatas, maka permasalahannya dirumuskan sebagai berikut:

1. Bagaimana hasil dari penerapan metode K-Means untuk pengelompokan kasus kepolisian di Kota Gorontalo?
2. Bagaimana hasil analisis setiap clustering kasus kepolisian sehingga bisa dijadikan keputusan penentuan tempat kasus terbanyak?

1.4 Tujuan Penelitian

Berdasarkan Rumusan permasalahannya diatas, maka tujuan dari penelitian ini adalah :

1. Untuk mengetahui hasil penerapan metode K-Means untuk pengelompokan kasus kepolisian.
2. Untuk mengetahui hasil analisis setiap clustering kasus kepolisian sehingga bisa dijadikan keputusan penentuan tempat kasus terbanyak.

1.5 Manfaat Penelitian

Penelitian ini diharapkan mempunyai manfaat yaitu :

1.5.1 Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan sumbangan pengembangan ilmu bidang kajian data mining tentang kemampuan metode K-Means dalam melakukan teknik *clustering*

1.5.2 Manfaat Praktis

Hasil dari penelitian ini diharapkan dapat digunakan sebagai *alternative* atau masukan kepada semua pihak yang berkepentingan khususnya dalam mengelompokkan kasus kepolisian di Kota Gorontalo.

BAB II

LANDASAN TEORI

2.1 Tinjauan Studi

Adapun penelitian terdahulu yang berhubungan dengan tema peneliti yang dijadikan sebagai bahan referensi dalam menentukan metode yang akan digunakan.

Tabel 2. 1 Penelitian Tentang Clustering

No	Peneliti	Judul Penelitian	Deskripsi Singkat
1.	Wakhid Afifi, Dhiya'an Ramadhanty Nastiti, Qurrotul Aini, 2020 [5].	Clustering K-Means Pada Data Ekspor	Dari permasalahan pengelompokan negara untuk tempat ekspor mobil dapat diselesaikan dengan menggunakan metode clustering k-means. Walaupun mampu menyelesaikan data dengan jumlah banyak, namun masih belum efisien dalam mengelompokkan dokumen secara tepat. Hasil akhir juga sangat bergantung pada centroid yang ditentukan, karena jika berbeda centroid yang ditentukan maka hasilnya pun akan berbeda juga.
2.	Sugiono, Siti Nurdiani, Safitri Linawati, Rizky Ade Safitri, Elin Panca Saputra, 2019 [6].	Pengelompokan Perilaku Mahasiswa Pada Perkuliahan E-Learning dengan K-Means Clustering	Berdasarkan proses Clustering K-Means dengan Rapid Miner, diperoleh sejumlah perilaku mahasiswa dalam mengikuti perkuliahan berbasis E-learning antara lain; 1. kelompok mahasiswa dengan jumlah aktifitas yang

No.	Peneliti	Judul Penelitian	Deskripsi Singkat
			2. Kelompok mahasiswa dengan jumlah aktifitas yang sedang dan mendapat nilai perkuliahan yang tinggi diwakili dengan titik biru. 3. Kelompok mahasiswa dengan jumlah aktifitas sedikit dan mendapatkan nilai yang rendah diwakili dengan titik merah.
3.	Wulan Fajariyanti, 2017 [7]	Clustering Wilayah Pada Pelanggaran Berkendaraan Menggunakan Metode K-Means	Pengelompokan wilayah pada pelanggaran dalam berkendara menggunakan metode K-Means yang dimulai dari pengambilan centroid secara random. Jumlah kelompok cluster ada 3. Tiap cluster akan dihitung jarak terdekat/ kemiripannya dengan centroid yang telah diambil secara random sesuai dengan kelompoknya. Setelah itu akan dihasilkan hasil klasterisasi kelompok pengayaan berdasarkan kelompoknya atau cluster yang telah ditempati.
			banyak dan mendapatkan nilai yang tinggi diwakili dengan titik hijau.

2.2 Tinjauan Pustaka

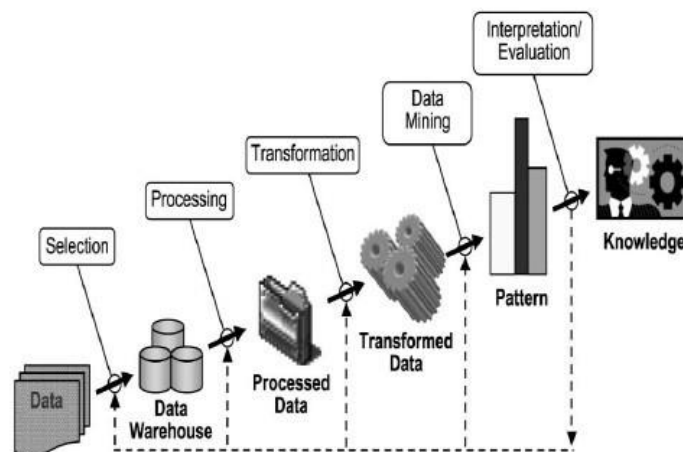
2.2.1 Kasus

Keadaan yang sebenarnya dari suatu urusan atau perkara, keadaan, atau kondisi khusus yang berhubungan dengan seseorang atau suatu hal, soal, perkara [8].

2.2.2 Data Mining

Data mining adalah perkembangan siklus mengamati contoh koneksi yang bertujuan untuk memisahkan informasi yang sangat besar untuk memukul paku di kepala, baru, berharga, dan mendapatkan pemahaman tentang contoh atau contoh dari informasi tersebut. (Han & Kamber, 2006) [10].

Gambar 2. 1 Proses *Knowledge Discovery in Data Database* (KDD) [1].



Tahapan yang dilakukan pada proses data mining diawali dari seleksi data dari data sumber ke data target, tahap *preprocessing* untuk memperbaiki kualitas data, transformasi, data mining serta tahap interpretasi dan evaluasi yang menghasilkan output berupa pengetahuan baru yang diharapkan memberikan kontribusi yang lebih baik. Secara detail dijelaskan sebagai berikut (Fayyad, 1996) [1]:

1. *Data selection*. Pilihan (seleksi) informasi dari sekumpulan informasi fungsional harus dilakukan sebelum fase mengekstrak data di KDD dimulai. Informasi yang dipilih yang digunakan untuk proses penambangan informasi disimpan dalam catatan, terpisah dari kumpulan data fungsional.

2. *Pre-processing/cleaning*. Sebelum siklus penambangan informasi dapat dilakukan, penting untuk melakukan interaksi pembersihan pada informasi yang menjadi titik fokus KDD. Sistem pembersihan mencakup, antara lain, menghilangkan informasi salinan, memeriksa informasi yang bertentangan, dan memperbaiki kesalahan dalam informasi.

2.2.3 Fungsi Data Mining

Penambangan data sebenarnya mempunyai banyak sekali manfaat atau fungsi. Namun, secara umum, terdapat 2 fungsi data mining yaitu: [11]

1. Deskriptif: Kapasitas yang berbeda mengacu pada kapasitas dalam memahami informasi yang terkonsentrasi lebih lanjut. Melalui siklus ini, adalah normal bahwa perilaku informasi dapat diketahui. Informasi ini nantinya dapat dimanfaatkan untuk mengungkap kualitas informasi yang ideal. Dengan menggunakan kapasitas ilustratif dari penggalian informasi, contoh tertentu yang awalnya tersembunyi dalam suatu informasi dapat ditemukan. Dengan demikian, dengan asumsi disadari bahwa ada contoh yang suram dan memiliki harga diri, itu menyiratkan bahwa atribut informasi dapat diketahui.
2. Prediktif: Fungsi Prediktif dari data mining adalah elemen interaksi yang pada akhirnya akan mengungkap contoh informasi tertentu. Contoh ini harus terlihat dari beberapa faktor yang terkandung dalam informasi. Setiap kali sebuah contoh ditemukan, contoh tersebut dapat digunakan untuk mengukur berbagai faktor yang nilai atau jenisnya pada saat ini belum jelas. Oleh karena itu, fungsi prediktif dipandang sebagai fungsi prediktif yang sebanding dengan analisis prediktif. Kemampuan ini juga dapat digunakan untuk menilai variabel luar biasa yang tidak ada dalam informasi. Selanjutnya, kapasitas ini dapat bekerja dengan dan menguntungkan setiap individu yang membutuhkan perkiraan yang tepat untuk meningkatkan hal-hal yang signifikan.

2.2.4 Clustering

Clustering disebut pengelompokan sejumlah data atau objek kedalam cluster (group) sehingga setiap *cluster* akan berisi data yang semirip mungkin dan berbeda dengan objek dalam *cluster* yang lainnya (Putri, Kom, & Kom, 2015). Ada dua metode *clustering* yang kita kenal, yaitu *Hierarchy* dan *Non Hierarchy* (Putri et al., 2015). Salah satu teknik pengelompokan dalam *data mining* adalah metode *clustering*. Pengertian *clustering* keilmuan dalam *data mining* adalah pengelompokan sejumlah data atau objek ke dalam *cluster (group)* sehingga setiap dalam *cluster* tersebut akan berisi data yang semirip mungkin dan berbeda dengan objek dalam *cluster* yang lainnya. Metode *clustering* yang mempunyai sifat efisien dan cepat yang dapat digunakan salah satunya adalah metode k-means, metode ini bertujuan untuk membuat *cluster* objek berdasarkan atribut menjadi k partisi. *Hierarchical clustering* adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan kedua [6].

2.2.5 Metode K-Means

K-Means adalah salah satu metode data clustering non hierarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster atau kelompok agar data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama sedangkan data yang mempunyai karakteristik berbeda dikelompokkan ke dalam kelompok yang lainnya (Benri dan Herlina, 2015) [12].

Langkah-langkah dalam algoritma K-means clustering adalah: [13]

- 1) Menentukan jumlah cluster
- 2) Menentukan nilai centroid. Dalam menentukan nilai centroid untuk awal iterasi dilakukan secara acak. Sedangkan untuk menentukan nilai centroid yang merupakan tahap dari iterasi, digunakan rumus sebagai berikut

$$\bar{V}_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} , \dots\dots\dots (2.1)$$

Dimana :

V_{ij} merupakan centroid/ rata-rata cluster ke-I untuk variable ke-j

N_i merupakan jumlah data yang menjadi anggota cluster ke-i i,

k adalah indeks dari cluster

j adalah indeks dari variabel

X_{kj} adalah nilai data ke-k di dalam cluster tersebut untuk variable ke-j

- 3) Menghitung jarak antara titik centroid dengan titik tiap objek Untuk menghitung jarak tersebut dapat menggunakan Euclidean Distance, yaitu

$$D_e = \sqrt{(X_i - S_i)^2 + (Y_i - t_i)^2}, \dots \dots \dots (2.2)$$

Dimana : D_e adalah *Euclidean Distance*

2.2.6 Penerapan Metode K-Means

Berikut Penerapan *K-Means* dalam penyelesaian berdasarkan rumus yang akan digunakan untuk Pengelompokkan Siswa Berprestasi berdasarkan studi yang pernah dilakukan oleh (Rini, Eka dan Sitty :2020)[14]

Tabel 2. 2 Data Siswa yang Mengajukan Beasiswa

No	NIS	Nilai Tugas (A)	Nilai UH (B)	Nilai UTS (C)	Nilai UAS (D)	Rata-rata (E)
1	1711	82	91	70	85	82
2	1715	89	68	89	60	76.5
3	1726	90	85	95	98	92
4	1747	87	80	87	95	87.25
5	1751	65	75	65	65	67.5
6	1760	70	75	61	54	65
7	1767	77	76	88	85	81.5
8	1768	84	82	65	55	71.5
9	1769	86	87	95	93	90.25
10	1771	78	78	71	70	74.25
11	1773	84	88	80	85	84.25

No	NIS	Nilai Tugas (A)	Nilai UH (B)	Nilai UTS (C)	Nilai UAS (D)	Rata- rata (E)
12	1774	78	77	65	85	76.25
13	1775	77	87	80	78	80.5
14	1780	96	98	78	76	87
15	1791	84	81	76	98	84.75
16	1802	82	98	70	95	86.25
17	1804	80	88	63	55	71.5
18	1805	83	71	88	65	76.75
19	1809	87	68	65	84	76
20	1811	86	76	68	76	76.5

Tabel 2. 3 Jarak Rata-Rata (E) Ke C1 dan C2

	C1	C2		C1	C2
E1	3	4	E11	5.25	6.25
E2	2.5	1.5	E12	2.75	1.75
E3	13	14	E13	1.5	2.5
E4	8.25	9.25	E14	8	9
E5	11.5	10.5	E15	5.75	6.75
E6	14	13	E16	7.25	8.25
E7	2.5	3.5	E17	7.5	6.5
E8	7.5	6.5	E18	2.25	1.25
E9	11.25	12.25	E19	3	2
E10	4.75	3.75	E20	2.5	1.5

Tahapan selanjutnya yaitu kelompokkan nilai Ratarata (E) berdasarkan jarak terpendek ke C1 dan C2, sehingga diperoleh kelompok nilai Rata-rata (E) sebagai berikut :

Anggota C1 : E1, E3, E4, E7, E9, E11, E13, E14, E15, E16

Anggota C2 : E2, E5, E6, E8, E10, E12, E17, E18, E19, E20

Kemudian menentukan nilai centroid baru (C1baru dan C2baru) sebagai berikut:

$$C1 \text{ baru} = (3+13+8.25+2.5+11.25+5.25+1.5+8+5.75+7.25) / 10$$

$$C1 \text{ baru} = 85,575$$

$$C2 \text{ baru} = (1.5+10.5+13+6.5+3.75+1.75+6.5+1.25+2+1.5) / 10$$

$$C2 \text{ baru} = 73,175$$

Setelah mendapatkan nilai centroid baru, bandingkan nilai centroid baru dengan centroid lama. Jika centroid baru tidak sama dengan centroid lama, maka perhitungan jarak diulang menggunakan centroid baru. Hasil perhitungan dapat dilihat pada table

Tabel 2. 4 Jarak Rata-Rata (E) ke C1 Baru dan C2 Baru

	C1	C2		C1	C2
E1	3,575	8,825	E11	1,325	11,075
E2	9,075	3,325	E12	9,325	3,075
E3	6,425	18,825	E13	5,075	7,325
E4	1,675	14,075	E14	1,425	13,825
E5	18,075	5,675	E15	0,825	11,575
E6	20,575	8,175	E16	0,675	13,075
E7	4,075	8,325	E17	14,075	1,675
E8	14,075	1,675	E18	8,825	3,575
E9	4,675	17,075	E19	9,575	2,825
E10	11,325	1,075	E20	9,075	3,325

Kelompokkan kembali nilai Rata-rata (E) berdasarkan jarak terpendek ke C1baru dan C2baru, maka diperoleh kelompok nilai Rata-rata (E) sebagai berikut :

Anggota C1 : E1, E3, E4, E7, E9, E11, E13, E14, E15, E16

Anggota C2 : E2, E5, E6, E8, E10, E12, E17, E18, E19, E20

Karena tidak terjadi perubahan anggota klaster, maka proses K-Means selesai. Dari hasil perhitungan K-Means, terdapat 10 orang siswa yang masuk kategori siswa berprestasi seperti pada table 2. 4

Tabel 2. 5 Kelompok Siswa Berprestasi

No.	NIS	Nilai Tugas (A)	Nilai UH (B)	Nilai UTS (C)	Nilai UAS (D)	Rata-rata (E)
1	1711	82	91	70	85	82
2	1726	90	85	95	98	92
3	1747	87	80	87	95	87.25
4	1767	77	76	88	85	81.5
5	1769	86	87	95	93	90.25
6	1773	84	88	80	85	84.25
7	1775	77	87	80	78	80.5
8	1780	96	98	78	76	87
9	1791	84	81	76	98	84.75
10	1802	82	98	70	95	86.25

2.2.7 Perangkat Lunak

Tabel 2. 7 Perangkat Lunak Pendukung

No	Perangkat Lunak Pendukung	Berfungsi
1	<i>Python</i>	Bahasa pemrograman tingkat tinggi, Penulisan kode/sintaks lebih sederhana, Bersifat <i>open-source</i> , <i>cross-platform</i> dan cocok digunakan untuk Data Scientist serta didukung banyak library seperti : <i>Numpy</i> , <i>SciPy</i> , <i>Pandas</i> , <i>Scikit-Learn</i> , dan <i>Matplotlib</i>
2	<i>Google Colaboratory</i> (Google Colab)	<i>Integrated Development Environments</i> (IDE) berbasis Web yang akan digunakan untuk menuliskan coding Bahasa <i>Python</i>

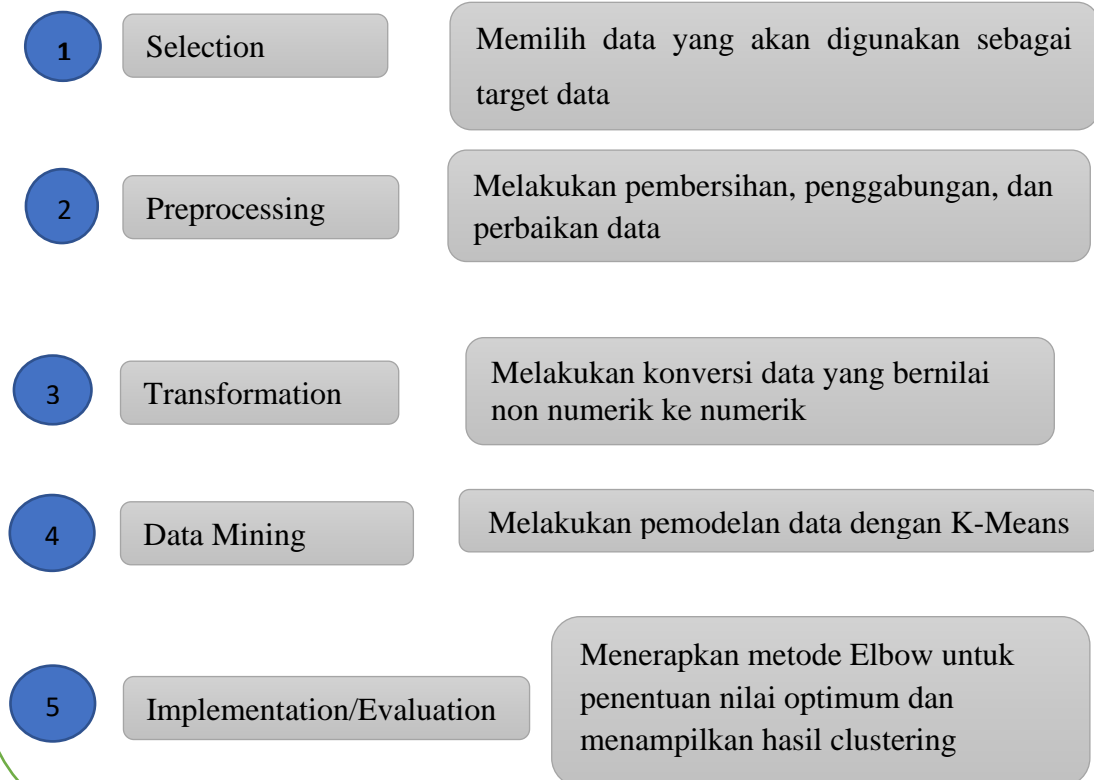
2.3 Kerangka Pikir

Masalah

1. Bagaimana hasil penerapan metode K-Means, untuk pengelompokan kasus kepolisian?
2. Bagaimana hasil analisis setiap clustering kasus kepolisian sehingga bisa dijadikan keputusan penentuan tempat kasus terbanyak?



Proses Clustering



Tujuan

1. Mengetahui hasil kinerja metode K-Means untuk pengelompokan kasus kepolisian.
2. Mengetahui hasil analisis setiap clustering kasus kepolisian sehingga bisa dijadikan keputusan penentuan tempat dengan kasus terbanyak.

BAB III

METODE PENELITIAN

3.1 Jenis, Metode, Subjek, Objek, Waktu dan Lokasi Penelitian

Penelitian ini menggunakan metode penelitian studi kasus Polsek Kota Gorontalo. Dengan demikian jenis penelitian ini adalah kuantitatif. Subjek penelitian ini adalah Clustering Kasus Kepolisian dengan Metode K-Means. Penelitian ini dimulai dari Oktober 2021 – Januari 2022 mencakup pengumpulan dan pengolahan data beserta bimbingan. Tempat dan lokasi penelitian ini berada di Polsek Kota Gorontalo.

3.2 Pengumpulan Data

1. Data Sekunder

Data Sekunder adalah data yang diambil dari studi literatur yang dilakukan. Sumber dari studi literatur yaitu jurnal. Makalah ilmiah atau buku yang membahas tentang Metode K-Means.

2. Data Primer

Untuk memperoleh data primer yang merupakan data lapangan atau langsung pada objek penelitian yang bertempat di Polsek Kota Gorontalo.

Teknik yang digunakan :

- a. Observasi adalah meninjau atau mengamati secara langsung. Adapun pada penelitian ini dilakukan dengan cara mengumpulkan data jumlah kasus di Polsek Kota Gorontalo.
- b. Wawancara, dengan cara mengajukan pertanyaan kepada bagian yang terkait di Polsek Kota Gorontalo untuk proses Clustering kasus kepolisian.

Tabel 3. 1 Variabel Data

No	Nama	Type	Value	Keterangan
1	Tanggal Laporan	<i>Date</i>	1 Januari – 31 Desember	<i>Parameter Input</i>
2	Tempat Kejadian	Nominal	<ul style="list-style-type: none"> - Dumbo Raya - Kota Selatan - Kota Timur - Kota Barat - Kota Tengah - Kota Utara - Duingi - Hulonthalangi - Sipatana 	<i>Parameter Input</i>
3	Jenis Kasus	Nominal	<ul style="list-style-type: none"> - Kekerasan - Pencurian - Pembunuhan - Pemerkosaan - Penculikan 	<i>Parameter Input</i>
4	Banyak Kasus	Numerik	1-100	<i>Parameter Input</i>

3.3 Tahapan *Clustering*

Pada tahap ini dilakukan serangkaian proses sesuai dengan kerangka pikir sebagai berikut :

1. Tahap *Selection* : pada tahapan ini akan dilakukan seleksi data terhadap dataset yang dikumpulkan dari data jumlah kasus kepolisian kota gorontalo 1 tahun terakhir untuk dijadikan sebagai target data (dataset)
2. Tahap *Preprocessing* : pada tahapan ini dilakukan beberapa proses yaitu penghapusan data yang duplikat, pengecekan data yang tidak konsisten, melengkapi data yang kosong dan menghapus atribut yang tidak diperlukan dalam pemodelan.
3. Tahap *Transformation* : pada tahapan ini dilakukan perubahan atau konversi data yaitu atribut yang bernilai data nominal akan dikonversi ke nilai rasio agar dapat digunakan dalam pemodelan *K-Means*. Contoh untuk atribut Jenis Kasus.

Tabel 3. 2 Konversi Nilai Atribut

Jenis Kasus	Nilai Konversi
Pembunuhan	1
Penculikan	2
Pemeriksaan	3
Pencurian	4
Kekerasan	5

Semua atribut yang bertipe nominal akan dilakukan konversi nilai seperti pada table di atas.

- 4 Tahap *Data Mining* : pada tahapan ini dilakukan pengembangan model yaitu penggunaan metode *K-Means* untuk mendapatkan hasil *cluster*.
- 5 Tahap *Interpretation/Evaluation* : pada tahapan ini menentukan berapa jumlah *cluster* yang optimal dengan metode *Elbow*. Hasil yang didapat dari nilai K optimum pada metode *Elbow* tersebut diterapkan ke dalam metode *K-Means* kembali. Hasilnya divisualisasikan jumlah kluster beserta anggota klusternya,

sehingga berdasarkan visualisasi tersebut dengan mudah dapat dipahami arti dari setiap kelompok *cluster* tersebut yang selanjutnya dijadikan sebagai pengetahuan atau keputusan untuk *clustering* kasus kepolisian.

BAB IV

HASIL PENELITIAN

4.1 Hasil Pengumpulan Data

Tabel 4. 1 Data Set Kepolisian

No	Periode	Tempat Kejadian	Jenis Kasus
1	Jan 2021	Dumbo Raya	Kekerasan
2	Jan 2021	Kota Selatan	Pencurian
3	Jan 2021	Hulonthalangi	Kekerasan
4	Jan 2021	Kota Timur	Kekerasan
5	Jan 2021	Dungingi	Pembunuhan
6	Jan 2021	Kota Barat	Kekerasan
7	Jan 2021	Kota Tengah	Kekerasan
8	Jan 2021	Sipatana	Kekerasan
9	Jan 2021	Kota Selatan	Kekerasan
10	Jan 2021	Kota Barat	Pemeriksaan
11	Jan 2021	Sipatana	Pencurian
12	Jan 2021	Dumbo Raya	Pencurian
13	Jan 2021	Dumbo Raya	Kekerasan
14	Jan 2021	Sipatana	Kekerasan
15	Jan 2021	Sipatana	Penculikan
16	Jan 2021	Dungingi	Pembunuhan
17	Jan 2021	Dungingi	Kekerasan
18	Jan 2021	Kota Selatan	Kekerasan
19	Jan 2021	Kota Barat	Kekerasan

No	Periode	Tempat Kejadian	Jenis Kasus
...
338	Des 2021	Kota Barat	Kekerasan

4.2 Proses Pengolahan Data

4.2.1 Tahap Selection

Pada tahap ini dilakukan seleksi data terhadap data set yang sudah dikumpulkan dan hanya mengambil data untuk tahun 2021 dari bulan Januari sampai dengan bulan Desember dengan jumlah data set sebanyak 338.

4.2.2 Tahap Preprocessing

Pada tahap ini dilakukan serangkaian proses diantaranya :

a. Melakukan Rekapitulasi Kasus

Pada tahap ini dilakukan rekapitulasi kasus berdasarkan bulan, tempat kejadian dan jenis kasus sehingga didapatkan hasil sebagai berikut :

Tabel 4. 2 Hasil Rekapitulasi Kasus

No	Bulan	Tempat Kejadian	Jenis Kasus	Jumlah
1	Jan 2021	Dumbo Raya	Kekerasan	1
2	Jan 2021	Kota Selatan	Kekerasan	1
3	Jan 2021	Kota Utara	Kekerasan	1
4	Jan 2021	Kota Timur	Kekerasan	1
5	Jan 2021	Dungingi	Kekerasan	1
6	Jan 2021	Kota Barat	Kekerasan	1
7	Jan 2021	Kota Tengah	Kekerasan	1
8	Jan 2021	Sipatana	Kekerasan	1
9	Jan 2021	Hulonthalangi	Kekerasan	1
10	Jan 2021	Dumbo Raya	Pencurian	1
11	Jan 2021	Kota Selatan	Pencurian	1

No	Bulan	Tempat Kejadian	Jenis Kasus	Jumlah
12	Jan 2021	Kota Utara	Pencurian	0
13	Jan 2021	Kota Timur	Pencurian	1
14	Jan 2021	Dungingi	Pencurian	0
15	Jan 2021	Kota Barat	Pencurian	0
16	Jan 2021	Kota Tengah	Pencurian	1
17	Jan 2021	Sipatana	Pencurian	1
18	Jan 2021	Hulonthalangi	Pencurian	1
19	Jan 2021	Dumbo Raya	Pembunuhan	0

711	Des 2021	Hulonthalangi	Pemeriksaan	2

b. Menghapus Rekapitulasi Data Kasus Yang Bernilai 0

Data kasus yang bernilai 0 (nol) dihapus karena tidak memiliki bobot maka data dianggap tiada.

Tabel 4. 3 Hasil Rekapitulasi Kasus Akhir

No	Bulan	Tempat Kejadian	Jenis Kasus	Jumlah
1	Jan 2021	Dumbo Raya	Kekerasan	2
2	Jan 2021	Kota Selatan	Kekerasan	2
3	Jan 2021	Kota Utara	Kekerasan	1
4	Jan 2021	Kota Timur	Kekerasan	2
5	Jan 2021	Dungingi	Kekerasan	1
6	Jan 2021	Kota Barat	Kekerasan	2
7	Jan 2021	Kota Tengah	Kekerasan	2
8	Jan 2021	Sipatana	Kekerasan	4
9	Jan 2021	Hulonthalangi	Kekerasan	2

No	Bulan	Tempat Kejadian	Jenis Kasus	Jumlah
10	Jan 2021	Dumbo Raya	Pencurian	1
11	Jan 2021	Kota Selatan	Pencurian	1
12	Jan 2021	Kota Timur	Pencurian	1
13	Jan 2021	Kota Tengah	Pencurian	1
14	Jan 2021	Sipatana	Pencurian	1
15	Jan 2021	Hulonthalangi	Pencurian	1
16	Jan 2021	Dungingi	Pembunuhan	2
17	Jan 2021	Sipatana	Penculikan	1
18	Jan 2021	Kota Barat	Pemeriksaan	1
19	Feb 2021	Dumbo Raya	Kekerasan	2
...
259	Des 2021	Hulonthalangi	Pemeriksaan	3

4.2.3 Tahap Transformasi

Pada tahap ini dilakukan tahap konversi untuk atribut nilai non merik ke numerik yaitu atribut tempat kejadian dan jenis kasus dengan hasil sebagai berikut:

Tabel 4. 4 Konversi Nilai Atribut Tempat Kejadian

No	Tempat Kejadian	Nilai
1	Kota Utara	1
2	Hulonthalangi	2
3	Kota Tengah	3
4	Kota Selatan	4
5	Dungingi	5
6	Kota Timur	6

No	Tempat Kejadian	Nilai
7	Dumbo Raya	7
8	Kota Barat	8
9	Sipatana	9

Tabel 4. 5 Konversi Nilai Atribut Jenis Kasus

No	Jenis Kasus	Nilai
1	Pembunuhan	1
2	Penculikan	2
3	Pemeriksaan	3
4	Pencurian	4
5	Kekerasan	5

Setelah dilakukan perbaikan data set maka didapatkan hasil akhir dataset sebagai berikut :

Tabel 4. 6 Hasil Dataset

No	Bulan	Tempat Kejadian	Jenis Kasus	Jumlah
1	Jan 2021	7	5	2
2	Jan 2021	4	5	2
3	Jan 2021	1	5	1
4	Jan 2021	6	5	2
5	Jan 2021	5	5	1
6	Jan 2021	8	5	2
7	Jan 2021	3	5	2
8	Jan 2021	9	5	4

No	Bulan	Tempat Kejadian	Jenis Kasus	Jumlah
9	Jan 2021	2	5	2
10	Jan 2021	7	4	1
11	Jan 2021	4	4	1
12	Jan 2021	6	4	1
13	Jan 2021	3	4	1
14	Jan 2021	9	4	1
15	Jan 2021	2	4	1
16	Jan 2021	5	1	2
17	Jan 2021	9	2	1
18	Jan 2021	8	3	1
19	Feb 2021	7	5	2
...
259	Des 2021	2	3	3

yang akan digunakan maka tahap selanjutnya akan dilakukan pemodelan klustering dengan menggunakan metode k-means.

4.3 Hasil Implentasi K-Means Dengan Python

Implementasi pemodelan k-means menggunakan Bahasa pemrograman python dengan memakai Tools Notebook Colaboratory dengan tahapan sebagai berikut :

1. Import Library Python

Berikut Library Python yang perlu di import agar proses pemodelan k-means dapat dijalankan :

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
import seaborn as sns
from mpl_toolkits.mplot3d import Axes3D
import plotly.graph_objs as go
from plotly import tools
from plotly.subplots import make_subplots
import plotly.offline as py
```

2. Pengaktifan Google Drive

Data set yang akan digunakan disimpan di google drive untuk itu diaktifkan dengan perintah sebagai berikut :

```
#menagktifkan google drive sebagai tempat penyimpanan data
yang digunakan
from google.colab import drive
drive.mount('/content/drive')
```

3. Membaca Dataset

Dataset yang sudah dibersihkan yang akan digunakan pada proses klustering yang berbentuk file excel harus di upload terlebih dahulu kedalam google drive, selanjutnya dataset dibaca dengan menggunakan perintah sebagai berikut :

```
#Baca Dataset
dfKasus = pd.read_excel('/content/drive/MyDrive/nancy/Data
setKasusPolisi.xlsx')
dfKasus.head()
```

Setelah dijalankan perintah diatas maka didapatkan hasil sebagai berikut :

	Bulan	Tempat Kejadian	Jenis Kasus	Jumlah
0	Jan 2021	Dumbo Raya	Kekerasan	2
1	Jan 2021	Kota Selatan	Kekerasan	2
2	Jan 2021	Kota Utara	Kekerasan	1
3	Jan 2021	Kota Timur	Kekerasan	2
4	Jan 2021	Dungingi	Kekerasan	1

Gambar 4. 1 Dataset

4. Konversi Dataset Excel Ke CSV

Dataset dalam format excel perlu dikonversi ke bentuk CSV (COMMA Separated Values) agar bisa digunakan dalam pemodelan data dengan menggunakan perintah sebagai berikut :

```
#Konversi dataset Excel ke CSV
dfKasus.to_csv("/content/drive/MyDrive/nancy/KasusPolisi.csv", index=False)
dfKasus.info()
```

Setelah dijalankan perintah diatas maka didapatkan hasil sebagai berikut :

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 259 entries, 0 to 258
Data columns (total 4 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Bulan            259 non-null   object
1   Tempat Kejadian  259 non-null   object
2   Jenis Kasus      259 non-null   object
3   Jumlah           259 non-null   int64
dtypes: int64(1), object(3)
memory usage: 8.2+ KB
```

Gambar 4. 2 Hasil Konversi Dataset

5. Menghitung Banyak Data

Sebelum melakukan konversi nilai atribut yang bernilai non numerik perlu di hitung banyaknya data untuk setiap nilai atribut. Berikut perintah untuk menghitung banyaknya data pada atribut tempat kejadian :

```
#menghitung banyaknya data berdasarkan tempat kejadian
dfKasus['Tempat Kejadian'].value_counts()
```

Setelah dijalankan perintah diatas maka didapatkan hasil sebagai berikut :

```
Sipatana      39
Kota Selatan  36
Dungingi     30
Dumbo Raya   28
Kota Barat    28
Kota Tengah   28
Kota Timur    27
Hulonthalangi 23
Kota Utara     20
Name: Tempat Kejadian, dtype: int64
```

Gambar 4. 3 Jumlah Data

Cara perhitungan gambar di atas yaitu di ambil dari urutan tempat kejadian yang memiliki jumlah paling sedikit (dari bagian bawah)

Selanjutnya dihitung banyaknya data untuk atribut :

```
#menghitung banyaknya data berdasarkan jenis kasus
dfKasus['Jenis Kasus'].value_counts()
```

Setelah dijalankan perintah diatas maka didapatkan hasil sebagai berikut :

```
Kekerasan      86
Pencurian       65
Pemerkosaan     61
Penculikan      28
Pembunuhan      19
Name: Jenis Kasus, dtype: int64
```

Gambar 4. 4 Urutan Dataset

Cara perhitungan gambar di atas yaitu di ambil dari urutan jenis kasus yang memiliki jumlah paling sedikit (dari bagian bawah)

6. Konversi Nilai Atribut

Semua atribut yang bertipe string atau non numerik harus dilakukan konversi nilai agar bisa digunakan dalam pemodelan k-means. Berikut perintah untuk konversi nilai atribut tempat kejadian dan jenis kasus :

```
# Konversi Data String ke Angka
dfKasus['Tempat Kejadian'].replace(['Kota Utara','Hulontha
langi','Kota Tengah',
                                   'Kota Selatan','Dungin
gi','Kota Timur',
                                   'Dumbo Raya','Kota Bar
at','Sipatana'],
                                   [1,2,3,4,5,6,7,8,9], i
nplace=True)
dfKasus['Jenis Kasus'].replace(['Pembunuhan','Penculikan',
                                'Pemeriksaan','Pencurian','Kekerasan'],
                                [1,2,3,4,5], inplace=True)
```

Berikut hasil konversi nilai atribut :

	Bulan	Tempat Kejadian	Jenis Kasus	Jumlah
0	Jan 2021	7	5	2
1	Jan 2021	4	5	2
2	Jan 2021	1	5	1
3	Jan 2021	6	5	2
4	Jan 2021	5	5	1
...
254	Des 2021	5	2	1
255	Des 2021	3	2	1
256	Des 2021	4	3	2
257	Des 2021	9	3	1
258	Des 2021	2	3	3

259 rows x 4 columns

Gambar 4.5 Hasil Konversi Data

7. Penentuan Kolom Dataset

Sebelum dilakukan pemodelan k-means atribut yang akan digunakan perlu ditentukan terlebih dahulu yaitu atribut : Tempat Kejadian, Jenis Kasus, Jumlah dengan perintah sebagai berikut :

```
# Memilih/menentukan kolom 1, 2 dan 3 sebagai dataset
x = dfKasus[['Tempat Kejadian', 'Jenis Kasus', 'Jumlah']].values
x
```

Setelah dijalankan perintah diatas mendapatkan hasil sebagai berikut :

```
array([[7, 5, 2],
       [4, 5, 2],
       [1, 5, 1],
       [6, 5, 2],
       [5, 5, 1],
       [8, 5, 2],
       [3, 5, 2],
       [9, 5, 4],
       [2, 5, 2],
       [7, 4, 1],
       [4, 4, 1],
       [6, 4, 1],
       [3, 4, 1],
       [9, 4, 1],
       [2, 4, 1],
       [5, 1, 2],
       [9, 2, 1],
       [8, 3, 1],
       [7, 5, 2],
       [4, 5, 2],
       [1, 5, 4],
```

Gambar 4. 6 Penentuan Kolom Dataset

Selanjutnya dilakukan pemodelan metode K-Means dan akan dibahas pada bab selanjutnya.

BAB V

PEMBAHASAN PENELITIAN

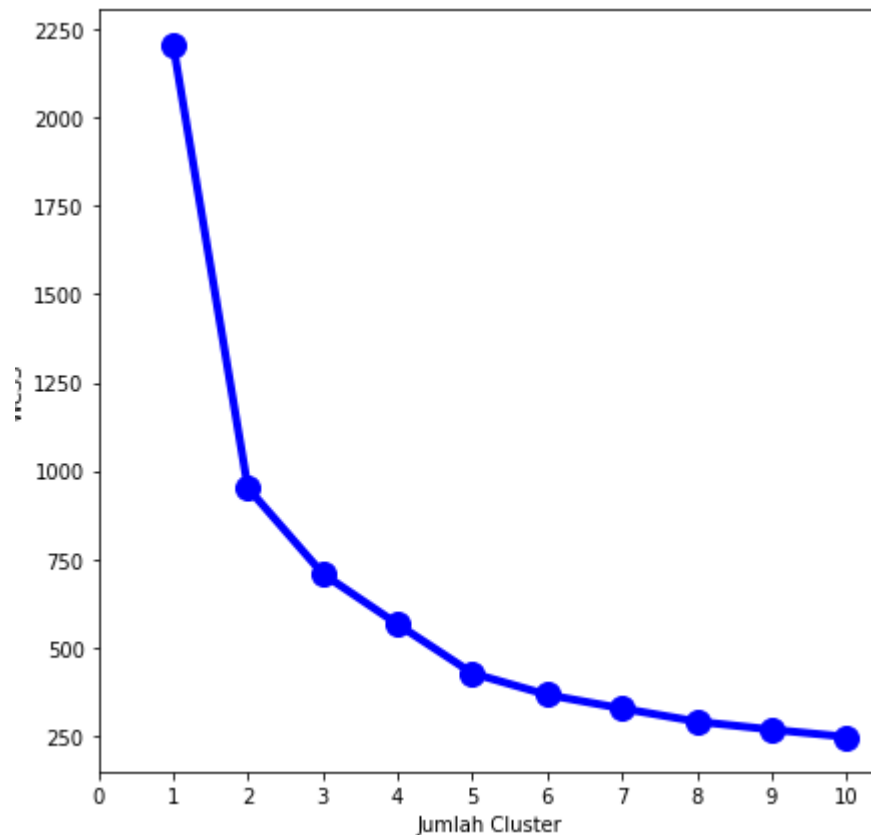
5.1 Kinerja Model

Kinerja metode *K-Means* dalam pengelompokkan data sangat ditentukan dari jumlah cluster yang ditentukan, agar tidak membutuhkan waktu yang lama dalam mengujicoba berapa jumlah cluster yang tepat, maka diperlukan suatu teknik untuk menentukan jumlah cluster yang optimum yaitu dengan menggunakan teknik *Elbow*. Berikut potongan coding penggunaan teknik *Elbow*:

```
# Menentukan jumlah cluster yang optimum dengan teknik Elbow

WCSS = []
for i in range(1,11):
    model = KMeans(n_clusters = i,init = 'k-means++')
    model.fit(x)
    WCSS.append(model.inertia_)
fig = plt.figure(figsize = (7,7))
plt.plot(range(1,11),WCSS, linewidth=4, markersize=12,marker=
'o',color = 'blue')
plt.xticks(np.arange(11))
plt.xlabel("Jumlah Cluster")
plt.ylabel("WCSS")
plt.show()
```

Hasil dari potongan coding di atas adalah sebagai berikut :



Gambar 5.1 Visualisasi Teknik *Elbow*

Berdasarkan gambar 5.1 di atas dapat dilihat bahwa jumlah *cluster* yang ada sebanyak 10 cluster. Kemudian di ambil cluster yang nlainya optimum sebanyak 5 cluster. Sehingga untuk penerapan metode K-Means dapat langsung digunakan jumlah cluster = 5.

5.2 Pembahasan Model

5.2.1 Pemodelan Metode K-Means

Setelah didapatkan jumlah cluster yang optimum, maka tahap selanjutnya adalah merapkan dalam pemodelan K-Means dengan menggunakan potongan coding sebagai berikut :

```
# Pemodelan KMeans dengan jumlah cluster yang optimum
model = KMeans(n_clusters = 5, init = "k-
means++", max_iter = 300, n_init = 10, random_state = 0)
y_clusters = model.fit_predict(x)
```

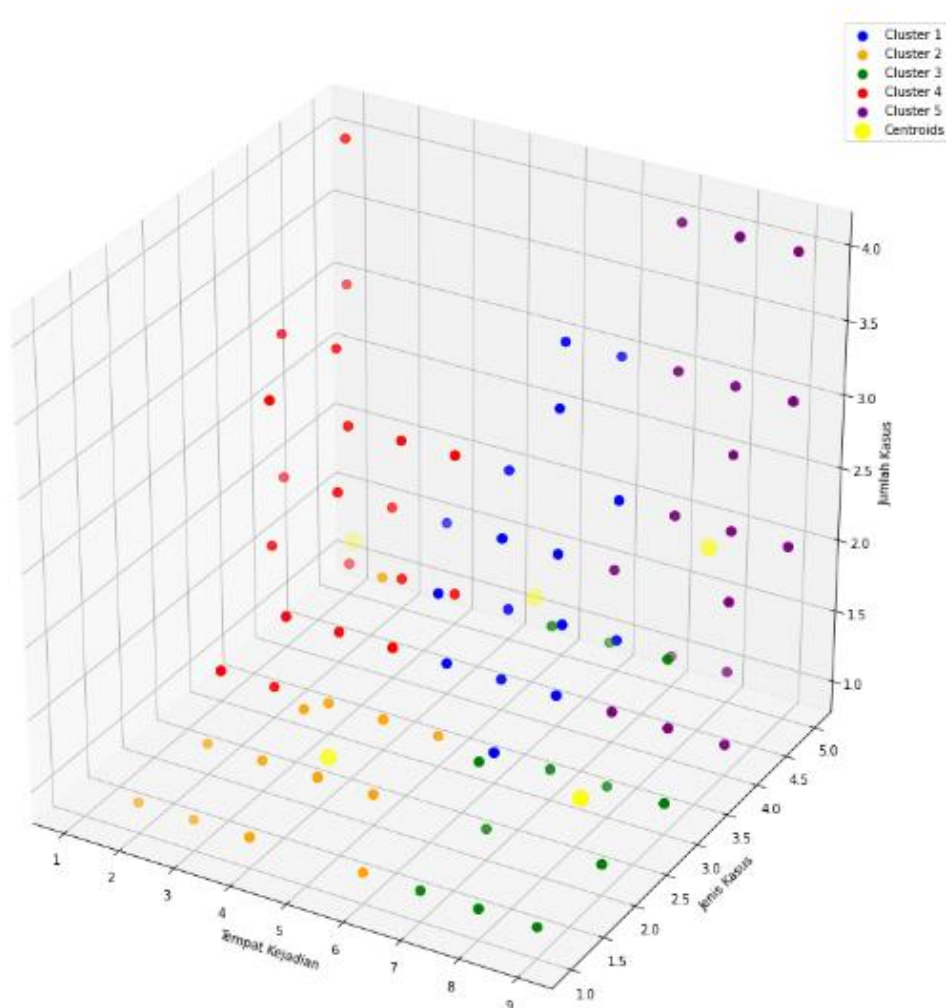
5.2.2 Visualisasi Hasil Pemodelan K-Means

Setelah dilakukan pemodelan K-Means, maka selanjutnya dapat dibuatkan visualisasi setiap cluster yang terbentuk dengan menggunakan potongan coding sebagai berikut :

```
# Visualisasi hasil cluster dengan 3d scatterplot menggunakan
matplotlib

fig = plt.figure(figsize = (15,15))
ax = fig.add_subplot(111, projection='3d')
ax.scatter(x[y_clusters == 0,0],x[y_clusters == 0,1],x[y_clusters == 0,2], s = 50 , color = 'blue', label = "Cluster 1")
ax.scatter(x[y_clusters == 1,0],x[y_clusters == 1,1],x[y_clusters == 1,2], s = 50 , color = 'orange', label = "Cluster 2")
ax.scatter(x[y_clusters == 2,0],x[y_clusters == 2,1],x[y_clusters == 2,2], s = 50 , color = 'green', label = "Cluster 3")
ax.scatter(x[y_clusters == 3,0],x[y_clusters == 3,1],x[y_clusters == 3,2], s = 50 , color = 'red', label = "Cluster 4")
ax.scatter(x[y_clusters == 4,0],x[y_clusters == 4,1],x[y_clusters == 4,2], s = 50 , color = 'purple', label = "Cluster 5")
ax.scatter(model.cluster_centers_[0],model.cluster_centers_[1],model.cluster_centers_[2], s = 150, c = "yellow", label = "Centroids")
ax.set_xlabel('Tempat Kejadian')
ax.set_ylabel('Jenis Kasus')
ax.set_zlabel('Jumlah Kasus')
ax.legend()
plt.show()
```

Hasil potongan coding di atas dapat dilihat pada gambar berikut :



Gambar 5.2 Visualisasi Hasil Clustering

Berdasarkan gambar di atas, dapat dilihat bahwa setiap cluster diwakili titik berwarna yaitu cluster 1 warna biru, cluster 2 warna coklat, cluster 3 warna hijau, cluster 4 warna merah, cluster 5 warna ungu dan pusat cluster warna kuning. Pada gambar di atas juga dengan bantuan visualisasi 3 dimensi secara sejelas dapat dilihat keterkaitan ketiga atribut yang digunakan dalam clustering.

5.2.3 Pusat Cluster dan Jumlah Data

Proses analisa untuk menarik kesimpulan terhadap setiap cluster, selain langsung melihat pada gambar di atas perlu diketahui nilai pusat cluster untuk masing-cluster beserta banyak data dalam setiap clusetr. Berikut potongan coding untuk mengetahui pusat masing cluster.

```
#Cetak Pusat Cluster
print(model.cluster_centers_)
```

Hasil output potongan coding di atas sebagai berikut :

```
[[5.18965517 4.36206897 1.46551724]
 [3.97916667 2.1875      1.0625     ]
 [8.23809524 2.35714286 1.21428571]
 [1.96551724 4.25862069 1.55172414]
 [8.01886792 4.60377358 2.03773585]]
```

Banyaknya data dalam setiap cluster dapat ditampilkan dengan potongan coding sebagai berikut :

```
#Hitung Jumlah Data Per Cluster
dfKasus['Kluster'].value_counts()
```

Hasil output potongan coding di atas sebagai berikut :

```
0      58
3      58
4      53
1      48
2      42
Name: Kluster, dtype: int64
```

Agar mudah dipahami data di atas, dibuatkan dalam bentuk tabel sebagai berikut:

Tabel 5. 1 Hasil Output

Cluster	Pusat Cluster			Jumlah Data
	Tempat Kejadian	Jenis Kasus	Jumlah Kasus	
1	5.18	4.36	1.46	58
2	3.97	2.18	1.06	48
3	8.23	2.35	1.21	42
4	1.96	4.25	1.55	58
5	8.01	4.60	2.03	53

Dari tabel di atas, dapat dianalisis atau disimpulkan bahwa :

- Cluster 1 adalah merupakan daerah yang memiliki kasus pencurian dan kekerasan dengan jumlah kasus kecil yaitu Duingi dan Kota Timur
- Cluster 2 adalah merupakan daerah yang memiliki kasus penculikan dan pemerkosaan dengan jumlah kasus rendah yaitu daerah Kota Tengah dan Kota Selatan
- Cluster 3 adalah merupakan daerah yang memiliki kasus penculikan dan pemerkosaan dengan jumlah kasus sedang yaitu daerah Dumbo Raya, Kota Barat dan Sipatana
- Cluster 4 adalah merupakan daerah yang memiliki kasus pencurian dan kekerasan dengan jumlah kasus sedang yaitu Kota Utara dan Hulondalangi
- Cluster 5 adalah merupakan daerah yang memiliki kasus pencurian dan kekerasan dengan jumlah kasus banyak yaitu Dumbo Raya, Kota barat dan Sipatana

5.2.4 Pelabelan Data Hasil Cluster

Setiap data dapat diketahui masuk dalam cluster berapa dengan menggunakan potongan coding sebagai berikut :

```
# Menampilkan hasil kluster
print(model.labels_)
# Menambahkan kolom "kluster" dalam data frame Data Nilai
dfKasus["Kluster"] = model.labels_
dfKasus
```

Hasil output potongan coding di atas adalah sebagai berikut :

```
[4 0 3 0 0 4 3 4 3 4 0 0 3 4 3 1 2 2 4 0 3 0 0 4 3 4 3 4 0 4 1 1 1 2 0 2 3
 0 3 0 0 4 3 4 3 4 3 0 0 4 3 3 2 1 2 1 1 1 3 0 1 2 4 3 0 4 3 4 3 4 3 0 3 2
 1 1 1 3 0 1 2 1 2 1 2 0 0 0 4 4 0 0 0 4 3 3 1 2 1 2 1 2 2 1 1 2 4 3 0 4 3
 4 0 4 3 3 1 2 2 1 1 2 1 1 2 4 4 3 4 3 3 0 0 4 4 1 1 2 2 2 4 0 3 0 0 4 3 4
 3 0 0 4 3 2 2 1 3 0 1 4 0 0 0 4 3 4 3 4 0 3 3 4 3 2 2 1 3 0 1 1 3 0 2 0 3
 0 0 4 3 4 3 4 0 3 0 4 3 4 2 1 2 1 3 0 2 1 1 2 2 1 2 3 4 0 0 0 4 3 4 3 0 3
 0 0 2 1 1 2 2 1 2 3 1 1 1 2 3 4 0 0 4 3 4 4 0 0 0 4 3 4 2 1 1 2 1 1 1 2 3]
```

Keterangan lebih lanjut untuk nilai cluster bisa dibuatkan kondisi dengan potongan coding sebagai berikut :

```
#Menambahkan kolom keterangan cluster
#Menambahkan kolom keterangan cluster
conditions = [
    (dfKasus['Kluster']==0),
    (dfKasus['Kluster']==1),
    (dfKasus['Kluster']==2),
    (dfKasus['Kluster']==3),
    (dfKasus['Kluster']==4)]
choices = ['C1', 'C2', 'C3', 'C4', 'C5']
dfKasus['Ket'] = np.select(conditions, choices)
dfKasus
```

```
#Simpan Hasil Cluster ke Excel
dfKasus.to_excel("/content/drive/MyDrive/nancy/HasilKasus.xls
x", index=False)
```

Hasil akhir setelah dilakukan kedua perintah di atas, maka hasilnya dapat dilihat pada tabel berikut :

Tabel 5. 2 Hasil Pelabelan Cluster

No	Bulan	Tempat Kejadian	Jenis Kasus	Jumlah	Kluster	Ket
1	Jan 2021	7	5	2	4	C5
2	Jan 2021	4	5	2	0	C1
3	Jan 2021	1	5	1	3	C4
4	Jan 2021	6	5	2	0	C1
5	Jan 2021	5	5	1	0	C1
6	Jan 2021	8	5	2	4	C5
7	Jan 2021	3	5	2	3	C4
8	Jan 2021	9	5	4	4	C5
9	Jan 2021	2	5	2	3	C4

No	Bulan	Tempat Kejadian	Jenis Kasus	Jumlah	Kluster	Ket
10	Jan 2021	7	4	1	4	C5
11	Jan 2021	4	4	1	0	C1
12	Jan 2021	6	4	1	0	C1
13	Jan 2021	3	4	1	3	C4
14	Jan 2021	9	4	1	4	C5
15	Jan 2021	2	4	1	3	C4
16	Jan 2021	5	1	2	1	C2
17	Jan 2021	9	2	1	2	C3
18	Jan 2021	8	3	1	2	C3
19	Feb 2021	7	5	2	4	C5
...
259	Des 2021	2	3	3	3	C4

BAB VI

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Berdasarkan hasil penelitian yang sudah diuraikan di atas tentang clustering kasus kepolisian dengan metode K-Means, maka dapat ditarik kesimpulan yaitu, jumlah cluster yang optimum untuk melakukan Clustering Kasus Kepolisian Dengan Metode *K-Means* yaitu sebanyak 5 cluster dengan menggunakan teknik Elbow. Hasil clustering untuk masing-masing cluster merupakan daerah yang memiliki kasus pencurian dan kekerasan dengan jumlah kasus kecil, daerah yang memiliki kasus penculikan dan pemerkosaan dengan jumlah kasus rendah, daerah yang memiliki kasus penculikan, pemerkosaan, pencurian dan kekerasan dengan jumlah kasus sedang, daerah yang memiliki kasus pencurian dan kekerasan dengan jumlah kasus banyak.

6.2 Saran

Berdasarkan kesimpulan tersebut diatas, peneliti dapat memberikan saran untuk penelitian selanjutnya, yaitu :

1. Penelitian selanjutnya dapat menggunakan atribut lain atau menambah dataset lebih banyak lagi agar hasil clusteringnya bisa dibandingkan.
2. Dapat dikembangkan dengan menggunakan metode clustering yang lain agar hasil bisa dibandingkan mana yang lebih baik.

DAFTAR PUSTAKA

- [1] “MEMAHAMI APA ITU DATA MINING ? – Accounting.” <https://accounting.binus.ac.id/2019/10/03/memahami-apa-itu-data-mining/> (accessed Jan. 06, 2022).
- [2] “Sepanjang 2019, Tingkat Kejahatan Menurun 19%.” <https://www.beritasatu.com/nasional/592780/sepanjang-2019-tingkat-kejahatan-menurun-19> (accessed Sep. 07, 2021).
- [3] W. Dhuhita, “Clustering Menggunakan Metode K-Mean Untuk Menentukan Status Gizi Balita,” *J. Inform. Darmajaya*, vol. 15, no. 2, pp. 160–174, 2015.
- [4] Asroni and R. Adrian, “Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang,” *J. Ilm. Semesta Tek.*, vol. 18, no. 1, pp. 76–82, 2015.
- [5] W. Afifi, D. R. Nastiti, and Q. Aini, “Clustering K-Means Pada Data Ekspor (Studi Kasus: Pt. Gaikindo),” *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 11, no. 1, pp. 45–50, 2020, doi: 10.24176/simet.v11i1.3568.
- [6] Sugiono, S. Nurdiani, S. Linawati, R. A. Safitri, and E. P. Saputra, “Pengelompokan Perilaku Mahasiswa Pada Perkuliahan E-Learning dengan K-Means Clustering,” *J. Kaji. Ilm.*, vol. 19, no. 2, pp. 126–133, 2019.
- [7] I. N. Farida, M. Kom, D. P. Pamungkas, and M. Kom, “ARTIKEL CLUSTERING WILAYAH PADA PELANGGARAN BERKENDARAAN MENGGUNAKAN METODE K-MEANS (Studi Kasus : Polres Kediri) Oleh : WULAN FAJARIYANTI Dibimbing oleh : SURAT PERNYATAAN ARTIKEL SKRIPSI TAHUN 2017,” vol. 01, no. 09, 2017.
- [8] “Arti kata kasus - Kamus Besar Bahasa Indonesia (KBBI) Online.” <https://kbbi.web.id/kasus> (accessed Jan. 19, 2022).
- [9] “Website Resmi Polri.” <https://www.polri.go.id/layanan-spkt> (accessed Sep. 07, 2021).
- [10] N. K. Afkarina, A. W. Widodo, and M. T. Furqon, “Implementasi Regresi Linier Berganda Untuk Prediksi Jumlah Peminat,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 11, pp. 10462–10467, 2019.
- [11] “Data Mining: Pengertian, Fungsi, Penerapan, dan Metode Pengambilannya - Cermati.com.” <https://www.cermati.com/artikel/data-mining-pengertian-fungsi-penerapan-dan-metode-pengambilannya> (accessed Dec. 04, 2021).
- [12] B. M. Metisen and H. L. Sari, “Analisis clustering menggunakan metode K-Means dalam mengelompokkan penjualan produk pada Swalayan Fadhila,”

J. Media Infotama, vol. 11, no. 2, pp. 110–118, 2015.

- [13] N. Wakhidah, “Clustering Menggunakan K-Means Algorithm,” *J. Transform.*, vol. 8, no. 1, p. 33, 2010, doi: 10.26623/transformatika.v8i1.45.
- [14] R. Sovia, E. P. W. Mandala, and S. Mardhiah, “Algoritma K-Means dalam Pemilihan Siswa Berprestasi dan Metode SAW untuk Prediksi Penerima Beasiswa Berprestasi,” *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 2, p. 181, 2020, doi: 10.26418/jp.v6i2.37759.
- [15] D. Putra and A. Wibowo, “Prediksi Keputusan Minat Penjurusan Siswa SMA Yadika 5 Menggunakan Algoritma Naïve Bayes,” *Pros. Semin. Nas. Ris. ...*, vol. 2, pp. 84–92, 2020, [Online]. Available: <http://tunasbangsa.ac.id/seminar/index.php/senaris/article/view/147>.
- [16] “SIKLUS HIDUP PENGEMBANGAN SISTEM.” <http://norhidayah19.blogspot.com/2018/04/siklus-hidup-pengembangan-sistem.html> (accessed Dec. 07, 2021).

Lampiran Data Operasional

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
1.	1 Jan 2021	Dumbo Raya	Kekerasan	2
		Kota Selatan	Pencurian	1
2.	2 Jan 2021	Hulothalangi	Kekerasan	1
3.	3 Jan 2021	Kota Timur	Kekerasan	1
4.	4 Jan 2021	Dungingi	Pembunuhan	1
5.	5 Jan 2021	Kota Barat	Kekerasan	4
6.	6 Jan 2021	Kota Tengah	Kekerasan	2
		Sipatana	Kekerasan	2
7.	8 Jan 2021	Kota Selatan	Kekerasan	1
8.	9 Jan 2021	Kota Barat	Pemeriksaan	1
9.	10 Jan 2021	Sipatana	Pencurian	2
10.	11 Jan 2021	Dumbo Raya	Pencurian	2
11.	12 Jan 2021	Dumbo Raya	Kekerasan	1
12.	14 Jan 2021	Sipatana	Kekerasan	3
13.	15 Jan 2021	Sipatana	Penculikan	1
14.	16 Jan 2021	Dungingi	Pembunuhan	1
15.	17 Jan 2021	Dungingi	Kekerasan	1
16.	19 Jan 2021	Kota Selatan	Kekerasan	2
17.	20 Jan 2021	Kota Barat	Kekerasan	2
18.	21 Jan 2021	Kota Timur	Pencurian	2
19.	22 Jan 2021	Kota Timur	Kekerasan	1
20.	23 Jan 2021	Kota Tengah	Kekerasan	2

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
21.	25 Jan 2021	Hulonthalangi	Pencurian	1
22.	26 Jan 2021	Hulonthalangi	Kekerasan	1
		Kota Tengah	Pencurian	1
		Sipatana	Kekerasan	1
23.	27 Jan 2021	Sipatana	Kekerasan	1
24.	30 Jan 2021	Kota Utara	Kekerasan	3
25.	1 Feb 2021	Kota Utara	Kekerasan	2
26.	2 Feb 2021	Kota Selatan	Kekerasan	1
		Dumbo Raya	Pencurian	1
		Kota Selatan	Kekerasan	1
27.	3 Feb 2021	Dungingi	Pencurian	1
28.	4 Feb 2021	Kota Tengah	Kekerasan	1
29.	5 Feb 2021	Kota Timur	Pembunuhan	1
30.	6 Feb 2021	Kota Utara	Kekerasan	1
31.	7 Feb 2021	Sipatana	Kekerasan	2
32.	9 Feb 2021	Kota Barat	Pemeriksaan	1
		Hulonthalangi	Pemeriksaan	1
33.	10 Feb 2021	Kota Utara	Kekerasan	1
34.	11 Feb 2021	Kota Barat	Kekerasan	2
35.	12 Feb 2021	Kota Selatan	Pembunuhan	1
36.	13 Feb 2021	Kota Tengah	Penculikan	1
37.	14 Feb 2021	Dumbo Raya	Kekerasan	2
38.	15 Feb 2021	Dungingi	Kekerasan	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
39.	16 Feb 2021	Hulonthalangi	Kekerasan	2
40.	17 Feb 2021	Dungingi	Pencurian	1
41.	18 Feb 2021	Sipatana	Kekerasan	1
42.	19 Feb 2021	Sipatana	Pencurian	2
43.	20 Feb 2021	Kota Timur	Pemeriksaan	1
44.	21 Feb 2021	Kota Barat	Pemeriksaan	1
45.	22 Feb 2021	Dumbo Raya	Pemeriksaan	1
46.	23 Feb 2021	Dumbo Raya	Kekerasan	2
47.	24 Feb 2021	Sipatana	Kekerasan	2
48.	25 Feb 2021	Kota Timur	Kekerasan	1
49.	27 Feb 2021	Kota Utara	Kekerasan	1
50.	28 Feb 2021	Hulonthalangi	Pemeriksaan	1
51.	1 Mar 2021	Dumbo Raya	Pencurian	2
52.	2 Mar 2021	Dungingi	Pencurian	1
53.	3 Mar 2021	Hulonthalangi	Kekerasan	1
		Hulonthalangi	Pencurian	1
54.	4 Mar 2021	Kota Barat	Kekerasan	1
55.	5 Mar 2021	Kota Selatan	Pembunuhan	4
56.	6 Mar 2021	Kota Timur	Kekerasan	1
57.	7 Mar 2021	Kota Utara	Kekerasan	1
58.	8 Mar 2021	Sipatana	Kekerasan	1
		Kota Utara	Kekerasan	1
59.	9 Mar 2021	Dungingi	Kekerasan	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
60.	10 Mar 2021	Dumbo Raya	Pencurian	1
61.	11 Mar 2021	Kota Tengah	Kekerasan	2
62.	12 Mar 2021	Kota Barat	Kekerasan	1
63.	13 Mar 2021	Sipatana	Kekerasan	1
64.	14 Mar 2021	Dungingi	Kekerasan	1
65.	15 Mar 2021	Kota Selatan	Kekerasan	1
66.	16 Mar 2021	Kota Timur	Pencurian	2
67.	17 Mar 2021	Kota Barat	Pencurian	1
68.	18 Mar 2021	Kota Utara	Kekerasan	1
69.	19 Mar 2021	Kota Utara	Pencurian	1
70.	20 Mar 2021	Sipatana	Kekerasan	2
71.	21 Mar 2021	Sipatana	Pembunuhan	1
72.	22 Mar 2021	Dumbo Raya	Pembunuhan	1
73.	23 Mar 2021	Dungingi	Kekerasan	1
74.	24 Mar 2021	Hulonthalangi	Kekerasan	1
75.	25 Mar 2021	Kota Barat	Kekerasan	2
76.	26 Mar 2021	Kota Selatan	Kekerasan	1
77.	27 Mar 2021	Kota Tengah	Pencurian	1
78.	28 Mar 2021	Kota Timur	Kekerasan	1
79.	1 Apr 2021	Kota Utara	Kekerasan	1
80.	2 Apr 2021	Sipatana	Kekerasan	2
81.	3 Apr 2021	Dumbo Raya	Kekerasan	1
82.	4 Apr 2021	Kota Barat	Kekerasan	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
83.	5 Apr 2021	Kota timur	Pencurian	1
84.	6 Apr 2021	Dungingi	Penculikan	1
85.	7 Apr 2021	Kota Tengah	Pemeriksaan	1
86.	8 Apr 2021	Kota Selatan	Pemeriksaan	1
87.	9 Apr 2021	Sipatana	Pemeriksaan	1
88.	10 Apr 2021	Sipatana	Kekerasan	1
89.	12 Apr 2021	Kota Tengah	Kekerasan	1
90.	13 Apr 2021	Kota Barat	Pembunuhan	1
91.	14 Apr 2021	Kota Utara	Pemeriksaan	1
92.	15 Apr 2021	Dumbo Raya	Pencurian	1
93.	16 Apr 2021	Hulonthalangi	Pencurian	1
94.	17 Apr 2021	Kota Timur	Pemeriksaan	1
95.	18 Apr 2021	Kota Selatan	Penculikan	1
96.	19 Apr 2021	Kota Tengah	Kekerasan	1
97.	20 Apr 2021	Dumbo Raya	Kekerasan	1
98.	21 Apr 2021	Kota Barat	Pembunuhan	1
99.	22 Apr 2021	Kota Utara	Pencurian	1
100.	24 Apr 2021	Dungingi	Kekerasan	1
101.	26 Apr 2021	Sipatana	Kekerasan	1
102.	27 Apr 2021	Hulonthalangi	Kekerasan	1
103.	28 Apr 2021	Sipatana	Pemeriksaan	1
104.	29 Apr 2021	Dumbo Raya	Kekerasan	1
105.	1 Mei 2021	Kota Barat	Kekerasan	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
		Kota Timur	Kekerasan	1
106.	2 Mei 2021	Sipatana	Kekerasan	1
107.	3 Mei 2021	Kota Selatan	Pencurian	1
108.	4 Mei 2021	Hulonthalangi	Pencurian	1
109.	5 Mei 2021	Dungingi	Pencurian	1
110.	6 Mei 2021	Dumbo Raya	Pemeriksaan	1
111.	8 Mei 2021	Kota Selatan	Kekerasan	1
112.	9 Mei 2021	Dungingi	Pencurian	1
113.	10 Mei 2021	Kota Barat	Kekerasan	2
114.	11 Mei 2021	Kota Tengah	Pencurian	1
115.	12 Mei 2021	Kota Timur	Kekerasan	1
116.	13 Mei 2021	Kota Selatan	Pemeriksaan	1
117.	14 Mei 2021	Dumbo Raya	Pemeriksaan	1
118.	15 Mei 2021	Kota Selatan	Penculikan	1
119.	17 Mei 2021	Sipatana	Pembunuhan	1
120.	19 Mei 2021	Kota Barat	Pencurian	1
121.	20 Mei 2021	Kota Barat	Kekerasan	1
122.	21 Mei 2021	Dungingi	Kekerasan	1
123.	22 Mei 2021	Hulonthalangi	Pencurian	1
124.	23 Mei 2021	Kota Timur	Pencurian	1
125.	24 Mei 2021	Sipatana	Pemeriksaan	1
126.	25 Mei 2021	Kota Barat	Kekerasan	1
127.	26 Mei 2021	Kota Tengah	Pembunuhan	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
128.	27 Mei 2021	Sipatana	Kekerasan	1
129.	1 Jun 2021	Kota Tengah	Kekerasan	1
130.	2 Jun 2021	Kota Timur	Kekerasan	1
131.	3 Jun 2021	Hulonthalangi	Pencurian	1
132.	4 Jun 2021	Dungingi	Pemeriksaan	1
133.	5 Jun 2021	Kota Barat	Kekerasan	1
134.	6 Jun 2021	Kota Tengah	Pemeriksaan	1
135.	7 Jun 2021	Kota Tengah	Kekerasan	1
136.	9 Jun 2021	Kota Utara	Kekerasan	1
137.	10 Jun 2021	Kota Timur	Pencurian	1
138.	11 Jun 2021	Sipatana	Penculikan	1
139.	12 Jun 2021	Kota Selatan	Pembunuhan	1
140.	13 Jun 2021	Kota barat	Kekerasan	1
141.	14 Jun 2021	Kota Tengah	Pencurian	1
142.	15 Jun 2021	Kota Barat	Kekerasan	1
143.	17 Jun 2021	Sipatana	Kekerasan	1
144.	18 Jun 2021	Sipatana	Pemeriksaan	1
145.	19 Jun 2021	Kota Tengah	Pencurian	1
146.	20 Jun 2021	Kota Barat	Kekerasan	1
147.	21 Jun 2021	Kota Utara	Kekerasan	1
148.	22 Jun 2021	Hulonthalangi	Pencurian	1
149.	23 Jun 2021	Sipatana	kekeraan	1
150.	24 Jun 2021	Kota Barat	Kekerasan	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
151.	26 Jun 2021	Dumbo Raya	Kekerasan	1
152.	28 Jun 2021	Kota Timur	Pencurian	1
153.	29 Jun 2021	Kota Barat	Kekerasan	1
154.	30 Jun 2021	Hulonthalangi	Pencurian	1
155.	1 Jul 2021	Kota Selatan	Penculikan	1
156.	2 Jul 2021	Kota Timur	Pencurian	1
157.	3 Jul 2021	Dumbo Raya	Kekerasan	1
158.	4 Jul 2021	Sipatana	Pencurian	1
159.	5 Jul 2021	Sipatana	Kekerasan	2
160.	6 Jul 2021	Kota Tengah	Kekerasan	1
161.	7 Jul 2021	Dungingi	Pencurian	1
162.	8 Jul 2021	Dumbo Raya	Kekerasan	1
163.	10 Jul 2021	Kota Barat	Kekerasan	1
164.	11 Jul 2021	Kota Timur	Pencurian	1
165.	12 Jul 2021	Kota Barat	Pencurian	1
166.	13 Jul 2021	Hulonthalangi	Kekerasan	2
167.	14 Jul 2021	Kota Barat	Kekerasan	1
168.	16 Jul 2021	Sipatana	Pencurian	1
		Kota Barat	Kekerasan	2
169.	18 Jul 2021	Sipatana	Pemeriksaan	1
170.	19 Jul 2021	Dumbo Raya	Kekerasan	1
171.	21 Jul 2021	Kota Tengah	Kekerasan	3
172.	22 Jul 2021	Kota Timur	Pencurian	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
173.	23 Jul 2021	Kota Utara	Pencurian	1
174.	24 Jul 2021	Sipatana	Kekerasan	1
175.	25 Jul 2021	Sipatana	Pemeriksaan	1
176.	27 Jul 2021	Dungingi	Pencurian	1
177.	28 Jul 2021	Hulonthalangi	Penculikan	1
178.	29 Jul 2021	Dumbo Raya	Kekerasan	1
179.	30 Jul 2021	Sipatana	Kekerasan	1
180.	1 Ags 2021	Kota Tengah	Kekerasan	1
181.	2 Ags 2021	Kota Timur	Pencurian	1
182.	3 Ags 2021	Sipatana	Pencurian	1
		Dumbo Raya	Pencurian	1
183.	4 Ags 2021	Dumbo Raya	Pemeriksaan	1
184.	5 Ags 2021	Sipatana	Penculikan	1
185.	6 Ags 2021	Hulonthalangi	Pencurian	1
186.	7 Ags 2021	Dungingi	Pencurian	1
187.	8 Ags 2021	Kota Barat	Kekerasan	2
188.	9 Ags 2021	Kota selatan	Kekerasan	1
189.	10 Ags 2021	Kota Tengah	Kekerasan	1
190.	11 Ags 2021	Kota Timur	Pencurian	1
191.	12 Ags 2021	Kota Utara	Pencurian	1
192.	13 Ags 2021	Dumbo raya	Kekerasan	1
193.	14 Ags 2021	Dumbo Raya	Kekerasan	1
194.	15 Ags 2021	Dungingi	Kekerasan	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
195.	16 Ags 2021	Kota Utara	Pencurian	1
196.	17 Ags 2021	Kota Utara	Kekerasan	1
197.	18 Ags 2021	Hulonthalangi	Pencurian	1
198.	19 Ags 2021	Kota Timur	Kekerasan	2
199.	20 Ags 2021	Kota Barat	Kekerasan	1
200.	21 Ags 2021	Kota Utara	Kekerasan	1
201.	23 Ags 2021	Dumbo Raya	Pencurian	1
202.	25 Ags 2021	Sipatana	Pencurian	1
203.	26 Ags 2021	Dungingi	Kekerasan	1
204.	27 Ags 2021	Dungingi	Kekerasan	1
205.	29 Ags 2021	Sipatana	Pencurian	2
206.	30 Ags 2021	Dumbo Raya	Kekerasan	1
207.	1 Sep 2021	Kota Barat	Kekerasan	1
208.	2 Sep 2021	Kota Timur	Kekerasan	1
209.	3 Sep 2021	Kota Tengah	Kekerasan	2
210.	4 Sep 2021	Kota Selatan	Pencurian	1
211.	5 Sep 2021	Kota Utara	Pencurian	1
212.	6 Sep 2021	Sipatana	Pencurian	1
213.	7 Sep 2021	Hulonthalangi	Kekerasan	1
214.	8 Sep 2021	Dumbo Raya	Penculikan	1
215.	9 Sep 2021	Dumbo Raya	Kekerasan	1
216.	10 Sep 2021	Dungingi	Pemeriksaan	1
217.	11 Sep 2021	Sipatana	Kekerasan	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
218.	12 Sep 2021	Kota Timur	Kekerasan	2
219.	13 Sep 2021	Kota Utara	Pencurian	1
220.	14 Sep 2021	Kota Selatan	Kekerasan	1
221.	15 Sep 2021	Kota Barat	Kekerasan	1
222.	16 Sep 2021	Kota Timur	Kekerasan	1
223.	17 Sep 2021	Kota Tengah	Pencurian	1
224.	18 Sep 2021	Hulonthalangi	Kekerasan	2
225.	19 Sep 2021	Kota Utara	Pemeriksaan	2
226.	20 Sep 2021	Dumbo Raya	Pencurian	1
227.	21 Sep 2021	Hulonthalangi	Pencurian	3
228.	22 Sep 2021	Kota Timur	Pemeriksaan	1
229.	23 Sep 2021	Kota Selatan	Penculikan	1
230.	24 Sep 2021	Kota Tengah	Kekerasan	1
		Dumbo Raya	Kekerasan	2
231.	26 Sep 2021	Kota Barat	Pembunuhan	2
232.	27 Sep 2021	Kota Utara	Pencurian	2
233.	28 Sep 2021	Dungingi	Kekerasan	1
234.	29 Sep 2021	Sipatana	Kekerasan	2
235.	1 Okt 2021	Hulonthalangi	Kekerasan	1
236.	2 Okt 2021	Kota Selatan	Pemeriksaan	1
237.	3 Okt 2021	Kota Timur	Kekerasan	1
238.	4 Okt 2021	Kota Barat	Kekerasan	1
239.	5 Okt 2021	Kota Utara	Kekerasan	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
		Kota Utara	Kekerasan	3
240.	6 Okt 2021	Sipatana	Pencurian	2
241.	7 Okt 2021	Sipatana	Pencurian	1
242.	8 Okt 2021	Dumbo Raya	Pencurian	1
243.	9 Okt 2021	Dungingi	Pemeriksaan	1
244.	11 Okt 2021	Hulonthalangi	Kekerasan	1
245.	13 Okt 2021	Kota Barat	Pencurian	1
246.	14 Okt 2021	Kota Selatan	Kekerasan	1
247.	15 Okt 2021	Kota Tengah	Pencurian	1
248.	17 Okt 2021	Kota Timur	Kekerasan	2
249.	18 Okt 2021	Kota Utara	Pemeriksaan	1
250.	19 Okt 2021	Sipatana	Pemeriksaan	1
251.	20 Okt 2021	Dumbo Raya	Penculikan	1
252.	21 Okt 2021	Kota Barat	Pembunuhan	2
253.	22 Okt 2021	Kota timur	Pencurian	1
		Dungingi	Kekerasan	1
254.	23 Okt 2021	Kota Tengah	Kekerasan	2
255.	25 Okt 2021	Kota Selatan	Pencurian	1
256.	26 Okt 2021	Kota Utara	Pencurian	2
257.	27 Okt 2021	Dungingi	Pemeriksaan	1
258.	28 Okt 2021	Sipatana	Kekerasan	1
259.	29 Okt 2021	Hulonthalangi	Pembunuhan	2
260.	30 Okt 2021	Sipatana	Kekerasan	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
261.	1 Nov 2021	Dumbo Raya	Kekerasan	1
262.	2 Nov 2021	Kota Barat	Kekerasan	1
263.	3 Nov 2021	Kota Timur	Pencurian	2
264.	4 Nov 2021	Sipatana	Pemeriksaan	2
265.	5 Nov 2021	Kota Selatan	Kekerasan	1
266.	6 Nov 2021	Hulonthalangi	Pemeriksaan	1
267.	7 Nov 2021	Dungingi	Kekerasan	1
268.	8 Nov 2021	Dumbo Raya	Kekerasan	2
269.	9 Nov 2021	Kota Selatan	Pencurian	1
270.	10 Nov 2021	Dungingi	Penculikan	1
271.	11 Nov 2021	Kota Barat	Pembunuhan	1
272.	12 Nov 2021	Kota Tengah	Pemeriksaan	1
273.	14 Nov 2021	Kota Timur	Pencurian	4
274.	15 Nov 2021	Kota Selatan	Pencurian	1
275.	16 Nov 2021	Dumbo Raya	Pemeriksaan	1
276.	17 Nov 2021	Kota Selatan	Penculikan	1
277.	18 Nov 2021	Sipatana	Kekerasan	1
278.	19 Nov 2021	Kota Barat	Kekerasan	1
279.	20 Nov 2021	Kota Barat	Pembunuhan	1
		Dungingi	Pencurian	1
280.	21 Nov 2021	Hulonthalangi	Kekerasan	1
		Kota Timur	Kekerasan	1
		Sipatana	Kekerasan	1

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
281.	22 Nov 2021	Kota Barat	Pemeriksaan	1
282.	23 Nov 2021	Kota Tengah	Kekerasan	1
283.	24 Nov 2021	Sipatana	Kekerasan	1
284.	25 Nov 2021	Kota Tengah	Kekerasan	1
285.	26 Nov 2021	Kota Timur	Kekerasan	1
286.	27 Nov 2021	Kota Utara	Pencurian	1
287.	28 Nov 2021	Dungingi	Pencurian	1
288.	1 Des 2021	Sipatana	Pencurian	1
289.	2 Des 2021	Hulonthalangi	Pemeriksaan	1
290.	3 Des 2021	Sipatana	Kekerasan	2
291.	4 Des 2021	Dumbo Raya	Pencurian	1
292.	5 Des 2021	Kota Barat	Kekerasan	1
293.	6 Des 2021	Kota Timur	Pencurian	1
294.	7 Des 2021	Sipatana	Kekerasan	1
295.	8 Des 2021	Kota Selatan	Pemeriksaan	1
296.	9 Des 2021	Hulonthalangi	Pemeriksaan	1
297.	12 Des 2021	Dungingi	Penculikan	2
298.	13 Des 2021	Dumbo Raya	Pembunuhan	1
299.	14 Des 2021	Kota Selatan	Pencurian	1
300.	15 Des 2021	Dungingi	Kekerasan	2
301.	16 Des 2021	Kota Barat	Kekerasan	1
		Kota Tengah	Pencurian	1
302.	17 Des 2021	Kota Timur	Pencurian	3

No	Tanggal Laporan	Tempat Kejadian	Jenis Kasus	Total
303.	18 Des 2021	Kota Selatan	Pemeriksaan	1
304.	19 Des 2021	Dumbo Raya	Kekerasan	1
305.	20 Des 2021	Kota Selatan	Pembunuhan	1
306.	21 Des 2021	Sipatana	Kekerasan	1
307.	22 Des 2021	Kota Barat	Kekerasan	1
308.	23 Des 2021	Kota Barat	Kekerasan	1
309.	24 Des 2021	Dungingi	Pencurian	1
310.	25 Des 2021	Hulonthalangi	Pemeriksaan	1
311.	26 Des 2021	Kota Timur	Kekerasan	1
312.	27 Des 2021	Sipatana	Pemeriksaan	1
313.	28 Des 2021	Kota Barat	Kekerasan	1
314.	29 Des 2021	Kota Tengah	Kekerasan	3
315.	30 Des 2021	Sipatana	Pencurian	2
316.	31 Des 2021	Kota Tengah	Penculikan	1
		Kota Timur	Pembunuhan	1
		Sipatana	Kekerasan	2
		Kota Barat	Kekerasan	1

Lampiran Surat Keterangan Telah Melakukan Penelitian



SPKT POLDA GORONTALO
KABUPATEN GORONTALO

Jl. Ahmad A. Wahab, Pantungo Kcc. Telaga Biru – Kab Gorontalo 96211
E-mail: spktpoldagtlo@gmail.com

SURAT KETERANGAN

No: 050622/01

Yang bertanda tangan dibawah ini :

Nama : Muamin D. Hilipito
NRP : 64030383
Jabatan : Kepala SPKT Poldo Gorontalo
Unit Kerja : Poldo Gorontalo

Mencerangkan dengan sesungguhnya bahwa :

Nama : Nancy M. Hilipito
NIM : T3117071
Fakultas : Ilmu Komputer
Program Studi : Teknik Informatika

Telah melaksanakan penelitian di SPKT Poldo Gorontalo mulai Oktober 2021 sampai dengan Januari 2022 untuk memperoleh data guna penyusunan Tugas akhir Skripsi dengan judul : Clustering Kasus Kepolisian Dengan Metode *K-Means*

Demikian Surat Keterangan ini di buat untuk dapat digunakan sebagaimana mestinya.

Gorontalo, 05 Juni 2022

Kepala SPKT



MUAMIN D. HILIPITO

NRP. 64030383

Potongan Kode Program

Coding Konversi Data

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 259 entries, 0 to 258
Data columns (total 4 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Bulan                 259 non-null   object
1   Tempat Kejadian       259 non-null   object
2   Jenis Kasus           259 non-null   object
3   Jumlah                259 non-null   int64
dtypes: int64(1), object(3)
memory usage: 8.2+ KB
```

Coding Menentukan Jumlah Cluster Dengan Teknik Elbow

```
WCSS = []
for i in range(1,11):
    model = KMeans(n_clusters = i,init = 'k-means++')
    model.fit(x)
    WCSS.append(model.inertia_)
fig = plt.figure(figsize = (7,7))
plt.plot(range(1,11),WCSS, linewidth=4, markersize=12,marker='o',color = 'blue')
plt.xticks(np.arange(11))
plt.xlabel("Jumlah Cluster")
plt.ylabel("WCSS")
plt.show()
```

Coding Pemodelan K-Means dengan jumlah cluster yang optimum

```
model = KMeans(n_clusters = 5, init = "k-means++", max_iter = 300, n_init = 10, random_state = 0)
y_clusters = model.fit_predict(x)
```

Coding Visualisasi Hasil Cluster

```
fig = plt.figure(figsize = (15,15))
ax = fig.add_subplot(111, projection='3d')
ax.scatter(x[y_clusters == 0,0],x[y_clusters == 0,1],x[
y_clusters == 0,2], s = 50 , color = 'blue', label = "C
luster 1")
ax.scatter(x[y_clusters == 1,0],x[y_clusters == 1,1],x[
y_clusters == 1,2], s = 50 , color = 'orange', label =
"Cluster 2")
ax.scatter(x[y_clusters == 2,0],x[y_clusters == 2,1],x[
y_clusters == 2,2], s = 50 , color = 'green', label = "
Cluster 3")
ax.scatter(x[y_clusters == 3,0],x[y_clusters == 3,1],x[
y_clusters == 3,2], s = 50 , color = 'red', label = "Cl
uster 4")
ax.scatter(x[y_clusters == 4,0],x[y_clusters == 4,1],x[
y_clusters == 4,2], s = 50 , color = 'purple', label =
"Cluster 5")
ax.scatter(model.cluster_centers[:,0],model.cluster_ce
nters[:,1],model.cluster_centers[:,2], s = 150, c = "
yellow", label = "Centroids")
ax.set_xlabel('Tempat Kejadian')
ax.set_ylabel('Jenis Kasus')
ax.set_zlabel('Jumlah Kasus')
ax.legend()
plt.show()
```

Biodata

Nama : Nancy M. Hilipito

NIM : T3117071

Jenis Kelamin : Perempuan

Tempat Tanggal Lahir: Manado, 07 November 1999

Agama : Islam

Suku Bangsa : Indonesia

Alamat : JL. Sirsak, Duingi Kota Gorontalo

Fakultas : Ilmu Komputer

Program Studi : Teknik Informatika

Jenjang Pendidikan : S1 (Sastra Satu)

IPK : 3,03

Judul Skripsi : Clustering Kasus Kepolisian Dengan Metode K-Means



Surat Keterangan Bebas Plagiasi