

**PENERAPAN METODE *K-MEANS CLUSTERING* TERHADAP DAERAH
RAWAN KRIMINALITAS DI INDONESIA PADA TAHUN 2017**

TUGAS AKHIR



Novita Wulandari

13 611 062

**JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA
2020**

**HALAMAN PERSETUJUAN PEMBIMBING
TUGAS AKHIR**

Judul : Penerapan Metode *k-means clustering* terhadap daerah rawan kriminalitas di Indonesia pada tahun 2017.

Nama Mahasiswa : Novita Wulandari

Nomor Mahasiswa 13611062

**TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK
DIUJIKAN**

Yogyakarta, 14 September 2020

Pembimbing



(Rahmadi Yotenka, S.Si., M.Sc)

HALAMAN PENGESAHAN
TUGAS AKHIR

Penerapan Metode *k-means clustering* Terhadap Daerah Rawan Kriminalitas di
Indonesia pada tahun 2017.

Nama Mahasiswa : Novita Wulandari
Nomor Mahiswa : 13611062

TUGAS AKHIR INI TELAH DIUJIKAN
PADA TANGGAL 14 SEPTEMBER 2020

Nama Penguji

Tanda Tangan

1. Muhammad Hasan sidiq Kurniawan, S.Si., M.Sc.
2. Muhammad Muhajir, S.Si., M.Sc.
3. Rahmadi Yotenka, S.Si., M.Sc.



Mengetahui

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



(Prof. Riyanto, S.Pd., M.,Si., Ph.D.)



Saya yang bertanda tangan dibawah ini :

Nama : Novita Wulandari
NIM : 13611062
Program Studi : Statistika
Fakultas : MIPA
Judul Skripsi : PENERAPAN METODE *K-MEANS CLUSTERING* TERHADAP DAERAH
RAWAN KRIMINALITAS DI INDONESIA PADA TAHUN 2017

Menyatakan dengan sebenarnya bahwa karya ilmiah tersebut adalah benar hasil karya sendiri, kecuali kutipan yang sudah saya sebutkan sumbernya, belum pernah di ajukan pada institusi manapun, dan bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian saya buat pernyataan ini dengan sebenarnya, tanpa ada tekanan dan paksaan dari pihak mana pun serta bersedia mendapatkan sanksi akademik jika ternyata dikemudian hari pernyataan ini tidak benar.

Sleman, 14 Oktober 2020

yang menyatakan,

Novita Wulandari
NIM 13611062



KATA PENGANTAR



Assalamu'alaikum Wr.Wb

Alhamdulillahirabbil'alamin. Puji dan syukur yang tidak ada hentinya kepada Allah SWT yang telah banyak melimpahkan nikmat, rahmat, dan karunia-Nya yang tidak terhingga kepada umat-Nya. Sholawat serta salam tak lupa dicurahkan kepada Nabi Muhammad SAW dan tak lupa juga doa kepada keluarganya, sahabatnya, dan para pengikutnya hingga akhir zaman.

Penyusunan Tugas Akhir yang berjudul “Penerapan Metode *K-Means Clustering* Terhadap Daerah Rawan Kriminalitas di Indonesia Pada Tahun 2017” merupakan salah satu persyaratan yang harus dipenuhi dalam menyelesaikan jenjang strata satu di Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia.

Penyelesaian Tugas Akhir ini tidak terlepas dari bantuan, pengarahan, dan bimbingan berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof Fathul Wahid, ST., M.Sc., Ph.D. selaku Rektor Universitas Islam Indonesia, Yogyakarta.
2. Bapak Prof. Riyanto, S.Pd., M.,Si., Ph.D. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia, Yogyakarta.
3. Bapak Dr. Edy Widodo, S.Si., M.Si. selaku Ketua Program Studi Statistika beserta seluruh jajarannya
4. Bapak Rahmadi Yotenka, S.Si., M.Sc selaku dosen pembimbing yang telah sabar membimbing dan memberi arahan kepada penulis selama penyusunan Tugas Akhir ini.
5. Dosen-dosen pengampu mata kuliah yang telah mengajarkan ilmu- ilmu, sehingga dapat penulis terapkan pada tugas akhir ini.
6. Segenap Staff/Karyawan di FMIPA UII yang telah membantu segala keperluan penulis dalam perkuliahan dan keperluan administrasi.

7. Ayah, Ibu dan Adik tercinta yang selalu mendoakan dengan tulus, memberikan dorongan, semangat, motivasi dan kasih sayang yang luar biasa untuk penulis.
8. Agi Dempita yang telah memberi semangat dan bantuan kepada penulis dalam proses pembuatan Tugas Akhir ini.
9. Teman-teman kos dan seluruh teman seperjuangan di Statistika UII angkatan 2013 kalianlah yang selalu memberi inspirasi.
10. Teman – teman bimbingan dan teman-teman statistika 2013 terima kasih atas kebersamaannya selama ini. Semoga kelak menjadi Statistisi yang amanah dan bermanfaat.
11. Semua pihak yang tidak dapat penulis sebutkan satu per satu, terima kasih.

Penulis menyadari sepenuhnya bahwa tugas akhir ini masih jauh dari sempurna, oleh karena itu segala kritik dan saran yang sifatnya membangun selalu penulis harapkan. Semoga Tugas Akhir ini dapat bermanfaat bagi penulis khususnya dan bagi semua yang membutuhkan umumnya. Akhir kata, semoga Allah SWT selalu melimpahkan rahmat serta hidayah-Nya kepada kita semua, Aami n aami n yaa Robbal ‘alaamiin.

Wassalamu’alaikum, Wr.Wb

Yogyakarta, 14 September 2020

Novita Wulandari

DAFTAR ISI

PENERAPAN METODE <i>K-MEANS CLUSTERING</i> TERHADAP DAERAH RAWAN KRIMINALITAS DI INDONESIA PADA TAHUN 2017	i
HALAMAN PERSETUJUAN PEMBIMBING	ii
HALAMAN PENGESAHAN TUGAS AKHIR	iii
KATA PENGANTAR	iv
DAFTAR ISI	vi
DAFTAR TABEL	viii
DAFTAR GAMBAR	ix
DAFTAR LAMPIRAN	x
INTISARI	xi
ABSTRACT	xii
BAB I	1
PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian	3
1.4. Batasan Masalah	3
1.5. Manfaat Penelitian	3
1.6. Sistematika Penulisan	4
BAB II	5
TINJAUAN PUSTAKA	5
2.1. Penelitian Terdahulu	5
BAB III	8
LANDASAN TEORI	8
3.1. Kriminalitas	8
3.2. Kejahatan Terhadap Fisik/ Badan (<i>Violence</i>)	9
3.3. Kejahatan Terhadap Hak Milik/ barang penggunaan kekerasan dan tanpa kekerasan	10

3.4.	Kejahatan Terkait Narkotika	10
3.5.	Kejahatan Terkait Penipuan, Penggelapan dan Korupsi	11
3.6.	Analisis Deskriptif	11
3.7.	Principal Components Analysis (PCA).....	13
3.8.	Analisis <i>K-Means Clustering</i>	13
3.9.	Pemetaan	17
3.10.	Analisis Korelasi.....	18
3.11.	Pencilan	19
BAB IV	20
METODOLOGI PENELITIAN	20
4.1.	Populasi dan Sampel	20
4.2.	Metode Pengambilan Data	20
4.3.	Variabel Penelitian.....	20
4.4.	Alat dan Metode Analisis Data	22
4.5.	Diagram Alur Penelitian.....	23
HASIL DAN PEMBAHASAN	24
5.1.	Korelasi.....	24
5.2.	Uji KMO, MSA, Barlett's.....	25
5.3.	Pengelompokan Kriminalitas	26
5.4.	Pengelompokan Daerah Kriminalitas.....	28
5.5.	Pemetaan.....	36
BAB VI	38
KESIMPULAN	38
6.1.	Kesimpulan	38
6.2.	Saran	38
DAFTAR PUSTAKA	40
LAMPIRAN		

DAFTAR TABEL

Tabel 3.1 Data Body Mass Index (BMI).....	16
Tabel 4.1 Tabel Variabel Penelitian	20
Tabel 5.2 Tabel MSA	26
Tabel 5.2 Pengelompokan Kriminalitas	26
Tabel 5.3 Parameter Faktor.....	27
Tabel 5.4 Titik Pusat Cluster Awal.....	28
Tabel 5.5 Hasil Pengelompokan.....	28
Tabel 5.6 Titik Pusat Cluster Baru	29
Tabel 5.7 Hasil Pengelompokan Cluster Baru.....	30
Tabel 5.8 Pengelompokan Daerah Kriminalitas	31
Tabel 5.9 Final Cluster.....	34



DAFTAR GAMBAR

Gambar 4.1 Diagram Alur Penelitian	24
Gambar 5.1 Korelasi Antar Variabel	25
Gambar 5.2 Uji KMO dan Barlett's	26
Gambar 5.3 Grafik Metode <i>Gap Statistics</i>	28
Gambar 5.4 Pengelompokan Daerah Kriminalitas	33
Gambar 5.5 Peta Daerah Rawan Kriminalitas	37



DAFTAR LAMPIRAN

Lampiran 1. Data Penelitian

Lampiran 2. *Syntax*



PENERAPAN METODE K-MEANS CLUSTERING TERHADAP DAERAH RAWAN KRIMINALITAS DI INDONESIA PADA TAHUN 2017

Novita Wulandari

Program Studi Statistika Fakultas MIPA

Universitas Islam Indonesia

INTISARI

Kriminalitas atau kejahatan merupakan suatu perbuatan yang dapat mengakibatkan timbulnya masalah dalam masyarakat dan juga keresahan bagi kehidupan masyarakat. Di Indonesia jumlah kriminalitas meningkat tiap tahun. Tujuan dari penelitian ini adalah melakukan penerapan metode *k-means clustering* terhadap daerah yang rawan kriminalitas di Indonesia. Variabel yang digunakan adalah penganiayaan berat, penganiayaan ringan, pencurian, pencurian dengan pemberatan, pencurian kendaraan bermotor, kejahatan NAPZA, penipuan, penggelapan. Hasil dari penelitian ini kemudian akan dibagi kedalam 3 kelompok dimana kelompok I masuk kedalam kategori wilayah Aman, kelompok ke II masuk kedalam daerah rawan, dan kelompok III masuk kedalam sangat rawan. Pemetaan kemudian didasarkan pada hasil analisis kelompok, daerah yang sangat rawan kriminalitas akan digambarkan dengan warna yang gelap, sedangkan untuk daerah dengan tingkat kriminalitas yang rendah akan digambarkan dengan warna yang terang.

Kata kunci : *K-means clustering*, Kriminalitas, Pemetaan.

IMPLEMENTATION OF *K-MEANS CLUSTERING* METHOD TO PRONE CRIMINALITY AREAS IN INDONESIA ON 2017

Novita Wulandari

Departement of Statistics Faculty of Mathematics and Natural Sciences
Islamic University of Indonesia

ABSTRACT

Criminality is an act that can cause problems in society and also unrest in people's lives. In Indonesia the number of crimes increases every year. The purpose of this study is to apply *K-Means Clustering* method to crime-prone areas in Indonesia. The variables used are severe maltreatment, light maltreatment, theft, weighted theft, motor vehicle theft, narcotics crime pschotropic addictive substance, fraud and embezzlement. The research serult will then be divided into 3 groups. Where group I is included in the safe area category, group II is included in a vulnerable area, and group 3 is included in a very vulnerable area. The mapping is then based on the result of group analysis, areas that are very prone to crime will be depicted in dark colors, while areas with low crime rate will be described in light colors.

Keywords ; *K-Means clustering*, Crime, Mapping

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kejahatan atau tindak kriminalitas merupakan salah satu bentuk dari perilaku menyimpang yang selalu ada dan melekat pada masyarakat. Kriminalitas adalah tindakan kejahatan yang dilakukan secara tidak sadar baik oleh wanita ataupun pria yang merugikan orang lain. Kriminalitas bukanlah warisan atau bawaan sejak lahir (Kartono,1999).

Rasa aman merupakan hak setiap penduduk Indonesia, sebagaimana telah diatur dalam UUD Republik Indonesia 1945 pasal 28G ayat 1. Pasal ini menyebutkan bahwasanya “Setiap orang berhak atas perlindungan diri, pribadi, keluarga, kehormatan, martabat dan harta benda yang dibawah kekuasaannya, serta berhak atas rasa aman dan perlindungan dari ancaman ketakutan untuk berbuat atau tidak berbuat sesuatu yang merupakan hak asasi”. Hal ini kemudian membuktikan bahwa rasa aman merupakan hal yang penting dan krusial(BPS,2016).

Permasalahan kriminalitas di Indonesia adalah permasalahan yang pelik yang kemudian berdampak luas keseluruh lapisan masyarakat. Banyak jenis tindak kriminalitas yang terjadi di Indonesia yang kemudian menjadi permasalahan umum diberbagai wilayah dan waktu yang berbeda,sehingga menyebabkan sulitnya menentukan provinsi mana saja yang memiliki tingkat kerawanan tindak kriminalitas. Informasi tentang pemetaan wilayah dengan tingkat kriminalitas yang cukup tinggi sangat dibutuhkan khususnya oleh masyarakat dan penegak hukum dalam hal ini jajaran kepolisian. Informasi ini kemudian akan berguna khususnya masyarakat luas untuk melakukan tindak antisipasi terlebih bagi kepolisian dalam mengambil keputusan apakah suatu daerah memerlukan pengawasan ekstra atau tidak.

Dari tahun ketahun kasus kriminalitas dimungkinkan akan bertambah banyak, dengan data kriminalitas yang semakin bertambah tersebut proses klasifikasi daerah rawan kriminalitas tersebut menjadi kurang akurat, sehingga

dalam hal ini metode *k-means* dibutuhkan untuk memecahkan masalah tersebut. Metode *k-means* digunakan karena metode tersebut termasuk dalam metode yang cukup sederhana dan cepat dalam melakukan proses klasifikasi. Dengan metode tersebut penentuan provinsi mana saja yang rawan terjadi kriminalitas yang cukup tinggi akan lebih mudah dan juga optimal. Jenis data yang digunakan adalah data kuantitatif. Data kuantitatif sendiri merupakan jenis data berupa angka atau nominal sehingga data tersebut dapat diolah menggunakan metode *k-means*. Kategori kriminalitas yang digunakan sebagai parameter antara lain : penganiayaan berat, penganiayaan ringan, pencurian, pencurian dengan pemberatan, pencurian kendaraan bermotor, kejahatan napza, penipuan, penggelapan. 8 Variabel yang digunakan merupakan variabel yang jumlah kasus kriminalitas paling tinggi pada tahun 2017. Data tersebut bersumber dari publikasi Statistika Kriminal 2018 yang dikeluarkan oleh Badan Pusat Statistik. Dari publikasi tersebut menghasilkan gambaran umum mengenai tingkat kriminalitas di Indonesia selama periode 2015 – 2017. Oleh karena itu peneliti hanya mengambil 8 sampel masalah kriminalitas tertinggi pada tahun 2017 dan kemudian dilakukan pengelompokan tingkat kriminalitas di berbagai provinsi di Indonesia dengan menggunakan cara mencari nilai *k* terbaik menggunakan metode *gapstatistics*. Metode ini sudah cukup lama digunakan, dimana metode ini akan melihat fungsi dari nilai cluster pada suatu data.

Oleh karena itu penulis melakukan klasifikasi dan pemetaan menggunakan metode *k-means clustering* dan aplikasi *Qgis* ini diharapkan pihak kepolisian khususnya Kepolisian Republik Indonesia dapat mengklasifikasikan provinsi yang rawan kriminalitas dan memungkinkan untuk pihak Kepolisian dalam hal itu Polda, Polres, atau Polsek diprovinsi tersebut untuk lebih meningkatkan keamanan, pencegahan dan juga pemberantasan terhadap kriminalitas tersebut agar masyarakat lebih waspada lagi terhadap potensi kriminalitas yang dapat terjadi diwilayah masing-masing.

12 Rumusan Masalah

Dari latar belakang diatas, maka dapat dirumuskan masalahnya adalah bagaimana klasifikasi kriminalitas dan pemetaan terhadap daerah/provinsi di Indonesia dengan menggunakan metode *k-means clustering*.

13 Tujuan Penelitian

Dalam penyusunan tugas akhir ini sesuai dengan rumusan masalah yang telah diuraikan diatas, maka tujuan dari penelitian ini adalah Mengetahui pemetaan klusterisasi data kriminalitas di Indonesia yang kemudian diolah menggunakan metode *k-means clustering*.

14 Batasan Masalah

Agar bahasan dalam permasalahan penelitian ini tidak semakin meluas, maka peneliti memberikan batasan batasan masalah sebagai berikut :

1. Data yang digunakan peneliti adalah data sekunder yang diperoleh dari Badan Pusat Statistika Indonesia.
2. Data yang akan diolah merupakan data jumlah kejahatan menurut Kepolisian Daerah pada tahun 2017 yaitu data Penganiayaan Berat, Penganiayaan Ringan, Pencurian, Pencurian dengan Pemberatan, Pencurian Kendaraan Bermotor, Kejahatan NAPZA, Penipuan, Penggelapan.
3. Metode analisis yang digunakan adalah *k-means clustering*.
4. Data diolah menggunakan bantuan dari aplikasi *R-studio* dan Pemetaan menggunakan aplikasi *qgis*.

15 Manfaat Penelitian

Ada beberapa manfaat yang diharapkan dalam penelitian ini yaitu :

1. Bagi peneliti, lebih memahami dan mengetahui secara sistematis tentang penggunaan dari metode analisis *k-means clustering*.
2. Bagi para pembaca, diharapkan dengan adanya penelitian ini menambah wawasan dan menambah informasi berupa peta tingkat kriminalitas agar

pembaca dapat lebih waspada dan juga berhati hati guna mengantisipasi dan menghindari kasus kriminalitas yang terjadi disekitar kita.

3. Untuk Kepolisian Republik Indonesia diharapkan dapat membantu sebagai acuan dalam memberikan rasa aman secara lebih terhadap daerah daerah yang memiliki tingkat kriminalitas yang tinggi.

16. Sistematika Penulisan

Sistematika penulisan tugas akhir ini sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menjelaskan latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini pemaparan mengenai penelitian yang terdahulu yang kemudian topik dan bahasan berhubungan dalam menganalisa dan juga memecahkan masalah.

BAB III LANDASAN TEORI

Pada bab ini lebih menjelaskan tentang teori dasar yang relevan dengan teori statistika guna mendukung penulisan tugas akhir seperti algoritma kmeans pemetaan menggunakan qgis

BAB IV METODOLOGI PENELITIAN

Pada bab ini berisi variable digunakan dalam penelitian serta tahap yang digunakan sebagai kajian untuk penelitian. Tahapan tersebut merupakan rangka sebagai pedoman untuk mendapatkan tujuan yang telah ditetapkan oleh peneliti.

BAB V HASIL DAN PEMBAHASAN

Pada bab ini berisi uraian mengenai gambaran pemetaan tingkat kriminalitas di Indonesia.

BAB VI KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan dan saran dari keseluruhan kegiatan penelitian yang telah dilakukan.

DAFTAR PUSTAKA

BAB II

TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Pada menyusun tugas akhir ini sangat diperlukan beberapa acuan dari penelitian penelitian sebelumnya. Berikut ini adalah beberapa penelitian yang digunakan dalam tinjauan pustaka dalam penelitian ini :

1. Dermawanti, dkk (2015) melakukan penelitian yang berjudul “*Faktor-Faktor Yang Mempengaruhi Kriminalitas Di Kabupaten Batang Pada Tahun 2013 Dengan Analisis Jalur*” Variabel- variabel Yang di gunakan dalam Penelitian ini diantaranya jumlah penduduk, pendidikan moral, kemiskinan, pengangguran dan kriminalitas di Kabupaten Batang tahun 2013. Hasil penelitian menunjukkan faktor yang memiliki pengaruh besar terhadap kriminalitas adalah pengangguran. Faktor terbesar kedua adalah pendidikan, dan faktor yang memiliki pengaruh positif terkecil adalah faktor moral.
2. Edmira Rivani (2010), melakukan penelitian yang berjudul “*Aplikasi K-Means Cluster untuk Pengelompokan Berdasarkan Produksi Padi, Jagung, Kedelai, dan Kacang Hijau Tahun 2009*” dengan menggunakan metode analisis *K-Means cluster* pada data produksi tanaman pangan utama yaitu padi, jagung, kedelai, dan kacang hijau. Hasil penelitian didapatkan bahwa pada tahun 2009 provinsi pada klaster 3 (Jawa Barat, Jawa Tengah, dan Jawa Timur) merupakan provinsi dengan jumlah produksi terbanyak atau dapat dikatakan provinsi pada klaster 3 merupakan daerah produksi tinggi, sementara provinsi pada klaster 2 (Sumatera Utara, Sumatera Selatan, Lampung, dan Sulawesi Selatan) merupakan daerah produksi menengah, dan provinsi pada klaster 1 (Nanggroe Aceh Darussalam, Sumatera Barat, Riau, Jambi, Bengkulu, Bangka Belitung, Kepulauan Riau, DKI Jakarta, DI Yogyakarta, dan Banten) merupakan daerah produksi rendah karena

memproduksi padi, jagung, kedelai, dan kacang hijau dalam jumlah paling sedikit.

3. Aan Rosiatun (2010), melakukan penelitian yang berjudul “Analisis Klaster untuk Segmentasi Pemirsa Program Berita Sore Stasiun TV Swasta” dengan menggunakan metode analisis *K-Means cluster* pada data kuisisioner penilai dengan 12 variabel yaitu kekritisn berita, narasumber yang terpercaya, keanekaragaman berita, dinikmati oleh semua kalangan, jam tayang yang sesuai, kedalaman isi berita, bahasa yang mudah dimengerti, keaktualan berita, berita dapat dipercaya, berita disajikan secara menarik, penampilan penyaji berita dan durasi program. Hasil penelitian didapatkan bahwa pemirsa berita sore yang berusia lebih dari 20 tahun sebesar 40,3 %, yang berjenis kelamin perempuan sebesar 32,4%, yang berpendidikan SMA sebesar 18,7%, yang bekerja sebagai ibu rumah tangga sebesar 21,6% memberikan penilaian yang sedang ditonton.
4. Yanne Flowrensia (2010), melakukan penelitian yang berjudul “Perbandingan Penggerombolan *K-Means* dan *K-Medoid* pada Data yang Mengandung Pencilan” dengan menggunakan metode analisis gerombol *K-Means* dan *K-medoid* pada data pecilan tanaman bunga iris dengan variabel yaitu Varietas Setosa, Versicolor, dan Viginica. Hasil penelitian didapatkan bahwa metode *kmedoid* mempunyai nilai rata-ran tingkat salah klasifikasi yang lebih rendah dan signifikan pada kondisi proporsi pencilan 5%, sedangkan pada kondisi proporsi pencilan 10% dan 15% hasil nilai rata-ran salah klasifikasinya tidak berbeda signifikan dengan metode *K-Means*.
5. Resti Hidayah (2014), melakukan penelitian yang berjudul “Sistem Informasi Berbasis WEB untuk Produksi Kedelai di Indonesia Menggunakan *K-Means Clustering*” dengan menggunakan metode analisis *K-Means cluster* pada data produksi kedelai. Hasil penelitian didapatkan bahwa *clustering* pada data produksi kedelai diperoleh dengan jumlah 4 kelompok dan menghasilkan *sum square error* (SSE) sebesar 2,07, 1,53, dan 0,0037 berturut-turut untuk data luas panen, produksi, dan produktivitas.

6. Ulfa Khaira (2012), melakukan penelitian yang berjudul “Integrasi Self Organizing Maps dan Algoritme *K-Means* untuk *Clustering* Data Ketahanan Pangan Kabupaten di Wilayah Bali, Nusa Tenggara Barat, dan Nusa Tenggara Timur” dengan menggunakan metode algoritme *K-Means* pada data ketahanan pangan. Hasil penelitian didapatkan bahwa validasi dengan DBI sebesar 1,635. Jika dibandingkan nilai DBI hasil *clustering* algoritme SOM, S-K, dan *K-Means*, algoritme S-K memiliki nilai DBI yang minimal.
7. Neli Soraya (2009), melakukan penelitian yang berjudul “Analisis Gerombol *K-Means* dan *K-error* untuk Pengelompokan Desa-desa di Jawa Barat berdasarkan Indikator Kemiskinan” dengan menggunakan metode Analisis *K-Means* cluster pada data kemiskinan SUSENAS tahun 2005. Hasil dari penelitian didapatkan bahwa ada 4 klaster, desa-desa yang memiliki jumlah penduduk miskin tertinggi ada di klaster 2 dengan rata-rata *K-Means* jarak relatif lebih besar dibandingkan rata-rata K-Error. Kesimpulan dengan memasukkan galat dalam pengelompokan dapat meminimalkan rata-rata jarak observasi ke pusat klaster, sehingga metode *K-Error* dapat menghasilkan klaster yang lebih baik.

Berdasarkan penelitian terdahulu, peneliti akan melakukan penelitian yang berjudul “Penerapan Metode *K-means clustering* terhadap daerah rawan kriminalitas Di Indonesia pada Tahun 2017” dalam penelitian ini mempunyai kesamaan dalam menggunakan metode yaitu analisis *K-Means cluster* dan terdapat perbedaan pada data dimana menggunakan data kriminalitas di Indonesia pada tahun 2017 dengan 8 variabel yaitu, pencurian dengan pemberatan, kejahatan NAPZA, pencurian kendaraan bermotor, pencurian, penipuan, penganiayaan, penganiayaan berat. Penelitian ini akan menggunakan aplikasi di Software *r-studio* dimana hasil dari penelitian ini akan di perjelas pada bab selanjutnya.

BAB III

LANDASAN TEORI

3.1. Kriminalitas

Tindak kejahatan/kriminalitas atau pelanggaran merupakan perbuatan seseorang yang dapat diancam hukuman berdasarkan KUHP atau undang-undang serta peraturan lainnya yang berlaku di Indonesia. Kriminalitas merupakan segala macam bentuk tindakan dan perbuatan yang merugikan secara ekonomis dan psikologis yang melanggar hukum yang berlaku dalam negara Indonesia serta norma-norma sosial dan agama. Dapat diartikan bahwa, tindak kriminalitas adalah segala sesuatu perbuatan yang melanggar hukum dan melanggar norma-norma sosial, sehingga masyarakat menentangnya (Kartono, 1999). Secara sosiologis, tindak kriminalitas merupakan suatu pola tingkah laku yang merugikan masyarakat dan suatu pola tingkah laku yang mendapatkan reaksi sosial dari masyarakat. Reaksi sosial tersebut dapat berupa reaksi formal, reaksi informal, dan reaksi non-formal. Pengertian kejahatan sebagai unsur dalam pengertian kriminalitas, secara sosiologis mempunyai dua unsur, yaitu :

1. Kejahatan itu ialah perbuatan yang merugikan secara ekonomis dan merugikan secara psikologis.
2. Melukai perasaan susila dari suatu segerombolan manusia, dimana orang-orang itu berhak melahirkan celaan.

Banyak faktor yang dapat menyebabkan seseorang melakukan tindak kriminal. Menurut Hamzah, A (1986), faktor penyebab kriminalitas dikelompokkan menjadi faktor dari dalam diri dan faktor dari luar diri pelaku. Faktor dari dalam diri tersebut yaitu faktor biologis, faktor pembawaan kriminal, faktor umur dan kualitas manusia itu sendiri. Faktor dari luar diri yaitu faktor lingkungan, laju pertumbuhan penduduk, kemiskinan, bacaan harian atau film.

Salah satu indikator kriminalitas yang data nya dapat dihimpun oleh kepolisian adalah risiko penduduk terkena tindak kejahatan. Risiko penduduk terkena tindak kejahatan (*crime rate*) adalah angka yang mengindikasikan

banyaknya penduduk berisiko terkena tindak pidana per seratus ribu penduduk, Indikator ini mengindikasikan risiko penduduk terkena tindak pidana pada suatu waktu tertentu (BPS, 2016). Rumus untuk menghitung *crime rate* adalah sebagai berikut:

$$crime\ rate = \frac{jumlah\ peristiwa\ kejahatan\ pada\ tahun\ t}{jumlah\ penduduk} \times 100.000 \quad (3.1)$$

Contoh 3.1: Pada tahun 2017, penduduk Indonesia sebesar 261.890.872 jiwa dengan jumlah peristiwa kejahatan pada tahun tersebut 336.652 peristiwa. Maka *crime rate* dapat dihitung sebagai berikut:

$$crime\ rate = \frac{336.652}{261.890.872} \times 100.000 = 129$$

Artinya jumlah orang yang berisiko terkena tindak kejahatan pada 100.000 penduduk adalah 129 orang.

32 Kejahatan Terhadap Fisik/ Badan (*Violence*)

Kejahatan terhadap fisik atau badan merupakan tindak kejahatan yang sulit untuk dihilangkan dalam kehidupan bermasyarakat. Berbagai macam tindak kejahatan terhadap fisik dan nyawa atau yang biasa dikenal dengan penganiayaan dan pembunuhan seringkali mengakibatkan luka pada bagian tubuh, pemukulan fisik dan kekerasan lainnya bahkan menyebabkan korban cacat fisik seumur hidup hingga kematian.

Kejahatan fisik/Badan dapat dijumpai dilingkungan rumah tangga, atau tempat umum ataupun tempat mana saja dan bisa menimpa siapa saja. Hal tersebut dapat terjadi diduga dari beberapa faktor yaitu pengaruh pergaulan dan kenalakan, premanisme, kecemburuan sosial, tekanan, kesenjangan ekonomi, serta ketidak harmonisan dalam rumah tangga, persaingan ataupun konflik lainnya.

Dalam hal ini Penganiayaan Berat dan Penganiayaan Ringan menjadi masalah yang cukup tinggi di Indonesia. Pada tahun 2017 jumlah kejadian kejahatan terhadap fisik/ badan cenderung menurun menjadi 42.683 kejadian sedangkan pada tahun 2016 jumlah kejadian mencapai 46.706 kasus.

33. Kejahatan Terhadap Hak Milik/ barang penggunaan kekerasan dan tanpa kekerasan

Kejahatan terhadap hak milik/barang dengan menggunakan kekerasan merupakan suatu tindakan yang menyimpang. Menurut Robert M.Z Lawang (2017) penyimpangan perilaku adalah semua tindakan yang menyimpang dari norma yang berlaku dalam sistem sosial dan menimbulkan usaha dari mereka yang berwenang dalam system itu untuk memperbaiki perilaku menyimpang.

Dalam pasal 362 KUHP dikatakan “pengambilan suatu barang, yang seluruh atau sebagiannya kepunyaan orang lain, dengan maksud untuk dimiliki secara melawan hukum diancam dengan pencurian”. Dengan demikian perampokan juga disebut sebagai pencurian atas suatu barang.

Dalam klasifikasinya kejahatan terhadap hak milik/barang dengan menggunakan kekerasan meliputi Pencurian dengan Kekeraan, Pencurian Menggunakan Senjata Api(Senpi), Pencurian dengan Senjata Tajam(Sajam). Jumlah kejadian terhadap kejahatan hak/milik barang dengan penggunaan kekerasan terbanyak pada tahun 2017 yaitu Polda Sulawesi Selatan sebanyak 1.137 kejadian (BPS, 2018).

Sedangkan untuk kejahatan terhadap hak milik/barang meliputi Pencurian, Pencurian dengan pemberatan, Pencurian Kendaraan Bermotor, Pembakaran dengan sengaja, Pendadahan. Kejahatan terhadap hak milik barang tanpa penggunaan kekerasan pada tahun 2017 yaitu Polda Sumatera Utara sebanyak 12.361 kejadian (BPS, 2018).

34. Kejahatan Terkait Narkotika

Narkotika dapat merusak fisik dan mental tidak hanya disitu dampak buruk terhadap narkotika itu sendiri adalah mempengaruhi kehidupan sosial dalam bermasyarakat, berpotensi menjadi penghambat pembangunan nasional yang dapat mengancam keamanan dan kedaulatan Negara.

Pada tahun 2017 kasus kejahatan terkait Narkotika paling banyak terjadi diwilayah Polda Metro Jaya yaitu 7.214 Kejadian(BPS,2018). Kejahatan Narkotika menurut Undang-Undang Narkotika dikenakan sanksi pidana penjara,

pidana denda, pidana seumur hidup dan sanksi lainnya, Perlindungan bagi anak penyalahgunaan Narkotika ditinjau dari Aspek Viktimologi yaitu direhabilitasi karena anak tersebut disatu sisi menjadi pelaku dan sisi lainnya menjadi korban, Peredaran Narkotika ditinjau dari hukum internasional yaitu kebijakan penanggulangan kejahatan narkotika pada awalnya dituangkan dalam *The United Nation's Single Convention on Narcotic Drugs* 1961.

35. Kejahatan Terkait Penipuan, Penggelapan dan Korupsi

Kejahatan Terkait Penipuan merupakan Tindak pidana yang sering terjadi di lingkungan masyarakat, untuk memenuhi kebutuhan atau keuntungan seseorang dapat melakukan suatu tindak pidana penipuan. Di Indonesia seringkali terjadi tindak pidana penipuan dikarenakan banyak Faktor-faktor yang mendukung terjadinya suatu tindakan penipuan, misalnya karena kemajuan teknologi sehingga dengan mudah melakukan tindakan penipuan, keadaan ekonomi yang kurang sehingga memaksa seseorang untuk melakukan penipuan, terlibat suatu utang dan lain sebagainya.

Kejahatan penipuan di dalam bentuknya yang pokok diatur dalam Pasal 378 KUHP yang berbunyi sebagai berikut: “Barangsiapa dengan maksud untuk menguntungkan diri sendiri atau orang lain secara melawan hukum, dengan memakai nama palsu atau martabat palsu, dengan tipu muslihat, ataupun rangkaian kebohongan, menggerakkan orang lain untuk menyerahkan barang atau sesuatu kepadanya, atau memberikan hutang atau menghapus piutang, diancam dengan pidana penjara paling lama empat tahun “

Pada tahun 2017 kasus kejahatan terkait Penipuan Polda dengan jumlah kejadian terbanyak adalah Polda Metro Jaya sebanyak 6.169 Kejadian (BPS,2018).

36. Analisis Deskriptif

Metode deskriptif adalah suatu metode dalam penelitian status kelompok manusia, suatu objek, suatu set kondisi, suatu sistem pemikiran, ataupun suatu kelas peristiwa pada masa sekarang. Whitney (1960) berpendapat, penelitian

deskriptif adalah pencarian fakta dengan interpretasi yang tepat. Penelitian deskriptif mempelajari masalah- masalah dalam masyarakat, serta tata cara yang berlaku dalam masyarakat serta situasi-situasi tertentu, termasuk tentang hubungan, kegiatan-kegiatan, sikap-sikap, pandangan-pandangan serta proses-proses yang sedang berlangsung dan pengaruh-pengaruh dari suatu fenomena.

Sedangkan menurut M.Iqbal Hasan (2002:7) menjelaskan bahwa statistik deskriptif adalah bagian dari statistika yang mempelajari cara pengumpulan data dan penyajian data sehingga mudah dipahami. Statistika deskriptif hanya berhubungan dengan hal menguraikan atau memberikan keterangan-keterangan mengenai suatu data atau keadaan. Dengan kata statistika deskriptif berfungsi menerangkan keadaan, gejala, atau persoalan. Penarikan kesimpulan pada statistika deskriptif (jika ada) hanya ditujukan pada kumpulan data yang ada..

Sugiyono (2007) Analisis deskriptif adalah statistik yang digunakan untuk menganalisa data dengan cara mendeskripsikan atau menggambarkan data yang telah terkumpul sebagaimana adanya tanpa bermaksud membuat kesimpulan yang berlaku untuk umum atau generalisasi.

3.6.1 Rata-Rata Sampel (*mean*)

Jika suatu sampel berukuran n dengan elemen $x_1, x_2, x_3, \dots, x_n$. Maka mean sampel (\bar{x}) adalah (Walpole & Myers, 1995):

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (3.2)$$

dengan,

x_i : data ke i , untuk $i = 1, 2, 3, \dots, n$

n : jumlah sampel

\bar{x} : rata-rata sampel (*mean*)

3.6.2 Variansi Sampel

Bila $x_1, x_2, x_3, \dots, x_n$. itu merupakan sampel acak yang berukuran n , maka variansi sampel didefinisikan sebagai statistik (Walpole & Mayers, 1995):

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (3.3)$$

Dengan,

x_i : data ke i , untuk $i = 1, 2, 3, \dots, n$

n : jumlah sampel

\bar{x} : rata-rata sampel (*mean*)

37. Principal Components Analysis (PCA)

Principal Components Analysis (PCA) atau Analisis Komponen Utama digunakan untuk mengatasi multikolinearitas dalam data karena mengekstraksi faktor sehingga faktor tidak berkorelasi sesamanya. Sebelum melakukan analisis faktor, dilakukan pengujian terhadap data dengan menggunakan uji *barlett* dan KMO.

a. Uji Barlett

Pengujian ini digunakan untuk melihat apakah variabel yang digunakan berkorelasi dengan variabel lainnya. Pengujian ini dilakukan dengan menggunakan *statistics chi-square*, apabila nilai *chi-square* lebih dari nilai *chi-square* tabel maka variabel mengandung korelasi.

b. Uji Kaiser Mayer Olkin (KMO)

KMO merupakan indeks yang digunakan untuk meneliti ketepatan analisis faktor. Analisis faktor tepat untuk digunakan ketika nilai KMO berada diantara 0,5 hingga 1, apabila kurang dari 0,5 analisis faktor dikatakan tidak tepat.

38. Analisis K-Means Clustering

K-Means Cluster Analysis merupakan salah satu metode *Cluster analysis non-hierarki* yang berusaha untuk mempartisi objek yang ada ke dalam satu atau lebih *cluster* atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu *cluster* yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan

kedalam *cluster* yang lain. Tujuan pengelompokan adalah untuk meminimalkan *objective function* yang di *set* dalam proses *clustering*, yang pada dasarnya berusaha untuk meminimalkan variasi dalam satu *cluster* dan memaksimalkan variasi antar *cluster*.

Pada metode *K-Means clustering* memiliki beberapa kelebihan dan kekurangannya. Kelebihannya yaitu Algoritma *K-Means* dinilai cukup efisien, yang ditunjukkan dengan kompleksitasnya, dengan n adalah banyak obyek data, k adalah jumlah cluster yang dibentuk, dan t banyak iterasinya, Biasanya, nilai k dan t jauh lebih kecil dari pada nilai selain n . Selain itu kelebihan dari Algoritma *K-Means* yaitu mampu mengelompokkan objek besar dan penciran obyek dengan sangat cepat sehingga mempercepat proses pengelompokan..

Sedangkan untuk kelemahan algoritma ini adalah adanya keharusan menentukan banyaknya *cluster* yang akan dibentuk, hanya dapat digunakan dalam data yang *mean*-nya dapat ditentukan, dan tidak mampu menangani data yang mempunyai penyimpangan-penyimpangan (*noisy data dan outlier*). Sehingga dapat dikatakan bahwa kelemahan algoritma *k-means* ini sangat bergantung pada pemilihan nilai awal *centroid* dan juga hanya bisa digunakan pada attribute numeric.

Metode *cluster* ini meliputi *sequential threshold*, *pararel threshold*, dan *optimizing threshold*. *Sequential threshold* melakukan pengelompokan dengan terlebih dahulu memilih satu objek dasar yang akan dijadikan nilai awal *cluster*, kemudian semua *cluster* yang ada dalam jarak terdekat dengan *cluster* ini akan bergabung, lalu dipilih *cluster* kedua dan semua objek yang mempunyai kemiripan dengan *cluster* ini akan digabungkan, demikian seterusnya sehingga terbentuk beberapa *cluster* dengan keseluruhan objek yang terdapat didalamnya. Jika diberikan sekumpulan objek $X = (X_1, X_2, \dots, X_n)$ maka algoritma *K-Means Cluster Analysis* akan mempartisi X dalam k buah *cluster*, setiap *cluster* memiliki *centroid* dari objek-objek dalam *cluster* tersebut. Pada tahap awal algoritma *K-Means Cluster Analysis* dipilih secara acak k buah objek sebagai *centroid*, kemudian jarak antara objek dengan *centroid* dihitung dengan menggunakan jarak *euclidian*, objek ditempatkan dalam *cluster* yang terdekat dihitung dari titik

tengah *cluster*. *Centroid* baru ditetapkan jika semua objek sudah ditempatkan dalam *cluster* terdekat. Proses penentuan *centroid* dan penempatan objek dalam *cluster* diulangi sampai nilai *centroid* konvergen (*centroid* dari semua *cluster* tidak berubah lagi).

Secara umum metode *K-Means Cluster Analysis* menggunakan algoritma sebagai berikut:

1. Tentukan k sebagai jumlah *cluster* yang dibentuk. Untuk menentukan banyaknya *cluster* k dilakukan dengan beberapa pertimbangan seperti pertimbangan teoritis dan konseptual yang mungkin diusulkan untuk menentukan berapa banyak *cluster*.
2. Bangkitkan k *centroid* (titik pusat *cluster*) awal secara *random*. Penentuan *centroid* awal dilakukan secara *random* atau acak dari objek-objek yang tersedia sebanyak k *cluster*, kemudian untuk menghitung *centroid cluster* ke- i berikutnya, digunakan rumus sebagai berikut :

$$v = \frac{\sum_{i=1}^n \bar{x}_i}{n} ; i = 1, 2, 3, \dots, n \quad (1) \quad (3.4)$$

dengan,

v : *centroid* pada *cluster*

\bar{x}_i : objek ke- i

n : banyaknya objek yang menjadi anggota *cluster*

3. Hitung jarak setiap objek ke masing- masing *centroid* dari masing- masing *cluster*. Untuk menghitung jarak antara objek dengan *centroid* dapat menggunakan *Euclidean Distance*.

$$d(x, y) = |\bar{x} - y| = \sqrt{\sum_{i=1}^n (\bar{x}_i - y_i)^2} \quad 1, 2, 3, \dots, n \quad (2) \quad (3.5)$$

dengan,

x_i : objek x ke- i

y_i : daya y ke- i

n : banyak objek

Contoh : Diberikan data *Body Mass Index* (BMI) dan ukuran kerangka 20 orang mahasiswa sebagai berikut,

Tabel 3.1. Data Body Mass Index (BMI)

Mahasiswa ke	Bmi	Ukuran kerangka	Mahasiswa ke	Bmi	Ukuran kerangka
1	22.21	11.64	11	22.94	10.60
2	43.25	8.95	12	26.89	10.44
3	19.71	10.93	13	24.91	10.63
4	21.05	10.38	14	22.99	11.47
5	17.93	12.85	15	26.81	9.17
6	17.72	12.00	16	19.14	12.11
7	18.71	11.53	17	21.09	10.67
8	25.85	9.33	18	18.71	12.36
9	19.15	11.80	19	20.58	10.80
10	18.42	11.20	20	27.66	9.94

Selanjutnya kita mencoba mengelompokkan data tersebut menjadi 3 kelompok. Dengan menggunakan algoritma K-Means, berikut langkah-langkah penyelesaiannya :

1. Menentukan jumlah *cluster*, dimana jumlah *cluster* = 3
2. Menentukan pusat *cluster* secara acak. Disini kita tentukan kita tentukan $c_1 = (20,9)$; $c_2 = (23,10)$; dan $c_3 = (27,11)$.
3. Menghitung jarak setiap data yang ada terhadap setiap pusat *cluster*. Berikut perhitungannya dengan menggunakan persamaan *Euclidean Distance Space* :

- a. Jarak antara data mahasiswa pertama dengan pusat *cluster* pertama

$$c_{11} = \sqrt{(22,21 - 20)^2 + (11,64 - 9)^2} = 4,97$$

- b. Jarak antara data mahasiswa pertama dengan pusat *cluster* ke-dua

$$c_{12} = \sqrt{(22,21 - 23)^2 + (11,64 - 10)^2} = 2,04$$

- c. Jarak antara data mahasiswa pertama dengan pusat *cluster* ke-tiga

$$c_{13} = \sqrt{(22,21 - 27)^2 + (11,64 - 11)^2} = 1,91$$

4. Alokasikan masing-masing objek ke dalam *centroid* yang paling dekat.

Untuk melakukan pengalokasian objek kedalam masing-masing *cluster* pada saat iterasi secara umum dapat dilakukan dengan dua cara yaitu dengan *hard k-means*, dimana secara tegas setiap objek dinyatakan sebagai anggota *cluster* dengan mengukur jarak kedekatan sifatnya terhadap titik pusat *cluster* tersebut.

5. Lakukan iterasi, tentukan posisi *centroid* baru dengan menggunakan persamaan 1.
6. Ulangi langkah 3 jika posisi *centroid* baru tidak sama. Pengecekan konvergensi dilakukan dengan membandingkan matriks *group assignment* pada iterasi sebelumnya dengan matriks *group assignment* pada iterasi yang sedang berjalan. Jika hasilnya sama maka algoritma *k-means cluster analysis* sudah konvergen, tetapi jika berbeda maka belum konvergen sehingga perlu dilakukan iterasi berikutnya.

Pada penerapan metode K-Means Clustering, data yang bisa diolah dalam perhitungan adalah data numerik yang berbentuk angka. Sedangkan data selain angka juga bisa diterapkan tetapi terlebih dahulu harus dilakukan pengkodean untuk mempermudah perhitungan jarak/kesamaan karakteristik yang dimiliki dari setiap objek. Setiap objek dihitung kedekatan jaraknya berdasarkan karakter yang dimiliki dengan pusat cluster yang sudah ditentukan sebelumnya, jarak terkecil antara objek dengan masing-masing cluster merupakan anggota cluster yang terdekat. Setelah jumlah cluster ditentukan, selanjutnya dipilih sebanyak 3 objek secara acak sesuai jumlah cluster yang dibentuk sebagai pusat cluster awal untuk dihitung jarak kedekatannya terhadap semua objek yang ada.

39. Pemetaan

Dalam kamus bahasa Indonesia pemetaan atau visualisasi adalah pengungkapan suatu gagasan atau perasaan dengan menggunakan gambar, tulisan, peta, dan grafik. Sementara itu Spasser (1997:78), mengatakan bahwa “peta adalah alat relasi (*relational tools*) yang menyediakan informasi antar hubungan entitas yang dipetakan.” Menurut Chen dalam Ristiyono (2008: 21) bahwa “peta ilmu pengetahuan menggambarkan suatu hubungan ruang antara batas penelitian dalam bidang kegiatan yang signifikan, juga dimana bidang penelitian itu didistribusikan serta dapat memberikan makna dari hubungan tersebut”.

Peta ilmu pengetahuan dapat menggambarkan dan memberikan makna dari hubungan ruang antara batas penelitian yang bidang kegiatannya signifikan dan

bidang kegiatan tersebut dapat didistribusikan. Peta ilmu pengetahuan tidak hanya merupakan suatu alat yang praktis untuk menyampaikan informasi mengenai aktivitas ilmiah, tetapi juga dapat dijadikan sebagai suatu dasar untuk mengkaji atau memahami aktivitas ilmiah dengan menggambarkannya secara tersusun dan terstruktur. Visualisasi ilmu pengetahuan dapat diwujudkan dalam bentuk peta, sehingga muncullah bidang pemetaan ilmu pengetahuan atau *knowledge mapping*. Pemetaan ilmu pengetahuan dapat dilakukan berdasarkan beberapa cara yang terkait erat dengan subjek dokumen.

3.10. Analisis Korelasi

Analisis korelasi merupakan kumpulan Teknik untuk mengukur hubungan antara dua variabel, sehingga hasil dari analisis korelasi adalah laporan hubungan antara dua variabel (Lind, dkk. 2008).

Nilai atau tanda positif (+) atau negatif (-) menunjukkan arah hubungan antara kedua variabel. Pada nilai positif (+), hubungan antara kedua variabel searah, jadi jika satu variabel naik maka variabel yang lain juga naik. Pada nilai negatif (-), kedua variabel berhubungan terbalik, artinya jika satu variabel naik maka variabel yang lain justru turun (Pratisto, 2004).

Menggunakan prosedur matematik untuk mengukur tingkat hubungan antar variabel tersebut dalam bentuk angka atau indeks koefisien korelasi yang bergerak antar -1 sampai +1. Sebuah nilai yang mendekati -1 menunjukkan hubungan kebalikan atau negative antar variabel (Sugiyono, 2007). Adapun rumus matematis untuk menghitung korelasi adalah :

$$r_{xy} = \frac{N(\sum XY) - (\sum X)(\sum Y)}{\sqrt{\{\sum X - \sum X\}^2 \{\sum Y - \sum Y\}^2}} \quad (3.6)$$

Dimana:

r_{xy} = Koefisien korelasi

X = Variabel Bebas

Y = Variabel terikat

N = Jumlah Populasi

3.11. Pencilan

Pencilan atau outlier adalah suatu pengamatan yang menyimpang cukup jauh dari pengamatan lainnya sehingga menimbulkan kecurigaan bahwa pengamatan tersebut berasal dari distribusi data yang berbeda. (Hawkins dalam Sujatmiko, 2005:4). (Hair, dkk. 1995) juga menyatakan bahwa pencilan adalah data yang muncul memiliki karakteristik unik yang terlihat sangat jauh berbeda dari observasi lainnya dan muncul dalam bentuk nilai ekstrim baik untuk sebuah variabel tunggal ataupun variabel kombinasi. Data pencilan muncul disebabkan karena berbagai kemungkinan antara lain:

1. Kesalahan prosedur dalam memasukan data atau mengkodekan.
2. Karena keadaan yang benar-benar khusus, seperti pandangan responden terhadap sesuatu yang menyimpang
3. Karena ada sesuatu alasan yang tidak diketahui penyebabnya oleh peneliti.
4. Muncul dalam range nilai yang ada, tetapi bila dikombinasi dengan variabel lain menjadi ekstrim (disebut multivariat outliers).

BAB IV

METODOLOGI PENELITIAN

41. Populasi dan Sampel

Populasi dalam penelitian ini merupakan data jumlah kriminalitas dengan unit pengamatan yaitu 33 Provinsi di Indonesia. Sampel yang digunakan adalah Data Kriminalitas pada Tahun 2017.

42. Metode Pengambilan Data

Data yang digunakan dalam penelitian ini adalah Data Kuantitatif dari instansi Badan Pusat Statistika Pusat Indonesia Pada Tahun 2017. Jenis data tersebut merupakan data sekunder yang diperoleh dari buku publikasi Statistika Kriminalitas pada tahun 2018 yang diterbitkan oleh BPS Pusat. Data yang digunakan merupakan data kriminalitas di Indonesia.

43. Variabel Penelitian

Variabel yang digunakan merupakan jumlah kriminalitas di Indonesia pada tahun 2017. Adapun variabel – variabel tersebut adalah sebagai berikut :

Tabel 4.1 Tabel Variabel Penelitian

No	Variabel	Kode	Definisi
1	Pencurian dengan Pemberatan	X1	Pencurian biasa yang dalam pelaksanaannya disertai oleh keadaan tertentu yang memberatkan.
2	Kejahatan NAPZA	X2	Kejahatan yang mengacu pada sekelompok zat yang mempunyai resiko kecanduan atau adiksi. Narkotika dan Psikotropika itulah yang secara umum biasa di kenal dengan Narkoba atau NAPZA
3	Pencurian Kendaraan Bermotor	X3	Kejahatan pencurian kendaraan bermotor tergolong ke dalam kejahatan

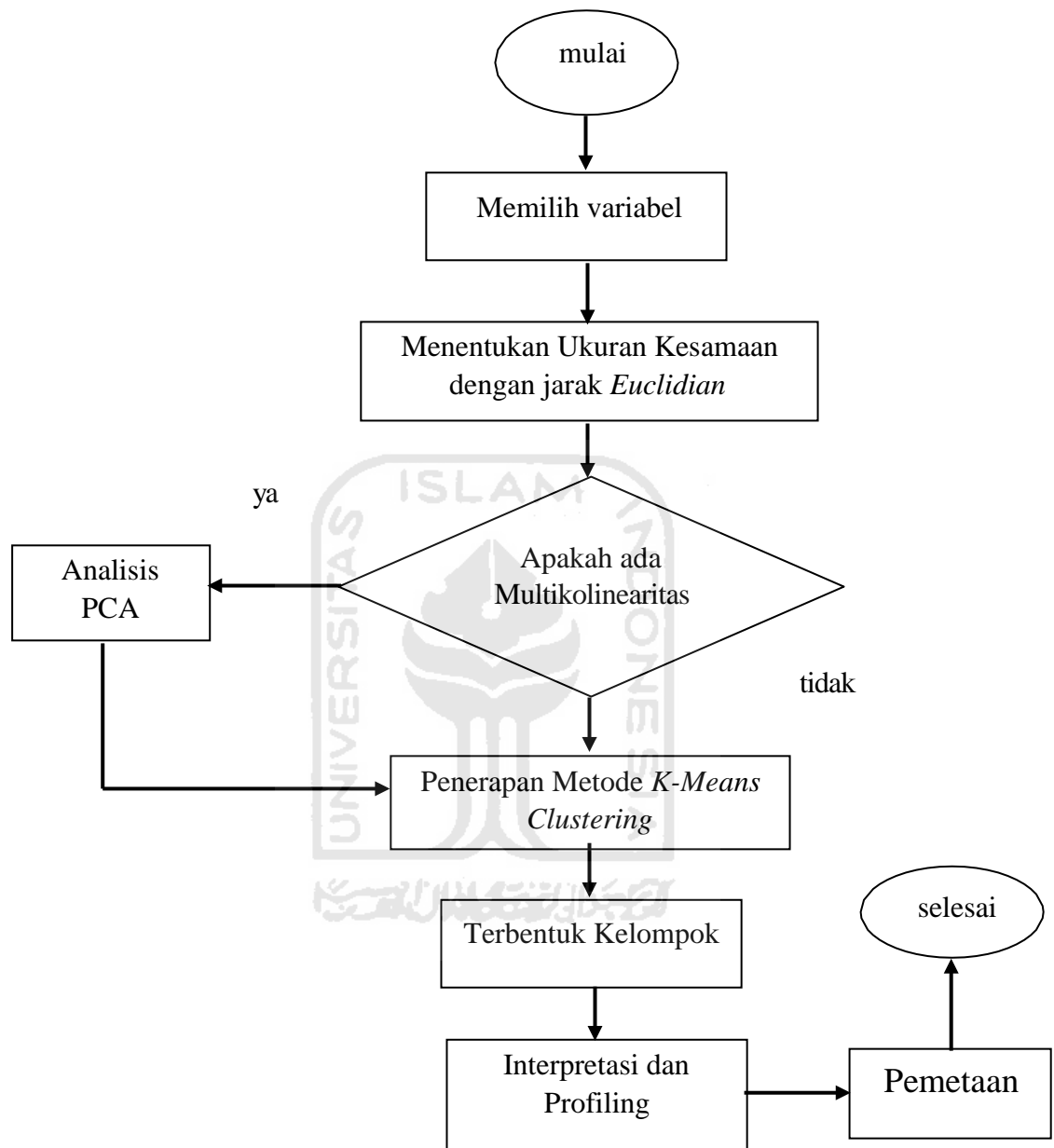
	(CURANMOR)		yang disebabkan karena faktor sosial dan pelakunya melakukan kejahatan karena kebiasaan, kesempatan, kesulitan ekonomi dan juga orang-orang banyak yang turut serta dalam kejahatan kelompok.
4	Pencurian	X4	Pencurian termasuk mengambil barang sesuatu, yang seluruhnya atau sebagian kepunyaan orang lain, dengan maksud untuk dimiliki secara melawan hukum.
5	Penipuan	X5	Penipuan merupakan sebuah kebohongan yang dibuat untuk keuntungan pribadi yang merugikan orang lain. Meskipun memiliki arti hukum yang lebih dalam, detail tentang penipuan bervariasi di berbagai wilayah hukum.
6	Penganiayaan Ringan	X6	Merupakan penganiayaan yang tidak menjadikan sakit atau menjadikan terhalang untuk melakukan jabatan atau pekerjaan sehari-hari.
7	Penggelapan	X7	Penggelapan merupakan suatu tindakan tidak jujur dengan menyembunyikan barang/harta orang lain oleh satu orang atau lebih tanpa sepengetahuan pemilik barang dengan tujuan untuk mengalih-milik, menguasai.
8	Penganiayaan Berat	X8	Penganiayaan berat berarti penganiayaan yang mengakibatkan luka berat.

44 Alat dan Metode Analisis Data

Dalam penelitian ini menggunakan beberapa alat bantu yang digunakan untuk mencapai tujuan dalam analisis ini yaitu perangkat lunak *software r-studio*, *software Microsoft Excel* dan Aplikasi *Qgis*. Metode yang digunakan merupakan analisis *cluster* dengan menggunakan metode *K- Means*, dimana metode *K- Means* merupakan salah satu metode *cluster* non- hirarki. Dari hasil analisis *cluster* dengan metode *K- Means* ini kemudian akan membentuk kelompok-kelompok dengan karakteristik tiap Provinsi untuk mendapatkan penyebaran pemetaan sesuai hasil yang sesuai.



45. Diagram Alur Penelitian



Gambar 4.1. Diagram Alur Penelitian

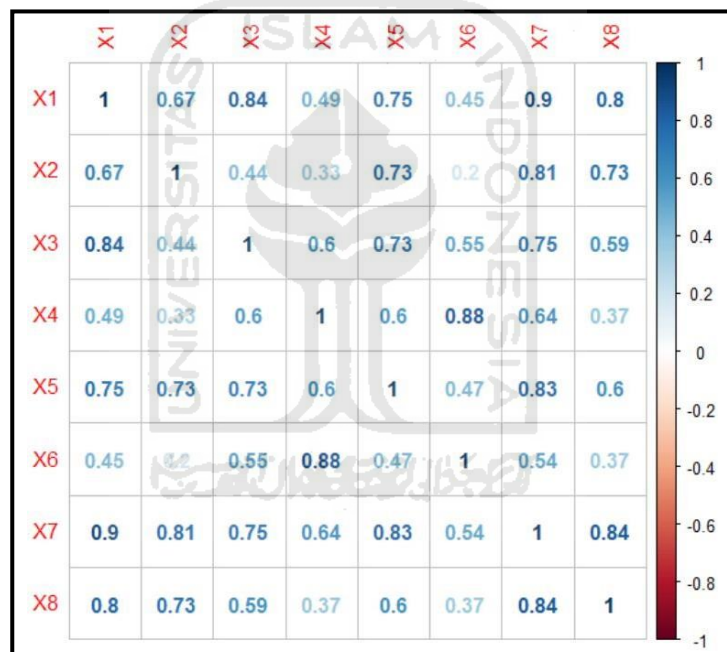
BAB V

HASIL DAN PEMBAHASAN

Pemetaan klasterisasi data kriminalitas di Indonesia dilakukan peneliti dengan menggunakan metode *k-means clustering* dengan mereduksi menggunakan *Principal Components Analysis* (PCA).

5.1 Korelasi

Tahapan analisis pertama adalah mengetahui pemetaan klasterisasi data kriminalitas di Indonesia, tahapan pertama dalam analisis ini adalah melakukan pengecekan korelasi antara variabel penelitian.



Gambar 5.1. Korelasi Antar variabel

Pada **Gambar 5.1** menunjukkan korelasi antara variabel berkorelasi kuat dan kurang berkorelasi antar variabel, namun memiliki hubungan yang positif. Hubungan positif ini ditandai dengan hasil yang menunjukkan seluruh variabel berwarna biru. Korelasi terendah adalah korelasi X2 atau Kejahatan NAPZA dan X6 atau Penganiayaan Ringan yaitu berkorelasi sebesar 0.2.

52 Uji KMO, MSA, Barlett's

Setelah melakukan pengecekan korelasi peneliti kemudian melanjutkan pengujian KMO dan Bartlett's. Pengujian KMO (Kaiser Meyer Olkin) dilakukan dengan tujuan memastikan apakah data yang akan dianalisis cukup memadai atau dapat digunakan dalam analisis PCA. Sementara pengujian Bartlett untuk menguji apakah analisis multivariat layak digunakan sebagai metode PCA. Berikut merupakan hasil uji KMO dan Bartlett's.

```
> uji_bart(novita[,2:9]) ##uji Bartlett's
      Bartlett's test of sphericity
data:  novita[, 2:9]
Khi-squared = 276.94, df = 28, p-value < 2.2e-16
> kmo(novita[,2:9]) ## Uji KMO
$KMO
[1] 0.808352
```

Gambar 5.2. Uji KMO dan Bartlett's

Uji KMO memiliki hipotesis sebagai berikut:

H_0 = Jumlah data telah cukup untuk dilakukan analisis

H_1 = Jumlah data tidak cukup untuk dilakukan analisis

Pada **Gambar 5.2.** diperoleh $KMO = 0.81$ dengan tingkat signifikan 0.05, maka gagal tolak H_0 karena $KMO (0.81) > \alpha (0.05)$, dengan menggunakan tingkat kepercayaan 95% dapat disimpulkan sebagai jumlah data yang dimiliki telah cukup untuk dilakukan analisis. Selanjutnya peneliti melakukan uji Bartlett's, uji ini digunakan untuk mengetahui apakah ada hubungan antar variabel dalam kasus multivariat. Jika X_1, X_2, \dots, X_{28} Independen (saling bebas) maka perpaduan antar variabel sama dengan label identitas. Hipotesis pada Uji ini adalah

H_0 : $p = 1$ (Analisis Multivariat tidak layak digunakan terutama metode PCA).

H_1 : $p \neq 1$ (Analisis Multivariat layak digunakan metode PCA)

Berdasarkan hasil pada **Gambar 5.2** diperoleh nilai *p-value* atau nilai *sig* adalah 0,000. Maka tolak H_0 karena *p-value* (0.000) < α (0.05), sehingga dengan tingkat kepercayaan 95% maka dapat disimpulkan analisis multivariat layak digunakan analisis PCA untuk data penelitian ini.

Tabel 5.1. Nilai MSA

No	Variabel	Nilai MSA	Keputusan
1	Pencurian dengan Pemberatan (X1)	0.81	Signifikan
2	Kejahatan NAPZA (X2)	0.82	Signifikan
3	Pencurian Kendaraan Bermotor (X3)	0.84	Signifikan
4	Pencurian (X4)	0.70	Signifikan
5	Penipuan (X5)	0.89	Signifikan
6	Penganiayaan Ringan (X6)	0.74	Signifikan
7	Penggelapan (X7)	0.80	Signifikan
8	Penganiayaan Berat (X8)	0.87	Signifikan

Selain pengecekan terhadap KMO dan uji Bartlett, dilakukan juga pengecekan pada nilai MSA (Tabel 5.1) dari 8 variabel memiliki nilai MSA lebih dari 0.5 yang berarti semua variabel layak untuk dilakukan analisis.

5.3 Pengelompokan Kriminalitas

Pengelompokan Kriminalitas yang terjadi di Indonesia pada tahun 2017 akan di faktorkan kedalam 3 pengelompokan menggunakan analisis PCA, dengan prinsip analisis PCA adalah mereduksi jumlah variabel yang banyak untuk kemudian diperkecil atau dikurangi jumlahnya namun tidak mengurangi informasi yang diberikan seperti yang diperoleh pada Tabel 5.2.

Variabel	Faktor1	Faktor2	Faktor3
Pencurian dengan Pemberatan (X1)	0,58	0,21	0,75
Kejahatan NAPZA (X2)	0,96	0,08	0,13
Pencurian Kendaraan Bermotor (X3)	0,27	0,38	0,85
Pencurian (X4)	0,24	0,92	0,21
Penipuan (X5)	0,66	0,39	0,43
Penganiayaan Ringan (X6)	0,1	0,93	0,24
Penggelapan (X7)	0,75	0,38	0,5
Penganiayaan Berat (X8)	0,74	0,12	0,47

Tabel 5.2. Pengelompokan Kriminalitas

Hasil pada Tabel 5.2 menunjukkan faktor 1 mengelompokkan 4 variabel yaitu X2, X5, X7, dan X8. Sedangkan variabel X1 dan X3 masuk kedalam faktor 3, serta faktor 2 adalah X4 dan X6. Seperti yang telah dibahas pada bab landasan teori bahwa prinsip mereduksi adalah meremperkecil jumlah variabel

tanpa mengurangi Informasi yang diberikan, maka melalui Tabel 5.3 menjelaskan seberapa jauh ketiga komponen (faktor) tersebut menjelaskan informasi yang terdapat dalam delapan variabel yang dianalisis. Peneliti bisa melihat pada bagian *proportion variance*, angka yang tertera pada kolom *proportion var* tersebut menunjukkan kemampuan tiap faktor dalam menjelaskan variabilitas keseluruhan data.

SS loadings atau *sum of squared loadings* digunakan untuk melihat kelayakan tiap faktor, jika nilai *SS loadings* lebih besar dari 1 menunjukkan faktor dapat dipertahankan, hasil tabel 5.3 menunjukkan semua faktor memiliki *SS loadings* lebih dari 1 artinya ketiga faktor ini dapat digunakan. Pada Tabel 5.3 *proportion variance* untuk faktor 1 adalah sebesar 0.37, artinya faktor pertama mampu menjelaskan variansi dari delapan variabel sebesar 37%, Kemudian faktor kedua mampu menjelaskan variansi dari keseluruhan variabel yang diamati sebesar 28%, sedangkan variansi dari delapan variabel mampu dijelaskan oleh faktor ketiga sebesar 26%.

Tabel 5.3. Parameter Faktor

Faktor	Faktor 1	Faktor 2	Faktor 3
<i>SS loadings</i>	2.96	2.21	2.06
<i>Proportion Var</i>	0.37	0.28	0.26
<i>Cumulative Var</i>	0.37	0.65	0.90
<i>Proportion Explained</i>	0.41	0.31	0.28
<i>Cumulative Proportion</i>	0.41	0.72	1.00

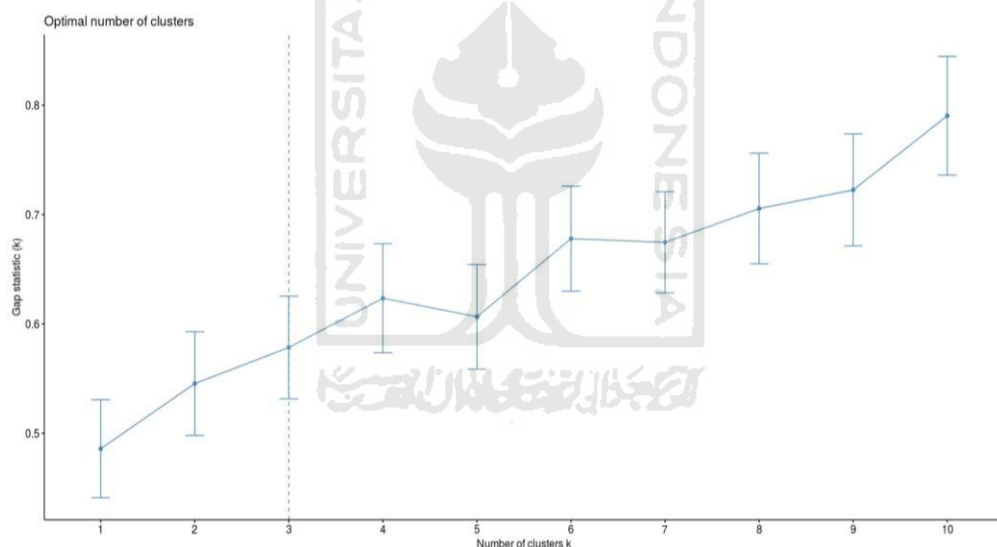
Hasil *Cumulative Variabel* pada Tabel 5.3 menunjukkan gabungan dari *proportion var*, dimana 0.62 menjelaskan faktor 1 dan faktor 2 yaitu $0.37 + 0.28 = 0.65$. Artinya adalah variansi dari delapan variabel mampu dijelaskan oleh faktor 1 dan faktor 2 sebesar 65%ss. Sedangkan untuk ketiga faktor yang terbentuk mampu menjelaskan variansi dari delapan variabel sebesar 90%.

54 Pengelompokan Daerah Kriminalitas

Pengelompokan berikutnya yang akan dilakukan adalah pengelompokan daerah kriminalitas, pengelompokan ini menggunakan analisis *cluster*. Pemilihan analisis *cluster* dikarenakan analisis ini berkerja dengan mengelompokkan objek-objek (daerah) berdasarkan karakteristik yang dimilikinya.

Berikut analisis *clustering* dalam mengelompokan daerah yang rawan terjadinya kriminalitas :

1. Menentukan jumlah cluster yang akan dibentuk disini peneliti menentukan 3 cluster atau 3 kelompok dimana cluster 1 aman, cluster 2 cukup rawan dan cluster 3 rawan terjadi kriminalitas. Sebelumnya jumlah 3 cluster tersebut didapatkan dari metode *GAP Statistic*.



Gambar 5.3. Grafik Metode *GAP Statistic*

Untuk penentuan nilai K pada *GAP Statistic* dilihat dari garis yang dihubungkan dengan garis putus-putus. Walaupun pada **Gambar 5.3** menunjukkan pada k=10 memiliki nilai tertinggi, namun kembali lagi pada logika bahwa data pada penelitian ini adalah 33 sehingga *cluster optimal* adalah 3.

2. Menentukan pusat cluster (centroid) penentuan titik pusat cluster yaitu dilakukan secara acak. Nilai pusat cluster dapat dilihat pada tabel berikut ini.

Tabel 5.4 Titik Pusat Cluster Awal

C1	539	486	489	568	351	362	322	197
C2	1.093	6.561	2.913	1.803	3.335	1.547	2.376	1.891
C3	2.327	1.296	2.523	1.582	1.761	1364	1.027	558

3. Menghitung jarak setiap data terhadap pusat cluster.

$$d(1,1) = \sqrt{\frac{(664 - 538,9)^2 + (1544 - 486,4)^2 + (968 - 489,1)^2 + (1351 - 567,9)^2 + (674 - 351)^2 + (1208 - 362)^2 + (643 - 322,2)^2 + (226 - 196,7)^2}{8}} = 1635,09$$

$$d(1,2) = \sqrt{\frac{(664 - 1092,5)^2 + (1544 - 6560,5)^2 + (968 - 2913)^2 + (1351 - 1803)^2 + (674 - 3335)^2 + (1208 - 1546,5)^2 + (643 - 2376)^2 + (226 - 1890,5)^2}{8}} = 6504,27$$

$$d(1,3) = \sqrt{\frac{(664 - 2326,7)^2 + (1544 - 1296,2)^2 + (968 - 2522,8)^2 + (1351 - 1582,4)^2 + (674 - 1760,5)^2 + (1208 - 1364)^2 + (643 - 1026,5)^2 + (226 - 196,7)^2}{8}} = 2599,83$$

$$d(2,1) = \sqrt{\frac{(5066 - 538,9)^2 + (5907 - 486,4)^2 + (4030 - 489,1)^2 + (2282 - 567,9)^2 + (2434 - 351)^2 + (2842 - 362)^2 + (2831 - 322,2)^2 + (2791 - 196,7)^2}{8}} = 9187,30$$

$$d(2,2) = \sqrt{\frac{(5066 - 1092,5)^2 + (5907 - 6560,5)^2 + (4030 - 2913)^2 + (2282 - 1803)^2 + (2434 - 3335)^2 + (2842 - 1546,5)^2 + (2831 - 2376)^2 + (2791 - 1890,5)^2}{8}} = 4604,5$$

$$d(2,3) = \sqrt{\frac{(5066 - 2326,7)^2 + (5907 - 1296,2)^2 + (4030 - 2522,8)^2 + (2282 - 1582,4)^2 + (2434 - 1760,5)^2 + (2842 - 1364)^2 + (2831 - 1026,5)^2 + (2791 - 196,7)^2}{8}} = 6511,94$$

Lakukan perhitungan sampai data ke-133.

Tabel 5.5 Hasil Pengelompokan

Kepolisian	(X1)	(X2)	(X3)	(X4)	(X5)	(X6)	(X7)	(X8)	C1	C2	C3	Kelas
Aceh	664	1.544	968	1351	674	1208	643	226	1635,09	6504,275077	2599,838489	1
Kalimantan Tengah	288	474	170	136	73	133	160	50	573,30	8302,554793	4222,846622	1
Kalimantan Selatan	594	1211	494	393	261	112	275	221	707,8272	7444,060982	3610,392096	1
Kalimantan Timur	767	2928	1059	584	357	245	491	454	2514,70	5865,445293	3428,508983	1
Riau	1078	1173	676	541	323	405	624	185	876,23	7209,409196	3058,938705	1
Jambi	1683	712	1048	995	942	467	905	293	1446,43	7091,70706	2198,972219	1
Bengkulu	810	192	665	466	365	63	342	186	410,55	8095,853176	3530,557707	1
Lampung	1504	1744	271	455	710	124	707	152	1630,15	6847,431624	3182,322427	1
Kep.Riau	271	400	416	715	404	345	288	154	0	7965,862742	3715,854144	1
DI Yogyakarta	604	290	573	521	510	144	296	93	0	8011,36669	3565,223392	1

Banten	580	389	810	153	421	208	193	106	460,5767141	7982,989341	3588,992648	1
Bali	405	665	328	447	203	418	235	6	120,5242836	7935,284518	3801,297541	1
Nusa Tenggara Barat	930	227	281	649	385	99	316	50	487,9320063	8174,813498	3639,554364	1
Nusa Tenggara Timur	276	2	349	1069	486	1533	274	484	1381,276354	8081,282688	3618,665076	1
Kalimantan Barat	684	490	560	970	223	0	446	79	474,3372302	7894,998794	3516,702717	1
Sulawesi Utara	70	111	250	809	691	0	576	585	836,5847708	8065,231397	3933,189273	1
Papua	628	41	1920	681	273	413	216	373	1479,343008	7911,616564	3073,392871	1
Kep.Bangka Belitung	479	224	126	128	108	3	80	50	695,5985465	8517,959057	4261,580613	1
Sulawesi Tenggara	167	65	279	440	149	744	70	9	705,8195745	8455,538841	4088,763705	1
Gorontalo	15	38	134	711	302	1250	239	34	1149,454367	8361,803102	4053,165805	1
Sulawesi Barat	106	120	133	359	173	451	134	15	667,4794423	8485,826543	4238,918807	1
Maluku	99	73	145	413	219	121	98	576	797,583359	8445,554367	4273,754924	1
Maluku Utara	8	63	11	76	18	38	22	107	1005,86501	8763,452549	4634,996312	1
Papua Barat	224	41	74	568	154	164	103	233	659,9079178	8531,546069	4242,937332	1
Sumatera Utara	5.066	5907	4030	2282	2434	2842	2831	2791	9187,304379	4604,41948	6511,943555	2
Metro jaya	2180	7214	1796	1324	4236	251	1921	990	7301,472812	2568,03879	6632,457234	2
Sumatera Barat	2521	592	2910	1604	967	1744	910	329	3608,931963	6923,654289	1234,544021	3
Sumatera Selatan	2771	1768	1050	508	1127	282	1237	1239	2958,08376	6255,325669	2413,373233	3
Jawa Barat	3727	1090	5414	960	3345	1009	1429	862	6074,720315	6787,935189	3693,310746	3
Jawa Tengah	2035	1243	1665	821	785	369	726	187	2073,176169	6723,042228	1890,253692	3
Jawa Timur	2822	3405	2072	1366	2602	693	981	1026	4240,690669	4219,670941	2514,327595	3
Sulawesi Tengah	778	259	2357	2378	667	1487	662	32	2856,531196	7345,827166	2398,97565	3
Sulawesi Selatan	1633	717	2192	3440	2831	3964	1241	228	5108,607598	6912,934278	3526,650614	3

4. Menentukan kembali titik pusat cluster yang baru berdasarkan rata-rata.

Menentukan cluster dapat ditentukan dengan rumus = Nilai hasil/banyak hasil kluster.

Tabel 5.6 Titik pusat Cluster Baru

Cluster 1	Cluster 2	Cluster 3
538,9166	3623	2326,714
550,708	6560,5	1296,285
489,166	2913	2522,857
567,9166	1803	1582,428
351	3335	1760,5714
362	1546,5	1364
322,208	2376	1026,571
196,7083	1890,5	557,57146

5. Menghitung jarak data dengan pusat cluster yang baru

$$d(1,1) = \sqrt{\frac{(664 - 538,9)^2 + (1544 - 486,4)^2 + (968 - 489,1)^2 + (1351 - 567,9)^2 + ((674 - 351)^2 + (1208 - 362)^2 + (643 - 322,2)^2 + (226 - 196,7)^2}{4}} = 1216,64$$

$$d(1,2) = \sqrt{\frac{(664 - 3623)^2 + (1544 - 6560,5)^2 + (968 - 2913)^2 + (1351 - 1803)^2 + ((674 - 3335)^2 + (1208 - 1546,5)^2 + (643 - 2376)^2 + (226 - 1890,5)^2)}{4}} = 7132,85$$

$$d(1,3) = \sqrt{\frac{(664 - 2326,7)^2 + (1544 - 1296,2)^2 + (968 - 2522,8)^2 + (1351 - 1582,4)^2 + ((674 - 1760,5)^2 + (1208 - 1364)^2 + (643 - 1026,5)^2 + (226 - 1557,5)^2)}{4}} = 2599,83$$

$$d(2,1) = \sqrt{\frac{(5066 - 538,9)^2 + (5907 - 486,4)^2 + (4030 - 489,1)^2 + (2282 - 567,9)^2 + ((2434 - 351)^2 + (2842 - 362)^2 + (2831 - 322,2)^2 + (2791 - 196,7)^2)}{4}} = 97693,26$$

$$d(2,2) = \sqrt{\frac{(5066 - 3623)^2 + (5907 - 6560,5)^2 + (4030 - 2913)^2 + (2282 - 1803)^2 + ((2434 - 3335)^2 + (2842 - 1546,5)^2 + (2831 - 2376)^2 + (2791 - 1890,5)^2)}{4}} = 2737,60$$

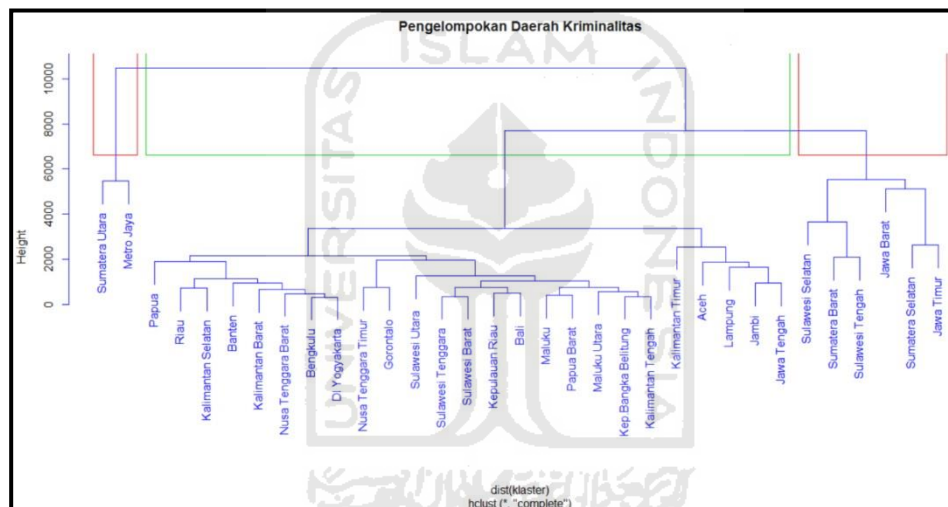
$$d(2,3) = \sqrt{\frac{(5066 - 2326,7)^2 + (5907 - 1296,2)^2 + (4030 - 2522,8)^2 + (2282 - 1582,4)^2 + ((2434 - 1760,5)^2 + (2842 - 1364)^2 + (2831 - 1026,5)^2 + (2791 - 557,6)^2)}{4}} = 6511,94$$

Hitung sampai data ke-33. sehingga diperoleh hasil pengelompokan baru sebagai berikut:

Tabel 5.7 Hasil Pengelompokan Baru

Kepolisian	(X1)	(X2)	(X3)	(X4)	(X5)	(X6)	(X7)	(X8)	c1	c2	c3	Kelas
Aceh	664	1.544	968	1351	674	1208	643	226	1216,643471	7132,856002	2599,838688	1
Kalimantan Tengah	288	474	170	136	73	133	160	50	475,5341752	8911,081177	4222,846813	1
Kalimantan Selatan	594	1211	494	393	261	112	275	221	0	8021,243654	3610,39229	1
Kalimantan Timur	767	2928	1059	584	357	245	491	454	430,1786471	6515,689737	3428,50916	1
Riau	1078	1173	676	541	323	405	624	185	343,4875084	7645,416584	3058,938925	1
Jambi	1683	712	1048	995	942	467	905	293	1486,700402	7328,523572	2198,972404	1
Bengkulu	810	192	665	466	365	63	342	186	0	8565,978213	3530,557897	1
Lampung	1504	1744	271	455	710	124	707	152	1012,810694	7155,988803	3182,322589	1
Kep.Riau	271	400	416	715	404	345	288	154	0	8603,252684	3715,854323	1
DI Yogyakarta	604	290	573	521	510	144	296	93	0	8547,379291	3565,223564	1
Banten	580	389	810	153	421	208	193	106	165,6074729	8527,911688	3588,992824	1
Bali	405	665	328	447	203	418	235	6	0	8535,312692	3801,297738	1
Nusa Tenggara Barat	930	227	281	649	385	99	316	50	0	8605,42897	3639,554553	1
Nusa Tenggara Timur	276	2	349	1069	486	1533	274	484	1229,188859	8708,777914	3618,665252	1
Kalimantan Barat	684	490	560	970	223	0	446	79	248,1038124	8414,382197	3516,702929	1
Sulawesi Utara	70	111	250	809	691	0	576	585	702,7522098	8753,638829	3933,189408	1
Papua	628	41	1920	681	273	413	216	373	1347,521058	8446,769249	3073,393086	1
Kep.Bangka Belitung	479	224	126	128	108	3	80	50	542,2815244	9058,914932	4261,5808	1
Sulawesi Tenggara	167	65	279	440	149	744	70	9	425,6894428	9087,544319	4088,763897	1
Gorontalo	15	38	134	711	302	1250	239	34	970,6728775	9043,027024	4053,165983	1
Sulawesi Barat	106	120	133	359	173	451	134	15	403,4889528	9132,649876	4238,91899	1
Maluku	99	73	145	413	219	121	98	576	549,9149192	9097,189497	4273,755097	1
Maluku Utara	8	63	11	76	18	38	22	107	873,6811976	9417,542872	4634,996493	1
Papua Barat	224	41	74	568	154	164	103	233	309,3211682	9142,545037	4242,937517	1
Sumatera Utara	5.066	5907	4030	2282	2434	2842	2831	2791	7693,267647	2737,606208	6511,943482	2
Metro jaya	2180	7214	1796	1324	4236	251	1921	990	4792,447879	2737,606208	6632,457058	2
Sumatera Barat	2521	592	2910	1604	967	1744	910	329	3634,865816	6863,736209	1234,544302	3
Sumatera Selatan	2771	1768	1050	508	1127	282	1237	1239	2744,193886	6085,864914	2413,373348	3
Jawa Barat	3727	1090	5414	960	3345	1009	1429	862	6731,986644	6256,713494	3693,310525	3
Jawa Tengah	2035	1243	1665	821	785	369	726	187	1932,462376	6843,449185	1890,25393	3
Jawa Timur	2822	3405	2072	1366	2602	693	981	1026	3788,559929	3931,433295	2514,327443	3
Sulawesi Tengah	778	259	2357	2378	667	1487	662	32	2833,268681	7871,230574	2398,975866	3
Sulawesi Selatan	1633	717	2192	3440	2831	3964	1241	228	5657,9962	7173,329475	3526,650504	3

Hasil dari tahapan yang pertama dan kedua tidak mengalami perubahan pada hasil pengelompokan, maka hasil sudah sesuai dengan pengelompokan cluster. Ditinjau dari landasan teori analisis *cluster* mengklasifikasi daerah-daerah ini sehingga setiap daerah yang paling dekat kesamaannya dengan daerah lain berada dalam *cluster* yang sama. Oleh karena itu dipilih 3 cluster karena Peneliti ingin melihat kategori daerah Aman, Cukup Rawan dan Rawan menggunakan metode *K-means Clustering*, dengan menggunakan jarak terjauh yang membentuk dari *cluster* atau metode *complete*. Visualisasi hasil *cluster* daerah kriminalitas seperti pada Gambar 5.4.



Gambar 5.4. Pengelompokan Daerah Kriminalitas

Pengelompokan daerah kriminalitas dibagi kedalam tiga kelompok, hasil pengelompokan untuk kelompok pertama terdiri dari 25 daerah, untuk kelompok kedua terdiri dari 2 daerah, serta kelompok ketiga terdiri dari 6 daerah. Lebih jelas mengenai daerah-daerah tiap kelompok dijelaskan dalam bentuk Tabel 5.8.

Tabel 5.8. Pengelompokan Daerah Kriminalitas

Kelompok	Daerah
Kelompok 1	Aceh
	Riau
	Jambi
	Bengkulu
	Lampung
	Kep.Bangka Belitung
	Kepulauan Riau
	Jawa Tengah
	DI Yogyakarta
	Banten
	Bali
	Nusa Tenggara Barat
	Nusa Tenggara Timur
	Kalimantan Barat
Kelompok 2	Kalimantan Tengah
	Kalimantan Selatan
	Kalimantan Timur
	Sulawesi Utara
	Sulawesi Tenggara
	Gorontalo
	Sulawesi Barat
	Maluku
	Maluku Utara
	Papua Barat
	Papua
Kelompok 3	Sumatera Utara
	Metro Jaya
Kelompok 3	Sumatera Barat
	Sumatera Selatan
	Jawa Barat
	Jawa Timur
	Sulawesi Tengah
	Sulawesi Selatan

Interpretasi dan Profiling Berdasarkan Tingkat Kriminalitas

Tabel 5.9 Final Cluster Center

Kelompok	X1	X2	X3	X4	X5	X6	X7	X8
1	538.91	550.7	489.1	567.9	351	362	322.2	196.7
2	3623	6560.5	2913	1803	3335	1546.5	2376	1890.5
3	2326.7	1296.2	2522.8	1582.4	1760.5	1364	1026.57	557.5
Rata-Rata	2162.8	2802.4	1975	1317.7	1815.5	1090.8	1241.5	881.5

Keterangan : 

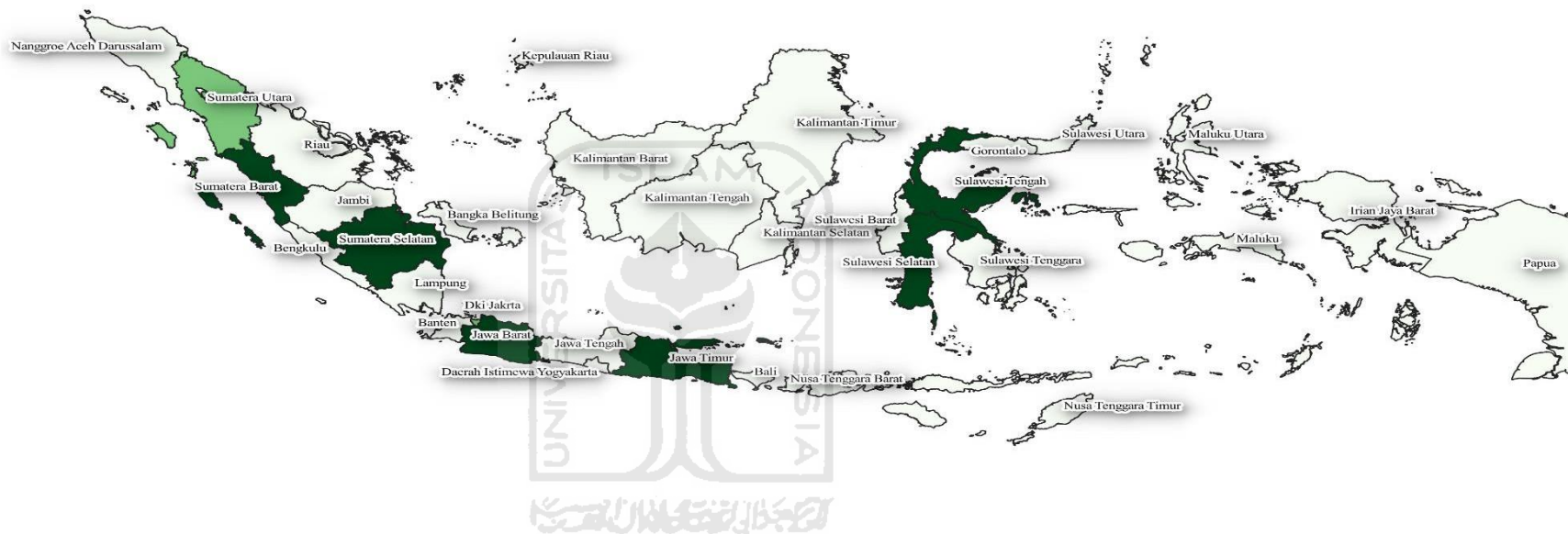
Berdasarkan tabel *final cluster center* diatas dapat disimpulkan dimana Provinsi pada cluster 1 menunjukkan tingkat kriminalitas pada daerah tersebut sangat rendah atau dapat dikatakan provinsi pada cluster 1 aman yaitu dilihat dari nilai cluster pada setiap variabelnya yang rendah. Pada cluster 1 walaupun masuk kedalam kategori provinsi yang aman akan tetapi perlu di waspadai untuk jenis kejahatan X4 (Pencurian) karena diwilayah yng masuk cluster 1 angka pencuriannya merupakan tindak kejahatan paling tinggi. Sedangkan untuk rata rata tindak kriminalitasnya kelompok 1 masuk dalam kategori daerah dengan kasus kriminalitas yang cukup rendah. Untuk Provinsi pada cluster 2 dilihat dari nilai cluster setiap variabelnya menunjukkan tingkat krimanalitas pada daerah cluster 2 ini sangat tinggi dibandingkan cluster 1 dan 3 atau dapat dikatakan provinsi pada cluster ke 2 ini dikategorikan rawan, untuk jenis kriminalitas provinsi pada cluster ke 2 yang tertinggi pada kasus kejahatan X2(kejahatan NAPZA). Sedangkan untuk Provinsi yang termasuk pada cluster 3 menunjukkan daerah dengan rawan kriminalitas hanya saja untuk rata - rata jenis kriminalitas pada cluster ini cukup rawan tidak seperti pada cluster ke 2 yang jenis kriminalitasnya tinggi. Hanya saja Untuk kejahatan tertinggi pada cluster 3 adalah kasus X3(Pencurian Kendaraan Bermotor).

55. Pemetaan

Kelompok yang terbentuk divisualisasikan dalam bentuk peta pada gambar 5.5. terdapat 3 kategori daerah rawan kriminalitas, yaitu aman, Cukup rawan, rawan. Daerah dengan tingkat kriminalitas yang tinggi memiliki warna biru yang gelap sedangkan daerah dengan tingkat kriminalitas yang rendah digambarkan dengan warna yang terang. Dengan kata lain bahwa semakin terang warna maka daerah tersebut semakin aman.



PETA DAERAH RAWAN KRIMINALITAS DI INDONESIA PADA TAHUN 2017



- Keterangan
- Aman
 - Cukup Rawan
 - Rawan

Gambar 5.5 Peta Daerah Rawan Kriminalitas

BAB VI

KESIMPULAN

61. Kesimpulan

Berdasarkan analisis yang telah dilakukan maka didapatkan kesimpulan bahwa:

Provinsi pada *cluster* 1 menunjukkan tingkat kriminalitas pada daerah tersebut sangat rendah atau dapat dikatakan provinsi pada *cluster* 1 aman yaitu dilihat dari nilai klaster pada setiap variabelnya yang rendah. Untuk Provinsi pada *cluster* 2 dilihat dari nilai *cluster* setiap variabelnya menunjukkan tingkat kriminalitas pada daerah *cluster* 2 ini sangat tinggi atau dapat dikatakan rawan terjadinya kriminalitas. Provinsi yang termasuk pada *cluster* 3 menunjukkan tingkat kriminalitasnya cukup tinggi atau dapat dikatakan cukup rawan terjadinya kriminalitas.

Provinsi yang masuk dalam *Custer* I yaitu kategori Daerah aman dengan anggota kelompok yaitu: Aceh, Riau, Bengkulu, Lampung, Kep. Bangka Belitung, Kepulauan Riau, Jawa Tengah, DI Yogyakarta, Banten, Bali, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Sulawesi Utara, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua. Provinsi pada *Cluster* II masuk kategori daerah cukup Rawan dengan anggota kelompok yaitu: Sumatera Utara, Metro Jaya. Provinsi pada *Cluster* III masuk kategori sangat rawan dengan anggota kelompok yaitu : Sumatera Barat, Sumatera Selatan, Jawa Barat, Jawa Timur, Sulawesi Tengah, Sulawesi Selatan.

62. Saran

Saran yang dapat diberikan dalam penelitian ini yaitu sebagai berikut:

1. Untuk Kepolisian Republik Indonesia dapat mengatur skala prioritas pada tiap daerah dan memberikan perhatian khusus terlebih diwilayah dengan tingkat kerawanan kriminalitas yang cukup tinggi. Sehingga diharapkan

daerah dengan angka kriminalitas yang sangat tinggi di daerah rawan tersebut bisa berkurang.

2. Penelitian ini juga dapat dilanjutkan dalam beberapa tahun kedepan dengan menambah jumlah variabel ataupun menggunakan data *crime-rate* sehingga memperoleh hasil yang lebih baik dan lebih akurat.



DAFTAR PUSTAKA

- Arif Pratisto. 2004. *Cara mudah mengatasi masalah statistik dan rancangan percobaan dengan SPSS 12*. PT Elex Media Computindo: Jakarta.
- Badan Pusat Statitistik Daerah Istimewa Yogyakarta. 2016. *Statistik Politik dan Kemanan DIY 2005- 2015*. Badan Pusat Statitistik: Kabupaten Sleman.
- Badan Pusat Statistik. 2018. *Statistik Kriminal*. Badan Pusat Statistik: Jakarta.
- Dermawanti, dkk. 2015. Faktor- Faktor Yang Mempengaruhi Kriminalitas di Kabupaten Batng pada Tahun 2013 Dengan Analisis Jalur.
- F. L, Whitney. 1960. *The Elements Of Resert*. Asian Eds. Osaka: Overseas Book Co.
- Flowrensia, Yanne. 2010. Perbandinga Penggerombolan K-Means dan K-Medoid pada Data yang Mengandung Pencilan. IPB: Bogor.
- Hamzah, A. 1986. *Kamus Hukum*. Ghalia Indonesia: Jakarta.
- Hasan, M Iqbal. 2002. *Pokok-Pokok Materi Metodologi Penelitian dan Aplikasinya*.Ghalia Indonesia, Bogor
- Hidayah, Resti. 2014. *Sistem Informasi Geografis berbasis Web Untuk produksi, produktivitas, dan luas panen kedelai di Indonesia menggunakan K-Means Clustering*. Semirata IPB: Bogor.
- Kartini, Kartono. 2005. *Patologi Sosial*. Raja Grafindo Persada: Jakarta.
- Khaira, Ulfa. 2012. Integritas Self Organizing Maps dan Algorit me K-Means untuk Clustering Data Ketahanan Pangan Kabupaten di Wilayah Provinsi Bali, NTB, dan NTT. IPB: Bogor.
- Lind, Marchal & Wathen. (2008). *Teknik-Teknik Statistika dalam Bisnis Dan Ekonomi Menggunakan Kelompok Data Global*. Buku 2, Edisi 13. Salemba Empat: Jakarta.
- M. Pandu Ristiyono. 2016. Pemetaan Bidang Ilmu Berdasarkan Artikel Jurnal Pendidikan Universitas Terbuka Tahun 1999-2007: Menggunakan Analisis Co-Words.Tesis., Institut Pertanian Bogor: Bogor.

- Rivani, Edmira. 2010. Aplikasi K-Means Cluster untuk Pengelompokkan Berdasarkan Produksi Padi, Jagung, Kedelai, dan Kacang Hijau Tahun 2009. *Jurnal Mat Stat*, 10(2), pp. 122-134.
- Rosiatun, Aan. 2012. Analisis Cluster untuk Segmentasi Pemirsa Program Berita Sore Stasiun Tv Swasta. *Media Statistika*, 3(2), pp. 93-102.
- Soraya, Neli. 2009. Analisis Gerombol K-Means dan K-Error untuk Pengelompokan Desa-desa di Jawa Barat berdasarkan Indikator Kemiskinan. IPB: Bogor.
- Spasser, M. 1997. Mapping the Terrain of Pharmacy: Co-classification Analytsis of The International Pharmaceutical Abstracts Database. *Scientometric*, 77-97
- Sugiyono. 2007. *Statistika untuk Penelitian*. Bandung: Alfabeta.
- Sujatmiko, Ponco. 2005. *Matematika Kreatif 1*. P.T Tiga Serangkai: Solo.
- Walpole, R. E. dan Myers, R. H. 1995. *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan*, Terj. *Probability and Statistics for Engineers and Scientists*, oleh RK Sembiring. Penerbit ITB: Bandung.

LAMPIRAN



Lampiran 1

Data Penelitian

Kepolisian	(X1)	(X2)	(X3)	(X4)	(X5)	(X6)	(X7)	(X8)
Aceh	664	1544	968	1351	674	1208	643	226
Sumatera Utara	5066	5907	4030	2282	2434	2842	2831	2791
Sumatera Barat	2521	592	2910	1604	967	1744	910	329
Riau	1078	1173	676	541	323	405	624	185
Jambi	1683	712	1048	995	942	467	905	293
Sumatera Selatan	2771	1768	1050	508	1127	282	1237	1239
Bengkulu	810	192	665	466	365	63	342	186
Lampung	1504	1744	271	455	710	124	707	152
Kep.Bangk a Belitung	479	224	126	128	108	3	80	50
Kep.Riau	271	400	416	715	404	345	288	154
Metro jaya	2180	7214	1796	1324	4236	251	1921	990
Jawa Barat	3727	1090	5414	960	3345	1009	1429	862
Jawa Tengah	2035	1243	1665	821	785	369	726	187
DI Yogyakarta	604	290	573	521	510	144	296	93
Jawa Timur	2822	3405	2072	1366	2602	693	981	1026
Banten	580	389	810	153	421	208	193	106
Bali	405	665	328	447	203	418	235	6
Nusa Tenggara Barat	930	227	281	649	385	99	316	50
Nusa Tenggara	276	2	349	1069	486	1533	274	484

Timur								
Kalimantan Barat	684	490	560	970	223	0	446	79
Kalimantan Tengah	288	474	170	136	73	133	160	50
Kalimantan Selatan	594	1211	494	393	261	112	275	221
Kalimantan Timur	767	2928	1059	584	357	245	491	454
Sulawesi Utara	70	111	250	809	691	0	576	585
Sulawesi Tengah	778	259	2357	2378	667	1487	662	32
Sulawesi Selatan	1633	717	2192	3440	2831	3964	1241	228
Sulawesi Tenggara	167	65	279	440	149	744	70	9
Gorontalo	15	38	134	711	302	1250	239	34
Sulawesi Barat	106	120	133	359	173	451	134	15
Maluku	99	73	145	413	219	121	98	576
Maluku Utara	8	63	11	76	18	38	22	107
Papua Barat	224	41	74	568	154	164	103	233
Papua	628	41	1920	681	273	413	216	373

Lampiran 2

Syntax

```
### _____ Pengelompokan
Variabel
data=read.delim("clipboard")
View(data) ## menampilkan data di tab baru
head(data) ## menampilkan data namun hanya beberapa
data teratas
data2=data.frame(novita[,2:9]) ## seleksi variabel yang
akan digunakan
colnames(data2)<-
c("X1","X2","X3","X4","X5","X6","X7","X8") ## mengubah
nama setiap variabel
rownames(data2)<-novita$Kepolisian.Daerah ## mengubah
nama baris jadi nama2 daerah
head(data2) ## tampilkan data yang telah diubah dengan
nama data2, data ini yang akan digunakan

## Pengecekan ada tidaknya data NA
summary(data)

library(psych)
library(factoextra)

## Uji korelasi pearson
cor(data2) ##Pengecekan korelasi
View(cor(data2))
## Korelasi secara visual
library(corrplot)
corrplot(cor(novita2), method="number")
```

```

## Melakukan uji Bartlett's dan Uji KMO
kmo <- function(x)
{
  x <- subset(x, complete.cases(x))
  r <- cor(x)
  r2 <- r^2
  i <- solve(r)
  d <- diag(i)
  p2 <- (-i/sqrt(outer(d, d)))^2
  diag(r2) <- diag(p2) <- 0
  KMO <- sum(r2)/(sum(r2)+sum(p2))
  MSA <- colSums(r2)/(colSums(r2)+colSums(p2))
  return(list(KMO=KMO, MSA=MSA))
}

uji_bart <- function(x)
{
  method <- "Bartlett's test of sphericity"
  data.name <- deparse(substitute(x))
  x <- subset(x, complete.cases(x))
  n <- nrow(x)
  p <- ncol(x)
  chisq <- (1-n+(2*p+5)/6)*log(det(cor(x)))
  df <- p*(p-1)/2
  p.value <- pchisq(chisq, df, lower.tail=FALSE)
  names(chisq) <- "Khi-squared"
  names(df) <- "df"
  return(structure(list(statistic=chisq, parameter=df,
p.value=p.value,

```

```

                                method=method,
data.name=data.name), class="hstest"))
}

uji_bart(novita[,2:9]) ##uji Bartlett's
kmo(novita[,2:9]) ## Uji KMO

### Hasil PCA atau AKU untuk 3 faktor
library('GPArotation') ## package rotasi
pcadata <-
principal(data2,nfactor=3,rotate="varimax",scores=T) ##
menggunakan rotasi varimax
pcadata

###
=====
=====
===== ###
###
-----
Pengelompokan Derah Kriminalitas

-----
###
#install.packages("tidyverse")
library(tidyverse)
#install.packages("cluster") #Algoritma klastering
library(cluster)
#install.packages("factoextra") #Algoritma klastering
dan visualisasi

```

```
library(factoextra)
## Menginput data
daerah=read.delim("clipboard")
klaster=data.frame(daerah[,2:9]) ## memilih variabel
yang digunakan hanya X1-X8 berda di kolom ke 2 sampai
ke kolom 9
rownames(klaster)<-daerah$Kepolisian.Daerah ## mengubah
penomoran data menjadi nama daerah
colnames(klaster)<-
c("X1","X2","X3","X4","X5","X6","X7","X8") ## mengubah
nama setiap variabel
head(klaster) ## menampilkan data yang telah di
diseleksi dan diubah nama kolom dan baris

plot(daerah, col="red",lwd=2)
text(daerah,rownames(daerah))
##Cluster Hirarki
## complete adalah jarak terjauh yang membentuk cluster
hasil<-hclust(dist(klaster),"complete")
hasil
plot(hasil, col="blue",main="Pengelompokan Daerah
Kriminalitas") ## Menampilkan plot
str(hasil)
hasil$height
## urutan pengelompokan dari pertama
hasil$merge

# Penentuan Cluster
set.seed(200) #Mengunci data
gap_stat <-clusGap(data2, FUN=kmeans, nstart=2,
K.max=10)
```

```
fviz_gap_stat(gap_stat)

#Pembagian menjadi 3 kelompok
rect.hclust(hasil,k=3,border=2:3)
hasil.cut<-cutree(hasil,3)
table(hasil.cut) ## jumlah daerah pada masing-masing
kelompok

#mengetahu anggota daerah tiap kelompok
rownames(klaster)[hasil.cut==1]
rownames(klaster)[hasil.cut==2]
rownames(klaster)[hasil.cut==3]
```

