

LAPORAN PRAKTEK KERJA LAPANGAN

Analisis Rekapitulasi Data Pemilih Tetap Pilkada 2024 Provinsi Lampung Menggunakan K-Means Clustering

Diajukan untuk memenuhi persyaratan kelulusan
Matakuliah IF4004 Praktek Kerja Lapangan

Oleh:
Arsyadana Estu Aziz
121140068



**PROGRAM STUDI TEKNIK INFORMATIKA
INSTITUT TEKNOLOGI SUMATERA
2024**

Lembar Pengesahan Program Studi Teknik Informatika

Analisis Rekapitulasi Data Pemilih Tetap Pilkada 2024 Provinsi Lampung Menggunakan K-Means Clustering

di KPU Provinsi Lampung

Oleh:

Arsyadana Estu Aziz

121140068

disetujui dan disahkan sebagai
Laporan Praktek Kerja Lapangan

Lampung Selatan, 15 November 2024
Pembimbing Praktek Kerja Lapangan Program Studi Teknik Informatika ITERA



Meida Cahyo Untoro, S.Kom., M.Kom.

NIP. 198905182019031011

Lembar Pengesahan Program Studi Teknik Informatika

Analisis Rekapitulasi Data Pemilih Tetap Pilkada 2024 Provinsi Lampung Menggunakan K-Means Clustering

di KPU Provinsi Lampung

Oleh:
Arsyadana Estu Aziz
121140068

disetujui dan disahkan sebagai
Laporan Praktek Kerja Lapangan

Bandar Lampung, 5 Agustus 2024
Kasubbag Data dan Informasi KPU Provinsi Lampung



Abstrak

Penelitian ini menyajikan analisis data pemilih dan populasi menggunakan metode clustering K-Means, dengan tujuan untuk mengungkap pola demografi pemilih yang bermanfaat bagi perencanaan pemilu oleh KPU (Komisi Pemilihan Umum). Dataset mencakup variabel seperti total populasi, rasio jenis kelamin, tingkat pertumbuhan populasi, dan kelayakan pemilih yang penting untuk memahami distribusi serta karakteristik pemilih di berbagai wilayah. Dengan menggunakan teknik clustering, khususnya K-Means, wilayah-wilayah dikategorikan ke dalam kelompok yang bermakna untuk mendukung pengambilan keputusan dalam konteks pemilu. Penelitian ini juga membahas tantangan terkait skala fitur, normalisasi, dan pemilihan jumlah cluster optimal untuk menyeimbangkan perbedaan antar wilayah dan mengidentifikasi wawasan yang berguna. Hasil penelitian menunjukkan bagaimana metode clustering dapat mengungkap profil demografi yang berbeda serta menyoroti faktor-faktor kunci yang memengaruhi partisipasi pemilih dan distribusi sumber daya. Temuan ini memberikan kontribusi dalam strategi Pilkada yang lebih terinformasi, memastikan representasi yang adil dan manajemen pemilu yang efisien.

Kata Kunci: K-Means Clustering, Unsupervised Learning, Pilkada, KPU

Kata Pengantar

Puji syukur ke hadirat Allah Swt. yang Maha Esa yang telah memberikan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan Laporan Kerja Praktik ini. Shalawat serta salam tak lupa juga penulis sampaikan kepada Nabi Muhammad SAW, beserta para sahabatnya. Semoga kita semua sebagai para pengikutnya bisa mendapatkan syafaat-Nya di akhir nanti.

Laporan ini disusun sebagai salah satu syarat Mata Kuliah Kerja Praktik Program Studi Teknik Informatika Institut Teknologi Sumatera. Dalam pelaksanaan kerja praktik ini, penulis tidak terlepas dari adanya bimbingan, arahan, dan bantuan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Ibu Ressy Silvia Dewi dan Ibu Nurwafa Finanda selaku pembimbing kerja praktik di KPU Provinsi Lampung.
2. Bapak Meida Cahyo Untoro selaku dosen pembimbing kerja praktik.
3. Rekan-rekan kerja praktik.
4. Kedua orang tua yang senantiasa memberikan doa dan dukungan kepada penulis sehingga dapat menyelesaikan laporan kerja praktik ini.
5. Semua pihak yang terlibat dalam membantu penyusunan laporan ini.

Demikian kata pengantar yang dapat penulis sampaikan. Semoga laporan ini dapat bermanfaat bagi semua pihak. Penulis juga memohon maaf apabila terdapat kesalahan dalam penulisan laporan kerja praktik ini. Atas perhatiannya, penulis mengucapkan terima kasih.

Daftar Isi

Lembar Pengesahan	ii
Abstrak	ii
Kata Pengantar	v
BAB I Pendahuluan	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Manfaat Penelitian	2
1.5 Ruang Lingkup	2
1.6 Sistematika Penulisan	2
BAB II Gambaran Umum Instansi	4
2.1 Profil Organisasi	4
2.2 Visi dan Misi Organisasi	4
2.3 Struktur Organisasi	5
2.4 Deskripsi Pekerjaan	6
2.5 Jadwal Kerja	6
BAB III Landasan Teori	7
3.1 Data Mining	7
3.2 Data Clustering	7
3.2.1 K-Means Clustering	7
3.2.2 Hierarchical Clustering	8
3.2.3 DBSCAN (Density-Based Spatial Clustering)	8
3.3 Evaluasi Jarak dalam Clustering	8
3.3.1 Euclidean Distance	8
3.3.2 Manhattan Distance	9
3.3.3 Cosine Similarity	9
3.4 Elbow Method	9
3.5 Silhouette Score	10
3.6 Data Normalization	10
3.6.1 Min-Max Scaling	10
3.6.2 Z-Score Normalization	11
3.7 Principal Component Analysis (PCA)	11
3.7.1 Proses PCA	11
3.7.2 Visualisasi dan Aplikasi PCA	11
BAB IV Metode Penelitian	12
4.1 Analisis Permasalahan	12
4.2 Alur Penyelesaian	12
4.3 Studi Literatur	12
4.4 Data Preparation	13

4.5	Data Preprocessing	13
4.5.1	Feature Selection	13
4.5.2	Normalisasi Dataset	14
4.5.3	Pengukuran Jarak: Euclidean Distance	15
4.5.4	Pengukuran Jarak: Manhattan Distance	16
4.6	Model Kluster Building	16
4.7	Penerapan Model	16
4.8	Analisis dan Kesimpulan	17
4.9	Alat dan Bahan	17
4.9.1	Alat	17
4.9.2	Bahan	18
BAB V Hasil Implementasi	19
5.1	Data Preparation	19
5.2	Data Preprocessing	19
5.3	Model Building	19
5.4	Analisis	21
5.4.1	Pair Plot Analysis	21
5.4.2	3d Visualization Analysis	22
5.4.3	PCA Analysis	23
BAB VI Kesimpulan dan Saran	24
6.1	Kesimpulan	24
6.2	Saran	24
Lampiran A. TOR (Term of Reference)	29
Lampiran B. Log Sheet	32
Lampiran C. Dokumen Teknik	34
Lampiran D. Dokumentasi Kegiatan	35

Daftar Gambar

2.1	Logo KPU Lampung	4
2.2	Struktur KPU Provinsi Lampung	5
4.1	Alur Penyelesaian Penelitian	12
5.1	Hasil Silhouette Score	20
5.2	Hasil Pair Plot	21
5.3	3D Visualization	22
5.4	PCA Analysis	23

Daftar Tabel

4.1	Dataset Rekapitulasi Daftar Pemilih	13
4.2	Dataset 5 Kecamatan	13
4.3	Data Hasil Normalisasi Min-Max pada Lima Kecamatan	15
5.1	Dataset Rekapitulasi Daftar Pemilih	19

Daftar Kode

5.1	Program kode untuk menentukan Silhouette Score	19
-----	--	----

BAB I

Pendahuluan

1.1 Latar Belakang

Demokrasi memberi rakyat hak untuk memilih pemimpin, sebagaimana tercantum dalam UUD 1945 Pasal 1 Ayat 2, bahwa "Kedaulatan berada di tangan rakyat dan dilaksanakan sepenuhnya oleh Majelis Permusyawaratan Rakyat" [1]. Pilkada atau Pemilihan Kepala Daerah di Indonesia merupakan momen penting dalam proses demokrasi, di mana pemimpin di tingkat lokal, seperti gubernur, bupati, dan walikota, dipilih secara langsung oleh masyarakat. Untuk menjamin keberlangsungan pemilihan yang demokratis dan adil, Daftar Pemilih Tetap (DPT) disusun sebagai instrumen resmi yang mencatat warga negara yang memenuhi syarat untuk memberikan suara [2]. Validitas data pemilih memainkan peran penting dalam menghindari permasalahan seperti pemilih ganda atau pemilih tidak sah, yang bisa mengganggu kredibilitas hasil pilkada [3].

Pengelompokan pemilih berdasarkan karakteristik demografis dapat memberikan wawasan penting mengenai perilaku pemilih, distribusi geografis, serta potensi keterlibatan politik di berbagai daerah. Analisis terhadap distribusi usia, jenis kelamin, dan populasi di suatu wilayah dapat membantu memfokuskan kampanye politik di area tertentu yang membutuhkan perhatian lebih [4].

Tantangan dalam menganalisis data pemilih adalah kompleksitas dan volume data yang besar. Metode seperti klasterisasi dalam data mining sangat penting untuk menganalisis data pemilih secara efektif. K-Means merupakan teknik klasterisasi yang efektif dalam mengelompokkan data besar berdasarkan atribut demografis seperti rasio pemilih, tingkat pertumbuhan populasi, dan distribusi usia [5].

Klasterisasi, teknik analisis data yang mengelompokkan data berdasarkan karakteristik serupa [6], mencakup metode populer seperti K-Means. Teknik ini berperan dalam berbagai bidang, membantu menemukan pola tersembunyi dalam data berskala besar. Dalam konteks Pilkada, K-Means mengelompokkan wilayah berdasarkan data pemilih dan populasi, mengidentifikasi area yang membutuhkan perhatian khusus untuk penyuluhan atau penanganan pemilih [7].

Metode K-Means membagi data ke dalam klaster berdasarkan atribut seperti rasio pemilih laki-laki dan perempuan, rasio pemilih terhadap populasi total, serta tingkat pertumbuhan populasi. Teknik ini populer karena kemampuannya menangani data berskala besar dan mengidentifikasi pola demografis penting [8]. Dengan K-Means, penyelenggara Pilkada dapat memetakan daerah berpotensi pemilih tinggi yang kurang terlayani, sekaligus merumuskan strategi yang lebih tepat untuk meningkatkan keterlibatan pemilih.

Analisis klasterisasi data DPT memberikan manfaat strategis bagi penyelenggara. Pengelompokan wilayah berdasarkan pertumbuhan atau kepadatan populasi memungkinkan prioritas pada area dengan jumlah pemilih tinggi dan sumber daya terbatas, serta peningkatan partisipasi politik di wilayah dengan rasio pemilih rendah [8]. Sosialisasi yang

lebih terarah dapat dilaksanakan di daerah-daerah teridentifikasi, sehingga efektivitas kampanye meningkat dan risiko pengabaian suara berkurang [9]. Selain itu, analisis ini mendukung peningkatan partisipasi politik melalui identifikasi wilayah yang membutuhkan sosialisasi pilkada lebih intensif, khususnya area dengan rasio pemilih rendah terhadap jumlah populasi [10].

Penelitian ini menerapkan metode klasterisasi pada data pemilih di beberapa kecamatan untuk mengevaluasi efektivitas pengelompokan dalam mengungkap pola-pola penting. Pendekatan ini diharapkan memberikan kontribusi bagi perencanaan logistik Pilkada dan pemahaman lebih mendalam tentang karakteristik demografis pemilih di berbagai wilayah. Dengan analisis ini, penyelenggara Pilkada dapat merumuskan strategi yang lebih tepat sasaran, mengoptimalkan sumber daya, dan meningkatkan partisipasi politik di area yang memerlukan perhatian lebih.

1.2 Rumusan Masalah

Bagaimana menerapkan K-Means pada dataset rekapitulasi daftar pemilih di Provinsi Lampung.

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk menerapkan metode K-Means pada data pemilih di Provinsi Lampung.

1.4 Manfaat Penelitian

Penelitian ini diharapkan memberikan wawasan mengenai penerapan metode klasterisasi K-Means dalam analisis data pemilih, mempermudah pengelompokan wilayah berdasarkan karakteristik pemilih, serta mendukung pengambilan keputusan yang lebih efektif terkait dengan distribusi dan partisipasi pemilih.

1.5 Ruang Lingkup

Dalam penelitian ini, penulis mengambil objek penelitian dari KPU Provinsi Lampung, dimana ruang lingkup penelitian hanya mencakup kegiatan analisis rekapitulasi data pemilih beserta data umum yang di dapat dari Badan Pusat Statistik (BPS). Penulis mendapat kesempatan untuk membuat sebuah klustering data tersebut dan membuat visualisasi informatif berdasarkan hasil klustering.

1.6 Sistematika Penulisan

Laporan Kerja Praktik ini disusun dari beberapa Bagian (Bab):

- **Bab I:** Pendahuluan, membahas latar belakang, rumusan masalah, tujuan, manfaat, ruang lingkup, dan sistematika penulisan.
- **Bab II:** Gambaran umum instansi, yang menjelaskan profil dan struktur organisasi KPU Provinsi Lampung.
- **Bab III:** Landasan teori, berisi mengenai topik seputar data pemilih dan data BPS, serta metode *Machine Learning* yang berhubungan dengan metode yang digunakan.
- **Bab IV:** Metode penelitian, berisi tentang segala metode yang diterapkan pada penelitian ini serta analisis data yang digunakan.
- **Bab V:** Hasil implementasi, berisi tentang penjabaran hasil dan visualisasi rekapitulasi daftar pemilih.
- **Bab VI:** Kesimpulan dan saran, berisi tentang kesimpulan yang didapat dari Hasil dan Implementasi yang telah dilakukan serta Saran yang diberikan penulis untuk pengembangan terhadap penelitian selanjutnya.

BAB II

Gambaran Umum Instansi

2.1 Profil Organisasi

Komisi Pemilihan Umum (KPU) Provinsi Lampung merupakan lembaga yang bertanggung jawab dalam menyelenggarakan pemilihan umum di tingkat provinsi. KPU Lampung menyelenggarakan pemilu untuk memilih anggota DPR, DPD, DPRD, Presiden dan Wakil Presiden, serta referendum [11].



Figure 2.1: Logo KPU Lampung

2.2 Visi dan Misi Organisasi

Komisi Pemilihan Umum (KPU) Provinsi Lampung memiliki visi dan misi yang dijalankannya. Visi utamanya adalah “Menjadi Penyelenggara Pemilu Serentak yang Mandiri, Profesional dan Berintegritas”. Fokus utamanya adalah mendukung tercapainya organisasi Komisi Pemilihan Umum (KPU) Provinsi Lampung yang mampu melaksanakan tugas dan fungsinya dengan baik, disertai dengan kewibawaan dan kejujuran tanpa dipengaruhi oleh entitas lain dan memberikan layanan terbaik di bidang Pemilihan Umum dan Pemilihan [12]. Selain itu Komisi Pemilihan Umum (KPU) Provinsi Lampung juga memiliki misi yang mencakup hal-hal berikut:

1. Meningkatkan kompetensi penyelenggara Pemilu Serentak dengan berpedoman kepada perundang-undangan dan kode etik penyelenggara pemilu.
2. Melaksanakan peraturan di bidang Pemilu Serentak yang memberikan kepastian hukum, progresif, dan partisipatif.

3. Meningkatkan kualitas penyelenggaraan Pemilu Serentak yang efektif dan efisien, transparan, akuntabel, serta aksesibel.
4. Mengoptimalkan pemanfaatan kemajuan teknologi informasi dalam menyelenggarakan Pemilu Serentak.
5. Meningkatkan partisipasi dan kualitas pemilih dalam Pemilu Serentak.
6. Meningkatkan kualitas pelayanan Pemilu Serentak untuk seluruh pemangku kepentingan.

2.3 Struktur Organisasi

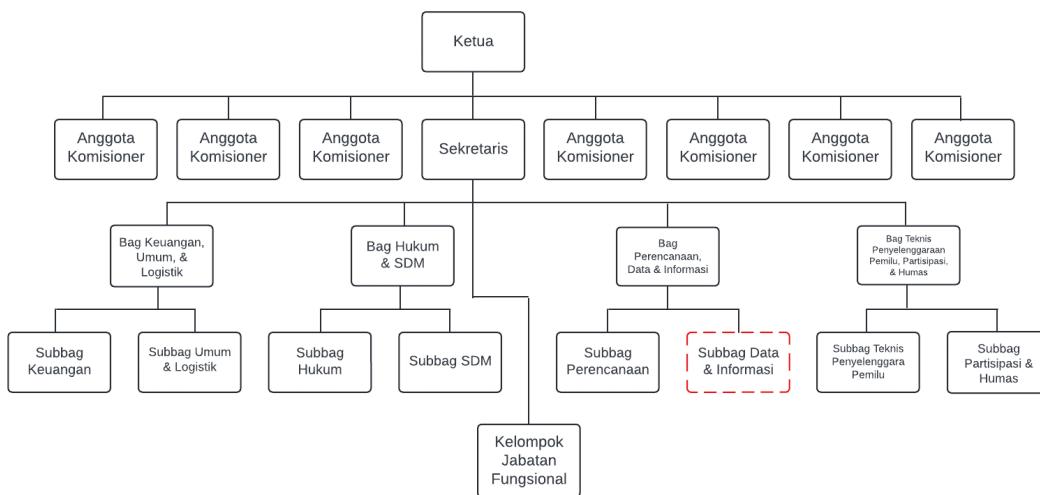


Figure 2.2: Struktur KPU Provinsi Lampung

Struktur organisasi KPU Provinsi Lampung terdiri dari 7 tingkatan, yaitu Ketua KPU Provinsi Lampung, 7 Anggota Komisioner KPU Provinsi Lampung, Sekretaris KPU Provinsi Lampung, Kabag Keuangan, Umum, & Logistik, Kabag Hukum & SDM, Kabag Perencanaan, Data & Informasi, dan Kabag Penyelenggaraan Pemilu, Partisipasi,& Humas. Masing-masing tingkatan memiliki tugas dan fungsi yang berbeda-beda untuk memastikan penyelenggaraan pemilu yang transparan, akuntabel, dan berkualitas.

Ketua KPU Provinsi Lampung memimpin dan mengkoordinasikan pelaksanaan tugas dan fungsi KPU Provinsi Lampung. Anggota KPU Provinsi Lampung membantu Ketua KPU Provinsi Lampung dalam melaksanakan tugas dan fungsi KPU Provinsi Lampung. Sekretaris KPU Provinsi Lampung melaksanakan administrasi dan ketatausahaan KPU Provinsi Lampung. Kabag Keuangan, Umum, & Logistik mengelola keuangan, umum, dan logistik KPU Provinsi Lampung. Kabag Hukum & SDM mengelola hukum dan sumber daya manusia KPU Provinsi Lampung. Kabag Perencanaan, Data & Informasi mengelola perencanaan, data, dan informasi KPU Provinsi Lampung. Kabag Penyelenggaraan Pemilu, Partisipasi, & Humas menyelenggarakan pemilu, partisipasi masyarakat, dan hubungan masyarakat KPU Provinsi Lampung. Pada Praktek Kerja ini Saya ditempatkan di Divisi Data & Informasi.

2.4 Deskripsi Pekerjaan

Dalam pelaksanaan kerja praktik di KPU Provinsi Lampung, penulis melakukan sebuah penelitian tentang klusterisasi pada data rekapitulasi data pemilih untuk Pilkada Provinsi Lampung 2024. Di sisi lain, penulis juga banyak membantu dan belajar mengenai administrasi di bagian Divisi Data dan Informasi KPU Provinsi Lampung.

Lingkup pekerjaan yang dilakukan oleh penulis adalah sebagai berikut:

1. Melakukan kegiatan administrasi seperti membuat Surat Tugas, Surat Pertanggung Jawaban, dll dari kegiatan-kegiatan yang dilakukan oleh Divisi Data dan Informasi.
2. Melakukan pengumpulan dan pengolahan data untuk klusterisasi data rekapitulasi daftar pemilih Pilkada 2024 Provinsi Lampung.

2.5 Jadwal Kerja

Kegiatan kerja praktik dilakukan pada tanggal 19 Juni 2024 – 05 Agustus 2024 di KPU Provinsi Lampung. Pelaksanaan kerja praktik dilakukan selama Lima hari dalam satu minggu, yaitu pada hari Senin hingga hari Jumat. Untuk waktu pelaksanaan kerja praktik dimulai dari pukul 08.00 – 16.00 WIB.

BAB III

Landasan Teori

3.1 Data Mining

Data mining merujuk pada proses mengekstraksi informasi berharga dari kumpulan data besar yang seringkali tidak terstruktur untuk mengungkap pola tersembunyi yang dapat mendukung proses pengambilan keputusan. Tujuan dari data mining adalah mengubah data mentah menjadi wawasan berharga yang dapat digunakan untuk analisis lebih lanjut, perencanaan strategis, dan peramalan. Proses ini melibatkan beberapa tahap, termasuk pembersihan data, integrasi data, transformasi data, penemuan pola, dan evaluasi [7]. Dalam data mining, terdapat beberapa metode untuk mengekstraksi informasi dan pola dari data. Beberapa teknik utama dalam data mining meliputi:

1. **Klasifikasi:** Teknik ini mengkategorikan data ke dalam kelompok yang telah ditentukan sebelumnya. *Supervised learning*, seperti *decision tree*, *support vector machine*, dan *neural network*, biasanya digunakan untuk klasifikasi. Metode ini banyak diterapkan dalam deteksi penipuan, filter spam, dan prediksi penyakit.
2. **Klastering:** Berbeda dengan klasifikasi, klastering adalah teknik pembelajaran tanpa pengawasan yang mengelompokkan data menjadi klaster objek serupa. Klastering sangat berguna dalam segmentasi pasar, deteksi anomali, dan segmentasi citra. Algoritma klastering yang umum digunakan antara lain K-means, DBSCAN, dan hierarchical clustering.
3. **Regresi:** Analisis regresi digunakan untuk memodelkan hubungan antara variabel, biasanya untuk prediksi dan peramalan. Regresi linear, regresi logistik, dan model yang lebih kompleks seperti *decision tree* dan *random forest* sering digunakan dalam tugas data mining prediktif.

3.2 Data Clustering

Clustering adalah teknik dalam data mining yang bertujuan untuk mengelompokkan data menjadi beberapa kelompok atau kluster berdasarkan kemiripan tertentu. Data yang berada dalam satu kluster memiliki kemiripan yang lebih tinggi satu sama lain dibandingkan dengan data di kluster lainnya [13]. Teknik ini sangat penting dalam eksplorasi data, terutama dalam aplikasi seperti segmentasi pasar, deteksi anomali, pengelompokan dokumen, dan analisis pola [14].

3.2.1 K-Means Clustering

K-Means adalah algoritma clustering yang membagi data menjadi k kelompok berdasarkan jarak terdekat antara titik data dan pusat kluster (centroid). Prosesnya melibatkan inisialisasi centroid, penghitungan jarak setiap data terhadap centroid, dan pembaruan posisi centroid sampai konvergen. K-Means sering digunakan karena kesederhanaan dan

efisiensinya dalam memproses data besar. Namun, algoritma ini memiliki keterbatasan, seperti sensitivitas terhadap outlier dan kebutuhan untuk menentukan jumlah kluster sebelumnya [15].

3.2.2 Hierarchical Clustering

Hierarchical clustering membentuk hierarki kluster dengan membangun pohon (dendrogram) yang mengilustrasikan tingkat kemiripan antar data. Terdapat dua pendekatan utama: agglomerative (pendekatan bottom-up) dan divisive (pendekatan top-down) [16]. Dalam pendekatan agglomerative, setiap data dimulai sebagai kluster tunggal, kemudian secara bertahap digabungkan berdasarkan kemiripan, sedangkan dalam pendekatan divisive, kluster besar dipecah menjadi kluster yang lebih kecil. Keuntungan utama dari hierarchical clustering adalah fleksibilitas dalam menentukan struktur kluster, namun memiliki keterbatasan dalam hal efisiensi pada dataset besar [17].

3.2.3 DBSCAN (Density-Based Spatial Clustering)

DBSCAN adalah algoritma clustering yang mengelompokkan data berdasarkan kepadatan. Algoritma ini efektif dalam mengidentifikasi kluster dengan bentuk arbitrer dan dapat mendeteksi outlier (data yang tidak termasuk dalam kluster mana pun). DBSCAN menggunakan dua parameter, yaitu radius pencarian (ϵ) dan jumlah titik minimum dalam radius tersebut ($MinPts$) untuk membentuk kluster. Algoritma ini sangat berguna dalam aplikasi yang melibatkan data spasial dan tidak memerlukan penentuan jumlah kluster sebelumnya, namun performanya sensitif terhadap parameter [18].

3.3 Evaluasi Jarak dalam Clustering

Pemilihan metrik jarak dan ukuran kemiripan sangat berpengaruh pada hasil clustering, terutama pada algoritma yang berbasis jarak seperti K-Means. Berbagai metrik jarak dapat digunakan tergantung pada karakteristik data dan tujuan analisis.

3.3.1 Euclidean Distance

Euclidean Distance merupakan salah satu metrik jarak paling umum yang digunakan dalam clustering, terutama untuk data numerik. Jarak Euclidean antara dua titik dalam ruang n-dimensi (x_i dan x_j) dihitung dengan rumus:

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}$$

Metode ini cocok untuk data dengan fitur berskala serupa, tetapi dapat memberikan hasil yang bias jika data memiliki perbedaan skala antar fitur yang signifikan [19].

3.3.2 Manhattan Distance

Manhattan Distance, juga dikenal sebagai L1 norm atau jarak taksi, menghitung jarak antara dua titik sebagai jumlah dari selisih absolut di setiap dimensi:

$$d(x_i, x_j) = \sum_{k=1}^n |x_{ik} - x_{jk}|$$

Manhattan Distance sering digunakan ketika data memiliki dimensi yang lebih tinggi atau untuk clustering di ruang non-Euclidean [20].

3.3.3 Cosine Similarity

Cosine Similarity mengukur kemiripan antara dua vektor berdasarkan sudut kosinus di antara mereka. Ini umumnya digunakan pada data berdimensi tinggi, seperti teks atau data spasial, di mana orientasi lebih penting daripada magnitude.

$$\text{Cosine Similarity} = \frac{x_i \cdot x_j}{\|x_i\| \times \|x_j\|}$$

Nilai yang mendekati 1 menunjukkan kemiripan yang tinggi, sementara nilai yang mendekati 0 menunjukkan kemiripan rendah. Cosine Similarity populer dalam text mining dan clustering dokumen karena mengabaikan panjang absolut vektor [21].

3.4 Elbow Method

Pemilihan jumlah cluster k yang optimal dalam K-Means adalah masalah yang penting. Salah satu pendekatan populer untuk menentukan jumlah cluster optimal adalah Metode Elbow [22]. Dalam metode ini, grafik *Within-Cluster Sum of Squares* (WCSS) diplot untuk beberapa nilai k , di mana WCSS adalah jumlah dari jarak kuadrat antara titik data dan centroid klusternya masing-masing. Nilai WCSS akan menurun seiring dengan bertambahnya jumlah cluster karena tiap cluster menjadi lebih kecil, tetapi pada titik tertentu, penurunan WCSS menjadi tidak signifikan. Jumlah cluster optimal adalah titik di mana WCSS mulai mengalami penurunan yang tidak signifikan, membentuk "titik siku" atau "elbow". Titik ini mengindikasikan jumlah cluster yang efisien untuk meminimalkan jarak intra-kluster sambil menghindari kompleksitas yang tidak perlu dalam pemodelan [23].

Formula untuk WCSS dalam K-Means adalah:

$$\text{WCSS} = \sum_{i=1}^k \sum_{x \in C_i} |x - \mu_i|^2$$

di mana C_i adalah cluster ke- i , x adalah data dalam cluster tersebut, dan μ_i adalah centroid dari cluster C_i . Dengan memplot nilai WCSS untuk berbagai nilai k , titik di mana terjadi perubahan drastis menuju penurunan yang lebih kecil dapat dilihat sebagai jumlah cluster optimal. Metode ini efektif tetapi subjektif, karena interpretasi "titik siku" tidak selalu jelas untuk semua dataset.

3.5 Silhouette Score

Silhouette Score adalah metrik lain yang sering digunakan untuk mengevaluasi kualitas clustering dan menentukan jumlah cluster yang optimal tanpa bergantung pada grafik visual [24]. Silhouette Score mengukur seberapa mirip sebuah titik data dengan cluster lainnya dibandingkan dengan cluster tempatnya berada saat ini. Metrik ini berkisar dari -1 hingga 1, di mana nilai mendekati 1 menunjukkan bahwa data berada jauh dari cluster yang berdekatan, sedangkan nilai mendekati 0 menunjukkan bahwa data berada di antara dua cluster. Nilai negatif menunjukkan bahwa data tersebut lebih dekat ke cluster tetangga dibandingkan dengan cluster tempatnya berada sekarang, yang menunjukkan pengelompokan yang tidak optimal.

Silhouette Score $s(i)$ untuk data i dapat dihitung dengan rumus:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

di mana: - $a(i)$ adalah jarak rata-rata antara data i dan semua data lain dalam cluster yang sama. - $b(i)$ adalah jarak rata-rata antara data i dan semua data dalam cluster terdekat yang bukan miliknya.

Nilai Silhouette Score rata-rata dari semua data dalam dataset memberikan ukuran keseluruhan dari kualitas clustering. Silhouette Score yang tinggi menunjukkan bahwa kluster bersifat kompak dan data antar kluster berbeda secara signifikan, sehingga metode ini memberikan indikator yang lebih kuantitatif untuk menentukan jumlah cluster yang optimal [25]. Penggunaan Silhouette Score sangat berguna untuk dataset kompleks di mana metode Elbow mungkin sulit diinterpretasi.

3.6 Data Normalization

Sebelum melakukan clustering, terutama untuk algoritma seperti K-Means, normalisasi data seringkali diperlukan untuk memastikan bahwa setiap fitur memiliki kontribusi yang seimbang dalam proses clustering. Algoritma K-Means sensitif terhadap skala antar fitur, sehingga fitur dengan rentang nilai yang lebih besar dapat mendominasi hasil clustering [26]. Dengan melakukan normalisasi, kita dapat menyamakan skala semua fitur sehingga algoritma dapat mengelompokkan data secara lebih akurat.

Beberapa metode normalisasi yang sering digunakan antara lain adalah:

3.6.1 Min-Max Scaling

Min-Max Scaling mengubah data ke rentang skala tertentu, biasanya antara 0 dan 1. Formula Min-Max Scaling adalah:

$$X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

di mana X adalah nilai asli, X_{\min} dan X_{\max} adalah nilai minimum dan maksimum dari fitur tersebut. Metode ini sangat efektif untuk data yang tidak memiliki outlier ekstrem, namun sensitif terhadap perubahan rentang data [27].

3.6.2 Z-Score Normalization

Z-Score Normalization atau Standarisasi mengonversi data berdasarkan mean (μ) dan standar deviasi (ω) dari data, menghasilkan distribusi dengan rata-rata 0 dan standar deviasi 1. Formula Z-Score Normalization adalah:

$$Z = \frac{X - \mu}{\omega}$$

Metode ini bermanfaat ketika data memiliki outlier atau perbedaan rentang nilai antar fitur yang besar. Standarisasi menjadikan semua fitur memiliki variabilitas yang setara sehingga algoritma clustering dapat lebih sensitif terhadap pola dalam data tanpa bias skala [28].

3.7 Principal Component Analysis (PCA)

Pada dataset berdimensi tinggi, analisis dan visualisasi dapat menjadi sulit dan tidak efisien. Oleh karena itu, teknik reduksi dimensi seperti Principal Component Analysis (PCA) sering digunakan untuk menyederhanakan data tanpa menghilangkan informasi yang signifikan. PCA mengurangi jumlah dimensi dengan memproyeksikan data ke dalam sejumlah komponen utama yang menangkap sebagian besar varians dalam data asli [29].

3.7.1 Proses PCA

PCA bekerja dengan mencari kombinasi linier dari fitur asli yang memaksimalkan varians data. Proses PCA melibatkan perhitungan matriks kovarians dari data, ekstraksi nilai eigen (eigenvalues) dan vektor eigen (eigenvectors), dan transformasi data berdasarkan komponen utama yang dihasilkan. Komponen pertama (PC1) menjelaskan variansi terbesar, dan komponen berikutnya (PC2, PC3, dst.) mengurangi dimensi lebih lanjut sambil menjaga sebanyak mungkin informasi [29].

3.7.2 Visualisasi dan Aplikasi PCA

Dengan mereduksi dimensi, PCA memungkinkan kita untuk memvisualisasikan dataset dalam ruang 2D atau 3D, yang berguna untuk interpretasi visual dalam clustering. Selain itu, PCA dapat membantu mengatasi masalah multikolinearitas dalam data yang memiliki fitur yang sangat berkorelasi. Dalam konteks clustering, PCA sering digunakan sebagai pra-pemrosesan untuk mengurangi kompleksitas data sebelum menjalankan algoritma clustering seperti K-Means, sehingga meningkatkan efisiensi dan akurasi [30].

BAB IV

Metode Penelitian

4.1 Analisis Permasalahan

KPU Provinsi Lampung saat ini sedang melaksanakan agenda lima tahunan Pilkada. Sebagai bagian dari persiapan Pilkada 2024, KPU Provinsi Lampung telah melakukan rekapitulasi daftar pemilih. Namun, untuk memahami distribusi dan karakteristik daftar pemilih antar kecamatan di 15 Kabupaten Lampung, diperlukan sebuah analisis yang lebih mendalam. Solusi yang diusulkan adalah penerapan teknik klusterisasi pada data rekapitulasi tersebut, serta penggunaan data opsional lain dari BPS untuk memperkaya analisis. Hasil klusterisasi ini diharapkan dapat memberikan wawasan tambahan melalui visualisasi interaktif yang lebih informatif.

4.2 Alur Penyelesaian

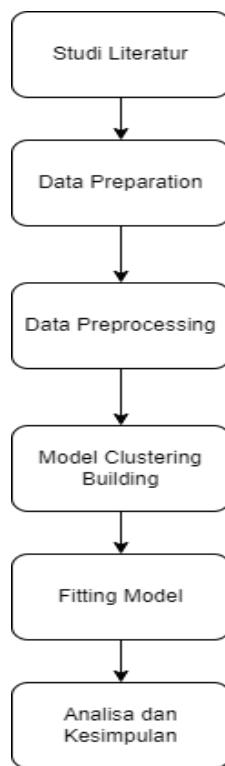


Figure 4.1: Alur Penyelesaian Penelitian

4.3 Studi Literatur

Pada tahap ini, penulis melakukan kajian literatur yang berkaitan dengan klusterisasi data, baik pada data pemilih maupun data umum lainnya. Penelitian yang ada menjadi

dasar untuk memilih metode klusterisasi yang tepat dan teknik validasi model. Selain itu, literatur yang mengkaji metode penentuan jumlah cluster optimal, seperti Elbow Method dan Silhouette Score, menjadi acuan dalam menentukan parameter optimal pada penelitian ini.

4.4 Data Preparation

Persiapan data (*data preparation*) adalah tahap awal sebelum data dapat digunakan dalam model. Pada tahap ini, data masih dalam bentuk mentah dan belum melalui proses normalisasi. Data yang digunakan pada penelitian ini adalah dataset rekapitulasi daftar pemilih untuk Pilkada 2024 Provinsi Lampung dan data tambahan dari BPS untuk seluruh Kecamatan di Provinsi Lampung.

Sub-Region	Male Voters	Female Voters	Total Voters	Voter Gender Ratio	Total Population	Population Growth Rate (%)	Population Density	Eligible Voter Ratio
KEDATON	19196	19657	38853	97.65	52400	-0.17	13896	74.146947
SUKARAME	23913	24623	48536	97.12	67100	0.65	6148	72.333830
TANJUNG KARANG BARAT	22329	22631	44960	98.67	63200	0.72	5476	71.139241
PANJANG	26859	26298	53157	102.13	74900	0.23	5488	70.970628
TANJUNG KARANG TIMUR	14017	14236	28253	98.46	38500	-0.37	18619	73.384416

Table 4.1: Dataset Rekapitulasi Daftar Pemilih

4.5 Data Preprocessing

Setelah data terkumpul, dilakukan tahap praproses data (*data preprocessing*). Pada tahap ini, penulis menentukan atribut yang relevan untuk proses klusterisasi dan melakukan normalisasi data guna memastikan keseragaman skala atribut. Tabel 4.1 berisikan 230 record baris data rekapitulasi per kecamatan di Provinsi Lampung. Proses normalisasi data dilakukan agar nilai dari setiap variable berada pada rentang yang sama, menghindari pengaruh dominasi fitur tertentu pada model klusterisasi.

4.5.1 Feature Selection

Untuk memastikan hasil klusterisasi yang lebih bermakna, dilakukan pemilihan fitur (*feature selection*) terhadap atribut data. Atribut yang dipilih mencakup data demografis seperti usia, jenis kelamin, dan kepadatan populasi, yang dapat memberikan insight terkait pola pemilih di berbagai kecamatan. Atribut-atribut ini ditentukan melalui analisis korelasi untuk menghindari redundansi dan memastikan keterkaitan yang erat dengan tujuan penelitian.

Sub-Region	Eligible Voter Ratio	Voter Gender Ratio	Population Growth Rate (%)	Population Density
Kecamatan 1	0.78	1.05	1.2	1000
Kecamatan 2	0.82	1.10	1.5	1200
Kecamatan 3	0.76	1.02	1.3	950
Kecamatan 4	0.80	1.08	1.6	1100
Kecamatan 5	0.85	1.07	1.4	1050

Table 4.2: Dataset 5 Kecamatan

4.5.2 Normaslisasi Dataset

Setiap fitur memiliki skala yang setara, sehingga menghindari dominasi fitur dengan rentang nilai yang lebih besar dalam proses klusterisasi. Penelitian ini menggunakan teknik normalisasi Min-Max pada fitur *Eligible Voter Ratio*, *Voter Gender Ratio*, *Population Growth Rate (%)*, dan *Population Density*. Normalisasi dilakukan untuk memastikan bahwa Dengan data yang telah dinormalisasi, analisis dapat dilakukan dengan lebih akurat menggunakan metrik jarak seperti Euclidean atau Manhattan distance.

Proses Normalisasi Min-Max Normalisasi dilakukan menggunakan formula Min-Max berikut:

$$X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

Berikut adalah perhitungan normalisasi untuk setiap fitur:

- **Eligible Voter Ratio** (Min: 0.76, Max: 0.85)

$$\begin{aligned}\text{Sub-Region 1 : } & \frac{0.78 - 0.76}{0.85 - 0.76} \approx 0.22 \\ \text{Sub-Region 2 : } & \frac{0.82 - 0.76}{0.85 - 0.76} \approx 0.67 \\ \text{Sub-Region 3 : } & \frac{0.76 - 0.76}{0.85 - 0.76} = 0.0 \\ \text{Sub-Region 4 : } & \frac{0.80 - 0.76}{0.85 - 0.76} \approx 0.44 \\ \text{Sub-Region 5 : } & \frac{0.85 - 0.76}{0.85 - 0.76} = 1.0\end{aligned}$$

- **Voter Gender Ratio** (Min: 1.02, Max: 1.10)

$$\begin{aligned}\text{Kecamatan 1 : } & \frac{1.05 - 1.02}{1.10 - 1.02} \approx 0.38 \\ \text{Kecamatan 2 : } & \frac{1.10 - 1.02}{1.10 - 1.02} = 1.0 \\ \text{Kecamatan 3 : } & \frac{1.02 - 1.02}{1.10 - 1.02} = 0.0 \\ \text{Kecamatan 4 : } & \frac{1.08 - 1.02}{1.10 - 1.02} \approx 0.75 \\ \text{Kecamatan 5 : } & \frac{1.07 - 1.02}{1.10 - 1.02} \approx 0.63\end{aligned}$$

- **Population Growth Rate (%)** (Min: 1.2, Max: 1.6)

$$\text{Kecamatan 1 : } \frac{1.2 - 1.2}{1.6 - 1.2} = 0.0$$

$$\text{Kecamatan 2 : } \frac{1.5 - 1.2}{1.6 - 1.2} = 0.75$$

$$\text{Kecamatan 3 : } \frac{1.3 - 1.2}{1.6 - 1.2} = 0.25$$

$$\text{Kecamatan 4 : } \frac{1.6 - 1.2}{1.6 - 1.2} = 1.0$$

$$\text{Kecamatan 5 : } \frac{1.4 - 1.2}{1.6 - 1.2} = 0.5$$

- **Population Density** (Min: 950, Max: 1200)

$$\text{Kecamatan 1 : } \frac{1000 - 950}{1200 - 950} = 0.2$$

$$\text{Kecamatan 2 : } \frac{1200 - 950}{1200 - 950} = 1.0$$

$$\text{Kecamatan 3 : } \frac{950 - 950}{1200 - 950} = 0.0$$

$$\text{Kecamatan 4 : } \frac{1100 - 950}{1200 - 950} = 0.6$$

$$\text{Kecamatan 5 : } \frac{1050 - 950}{1200 - 950} = 0.4$$

Sub-Region	Eligible Voter Ratio	Voter Gender Ratio	Population Growth Rate (%)	Population Density
Kecamatan 1	0.22	0.38	0.0	0.2
Kecamatan 2	0.67	1.0	0.75	1.0
Kecamatan 3	0.0	0.0	0.25	0.0
Kecamatan 4	0.44	0.75	1.0	0.6
Kecamatan 5	1.0	0.63	0.5	0.4

Table 4.3: Data Hasil Normalisasi Min-Max pada Lima Kecamatan

Tabel 4.3 menunjukkan contoh data normalisasi menggunakan Min-Max Scaling pada lima kecamatan di berdasarkan fitur *Eligible Voter Ratio*, *Voter Gender Ratio*, *Population Growth Rate*, dan *Population Density*.

4.5.3 Pengukuran Jarak: Euclidean Distance

Untuk menghitung jarak antara dua sub-region berdasarkan fitur yang telah dinormalisasi, digunakan Euclidean Distance, yang dapat didefinisikan sebagai berikut:

$$d(x_i, x_j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + (x_{i3} - x_{j3})^2 + (x_{i4} - x_{j4})^2}$$

Sebagai contoh, jarak antara Sub-Region 1 dan Sub-Region 2 adalah:

$$d(1, 2) = \sqrt{(0.22 - 0.67)^2 + (0.38 - 1.0)^2 + (0.0 - 0.75)^2 + (0.2 - 1.0)^2} \approx 1.32$$

4.5.4 Pengukuran Jarak: Manhattan Distance

Selain Euclidean, penelitian ini juga menggunakan Manhattan Distance untuk pengukuran jarak antar data. Rumus Manhattan Distance adalah sebagai berikut:

$$d(x_i, x_j) = |x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + |x_{i3} - x_{j3}| + |x_{i4} - x_{j4}|$$

Sebagai contoh, jarak Manhattan antara Sub-Region 1 dan Sub-Region 2 adalah:

$$d(1, 2) = |0.22 - 0.67| + |0.38 - 1.0| + |0.0 - 0.75| + |0.2 - 1.0| = 0.45 + 0.62 + 0.75 + 0.8 = 2.62$$

4.6 Model Kluster Building

Proses membangun model K-Means melibatkan beberapa tahapan untuk menentukan jumlah cluster optimal. Langkah-langkah dalam tahap ini mencakup:

1. Menentukan rentang nilai k dari $k=2$ hingga $k=n$ untuk proses K-Means.
2. Menerapkan algoritma K-Means pada setiap nilai k dengan berbagai metrik jarak.
3. Menghitung nilai *Silhouette Coefficient* untuk setiap nilai k pada setiap metrik jarak.
4. Menentukan nilai k dengan nilai *Silhouette Coefficient* tertinggi sebagai jumlah cluster optimal untuk setiap metrik jarak.
5. Membandingkan nilai k optimal dari berbagai metrik untuk memilih model terbaik.

Penggunaan beberapa metrik jarak dan metode validasi membantu memastikan bahwa jumlah cluster optimal benar-benar mewakili struktur data.

4.7 Penerapan Model

Setelah jumlah kluster yang sesuai telah ditentukan, data akan dimasukkan ke dalam model klusterisasi yang telah dibangun. Pada tahap ini, setiap kecamatan akan dikelompokkan ke dalam kluster yang sesuai berdasarkan karakteristik demografis dan data pemilih, yang mencakup atribut-atribut penting yang telah melalui tahap praproses dan normalisasi data sebelumnya. Model klusterisasi yang digunakan kemudian diterapkan

pada dataset sehingga setiap wilayah administratif dapat dianalisis menurut kluster yang ditentukan. Visualisasi hasil klusterisasi dilakukan dari berbagai perspektif untuk memperoleh pemahaman yang lebih mendalam mengenai pola distribusi pemilih di setiap kluster. Visualisasi ini mencakup representasi grafis seperti grafik 2D dan 3D, yang memetakan atribut-atribut utama antar kluster untuk melihat perbedaan dan persamaan antar kecamatan dalam satu kluster.

4.8 Analisis dan Kesimpulan

Tahap ini mencakup analisis hasil klusterisasi dan penyusunan kesimpulan berdasarkan model yang telah dibangun. Setelah model klusterisasi diterapkan, penulis akan menganalisis pola-pola yang muncul di setiap kluster, termasuk karakteristik unik dari masing-masing kelompok pemilih yang terbentuk. Analisis ini bertujuan untuk memahami distribusi pemilih antar kecamatan dan menemukan kesamaan atau perbedaan signifikan yang dapat memberikan wawasan lebih dalam mengenai struktur data pemilih di wilayah Lampung.

Penulis juga akan mengevaluasi performa model dengan mengidentifikasi kelebihan dan kekurangannya. Evaluasi ini mencakup penggunaan metrik validasi, seperti nilai Silhouette Score dan pengamatan terhadap visualisasi kluster, untuk memastikan bahwa hasil klusterisasi yang dihasilkan sesuai dengan tujuan penelitian.

4.9 Alat dan Bahan

4.9.1 Alat

Berikut merupakan alat-alat yang digunakan penulis dalam penelitian ini:

1. **Anaconda 3** sebagai package manager untuk mengatur dependensi dalam penelitian ini.
2. **Jupyter Notebook** Sebagai tempat untuk melakukan proses pembuatan model K-Means
3. **Visual Studio Code** digunakan sebagai code editor atau yang sering disebut dengan Integrated Development Environment (IDE).
4. **Acer Aspire 5** dengan spesifikasi chipset Intel i3-7020u, Memory (RAM) 122 GB dan SSD 512 GB.
5. **Git (Version Control System)** digunakan sebagai sistem kontrol versi untuk melacak perubahan kode selama pengembangan. Git memungkinkan kolaborasi yang lebih baik antar pengembang dengan memfasilitasi pengelolaan branch, penggabungan kode (merge), serta pengelolaan riwayat perubahan kode.
6. **Overleaf** Sebagai media untuk penulisan laporan berbasis LateX.

4.9.2 Bahan

Berikut merupakan alat-alat yang digunakan penulis dalam penelitian ini:

1. Dataset rekapitulasi daftar pemilih Pilkada 2024 Provinsi Lampung.
2. Dataset demografis penduduk dari BPS sebagai data pendukung rekapitulasi daftar pemilih.

BAB V

Hasil Implementasi

5.1 Data Preparation

Dalam penelitian ini, digunakan dataset rekapitulasi daftar pemilih dan data tambahan dari BPS untuk menambah wawasan tentang visualisasi yang dilakukan. Dataset disimpan di dalam file csv untuk memudahkan pemrosesan data tahap lanjut.

Sub-Region	Male Voters	Female Voters	Total Voters	Voter Gender Ratio	Total Population	Population Growth Rate (%)	Population Density	Eligible Voter Ratio
KEDATON	19196	19657	38853	97.65	52400	-0.17	13896	74.146947
SUKARAME	23913	24623	48536	97.12	67100	0.65	6148	72.333830
TANJUNG KARANG BARAT	22329	22631	44960	98.67	63200	0.72	5476	71.139241
PANJANG	26859	26298	53157	102.13	74900	0.23	5488	70.970628
TANJUNG KARANG TIMUR	14017	14236	28253	98.46	38500	-0.37	18619	73.384416

Table 5.1: Dataset Rekapitulasi Daftar Pemilih

5.2 Data Preprocessing

Tahap *Preprocessing* merupakan tahapan dimana data yang telah diambil sebelumnya akan dilakukan pengolahan lebih lanjut sebelum digunakan sebagai data untuk proses clustering. Bisa dilihat dari table 5.1 terdapat beberapa data mentah seperti jumlah pemilih laki-laki dan perempuan, maka dibuat atribut data baru *Voters Gender Ratio* sebagai atribut data baru dari kedua data sebelumnya untuk memudahkan analisis. Hasil akhir yang digunakan adalah 4 atribut data berupa *Eligible Voter Ratio*, *Voter Gender Ratio*, *Kepadatan Penduduk* dan *Percentase Pertumbuhan Penduduk*. Alasan penggunaan data tersebut adalah karena ke-4 atribut tersebut itu menggunakan data komposit dari data mentah sebelumnya.

1. Eligible Voter Ratio = $\frac{\text{Total Voter}}{\text{Total Populasi}}$
2. Voter Gender Ratio = $\frac{\text{Male Voter}}{\text{Female Voter}} * 100$
3. Kepadatan Penduduk = Hasil Kepadatan Penduduk (data dari BPS)
4. Persentase Pertumbuhan Penduduk = Hasil Persentase Pertumbuhan Penduduk (data dari BPS)

5.3 Model Building

Akan dibuat model clustering dan proses visualisasi dari 4 atribut ini dari berbagai macam sudut pandang, namun hal peratama yang perlu dilakukan adalah menentukan jumlah *cluster* yang sesuai untuk dataset kali ini.

```
1 ## Install the Necessary Package
2 import numpy as np
3 import pandas as pd
4 import matplotlib.pyplot as plt
```

```

5
6 ## Open and read the csv data
7 file_path = "dataset-rekapitulasi-dpt-kecamatan.csv"
8 df = pd.read_csv(file_path)
9
10 # Display the first few rows to understand the structure
11 df.head()
12
13 # Select the features for clustering
14 data = df[['Eligible Voter Ratio', 'Population Growth Rate (%)', 'Population
15 Density', 'Voter Gender Ratio']]
16
17 # Normalize the data
18 scaler = StandardScaler()
19 data_scaled = scaler.fit_transform(data)
20
21 # Perform KMeans can calculate SSE (Sum Squared Error) for K values
22 K_range = range(2, 10)
23 silhouette_scores = []
24
25 for k in K_range:
26     kmeans = KMeans(n_clusters=k, random_state=42)
27     kmeans.fit(data_scaled)
28     cluster_labels = kmeans.predict(data_scaled)
29     silhouette_avg = silhouette_score(data_scaled, cluster_labels)
30     silhouette_scores.append(silhouette_avg)
31
32 plt.plot(K_range, silhouette_scores, marker='o')
33 plt.title('Silhouette Score for Optimal K')
34 plt.xlabel('Number of clusters (K)')
35 plt.ylabel('Silhouette Score')
36 plt.show()

```

Kode 5.1: Program kode untuk menentukan Silhouette Score

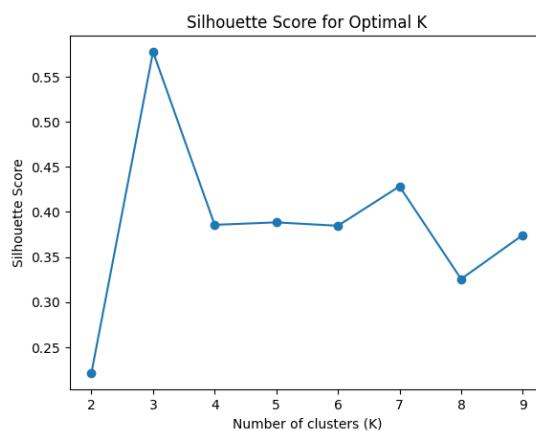


Figure 5.1: Hasil Silhouette Score

Berdasarkan hasil tersebut didapatkan bahwa nilai kluster yang optimal adalah 3, dimana dapat dilihat dari nilai *Silhouette Score*, Cluster 3 mempunyai nilai paling maximum. Nilai *Silhouette Score* menjelaskan kemiripan dari suatu object di dalam suatu *cluster*.

dengan *cluster* lain. Nilai tinggi (mendekati 1) mengindikasikan *cluster* yang lebih baik, sementara nilai rendah (mendekati 0) mengindikasikan adanya *cluster* yang saling tumpang tindih (*overlap*). Nilai 0.56 menunjukkan *cluster* yang cukup baik, namun terdapat beberapa *overlap* antar klaster.

5.4 Analisis

Hasil dari *Model Building* dan ketika dilakukan klusterisasi data rekapitulasi daftar pemilih, disini akan dijelaskan tentang hasil analisis dari masing-masing visualisasi.

5.4.1 Pair Plot Analysis

Metode analisis yang membandingkan nilai dari masing-masing atribut dan melihat sifat alami dari atribut data tersebut.

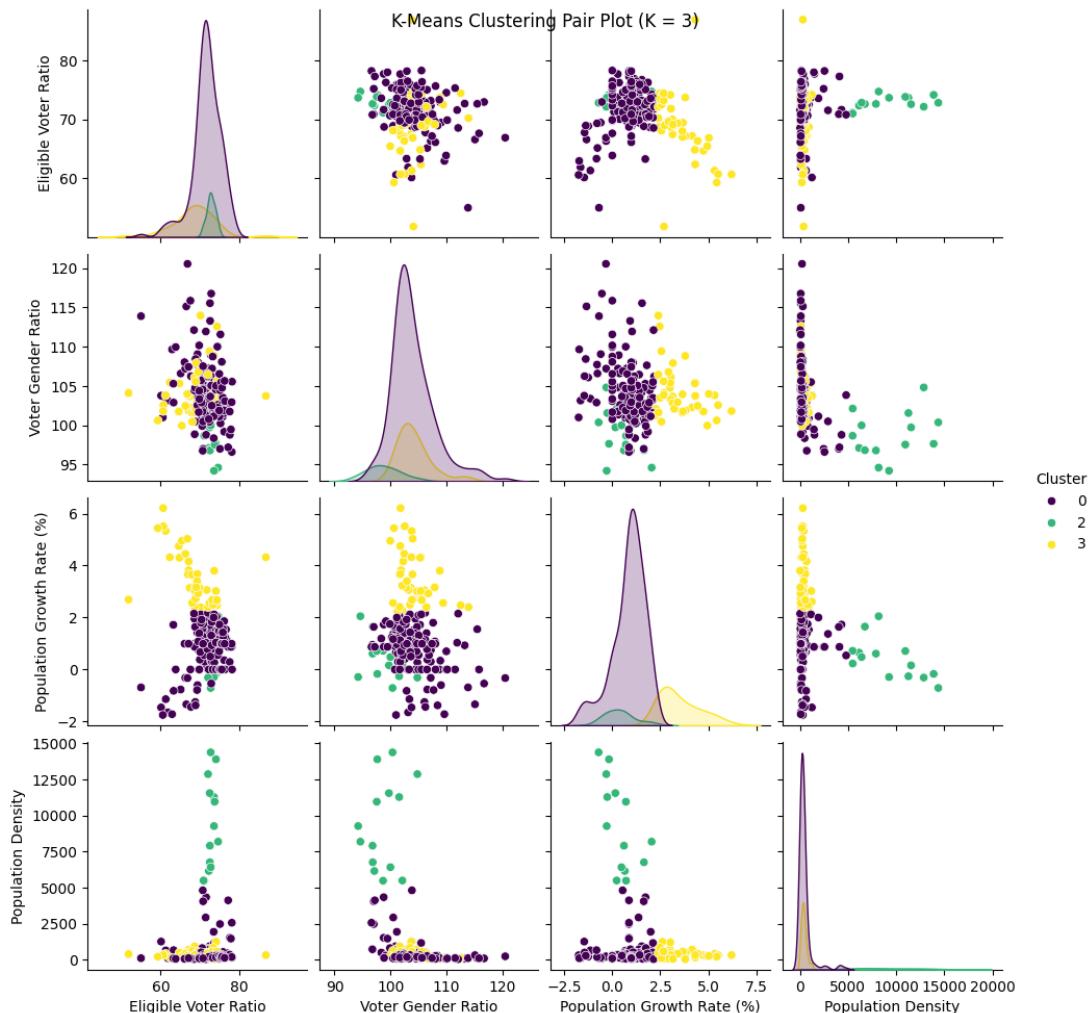


Figure 5.2: Hasil Pair Plot

Hasil analisis dari 5.2 menunjukkan bahwa interaksi antar atribut data, dari visualisasi

ini juga didapatkan bahwa terdapat homogenitas serta kurangnya standar deviasi (persebaran data) dalam beberapa atribut data sehingga didapatkan grafik yang tumpang-tindih (*overlapped*).

5.4.2 3d Visualization Analysis

Metode visualisasi dengan menggunakan visualisasi 3 dimensi untuk mendapatkan perspektif baru dari penelitian ini.

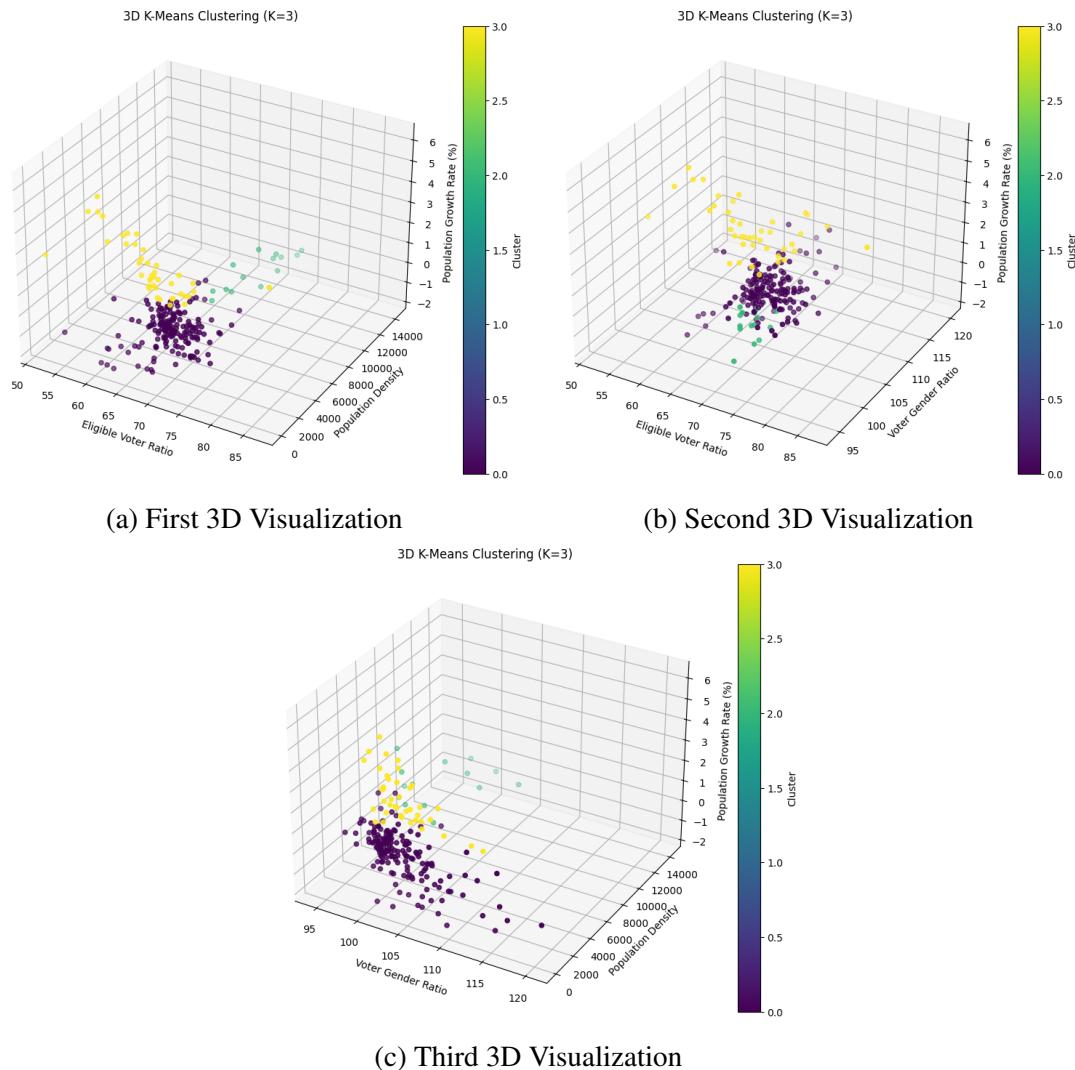


Figure 5.3: 3D Visualization

Pada visualisasi ini didapatkan hasil bahwa visualisasi 3D memberikan hasil yang lebih baik dibandingkan dengan *Pair Plot analysis*. Namun karena sifat alami dari data tersebut, masih terdapatnya Homogenitas data, dimana atribut data *voter ratio* dan *voter gender ratio*, menunjukkan adanya keterkaitan antar dua atribut sehingga data tersebut saling tumpang-tindih *overlap*.

5.4.3 PCA Analysis

Secara keseluruhan, proyeksi PCA ini menunjukkan bahwa Cluster 0 dan 1 tidak mudah dipisahkan dan memiliki tumpang tindih yang cukup besar, yang memperkuat homogenitas data dalam atribut tertentu. Variansi lebih banyak tertangkap pada Komponen Utama 1, yang kemungkinan dipengaruhi oleh atribut dengan variansi lebih tinggi seperti Tingkat Pertumbuhan Populasi dan Kepadatan Populasi, sementara Komponen Utama 2 tidak berkontribusi banyak pada pemisahan cluster.

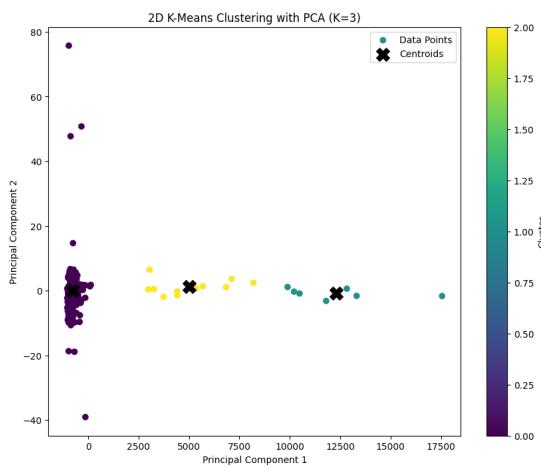


Figure 5.4: PCA Analysis

BAB VI

Kesimpulan dan Saran

6.1 Kesimpulan

Dari ketiga visualisasi yang dilakukan, baik dalam bentuk pair plot, 3D scatter plot, dan PCA 2D plot, dapat disimpulkan bahwa data yang dianalisis menunjukkan pola yang cukup homogen dengan variasi yang terbatas di beberapa atribut. Pada pair plot (visualisasi pertama), terlihat bahwa beberapa atribut seperti *Voter ratio* dan *voter gender ratio* memiliki distribusi yang sangat rapat dan tumpang tindih antar cluster, yang mengindikasikan sulitnya pemisahan yang jelas antar cluster dalam dimensi-dimensi tersebut. Pada 3D scatter plot (visualisasi kedua), meskipun terdapat sedikit pemisahan antar cluster, terutama pada atribut *growth population* dan *population density*, overlap antar cluster masih terlihat, terutama antara Cluster 0 (ungu) dan Cluster 1 (kuning). Ini menunjukkan bahwa atribut-atribut yang digunakan tidak memberikan pemisahan yang sangat kuat dalam ruang dimensi tinggi. Dalam PCA 2D plot (visualisasi ketiga), terlihat bahwa sebagian besar variabilitas dalam data ditangkap oleh Principal Component 1, sementara Principal Component 2 hanya menangkap sedikit variansi. Hal ini mengindikasikan bahwa sebagian besar atribut tidak memberikan cukup variasi untuk memisahkan cluster dengan tegas, dengan sebagian data cenderung tersebar di satu area tertentu.

Hasil dari nilai *Silhouette Score* juga menunjukkan nilai 0.56, yang menandakan bahwa kualitas pengelompokan dalam cluster ini cukup baik. Nilai *Silhouette Score* yang berkisar antara -1 hingga 1 ini dapat digunakan untuk mengevaluasi seberapa baik titik-titik data dikelompokkan. Nilai positif, terutama yang mendekati 1, menunjukkan bahwa data berada pada cluster yang tepat dan cukup jauh dari cluster lainnya. Pada nilai 0.56, hal ini berarti sebagian besar titik data memang berada dalam kelompok yang sesuai, meskipun terdapat beberapa titik yang overlap atau sulit dipisahkan dengan jelas. Overlap ini terlihat pada ketiga visualisasi yang dihasilkan, yang mungkin disebabkan oleh kemiripan karakteristik antar cluster di area tertentu. Namun, secara keseluruhan, hasil ini tetap menunjukkan kualitas clustering yang solid.

6.2 Saran

Berdasarkan penelitian yang sudah dilaksanakan, terdapat beberapa saran untuk dapat memaksimalkan klustering rekapitulasi daftar pemilih Pilkada 2024, diantaranya:

- Penambahan Atribut: Disarankan untuk menambahkan atribut lain yang lebih variatif, terutama yang dapat menangkap aspek yang mungkin lebih spesifik terhadap perbedaan antar wilayah atau sub-kelompok dalam data. Atribut seperti pendapatan per kapita atau tingkat pendidikan mungkin dapat memberikan dimensi baru yang memperjelas perbedaan antar cluster.
- Eksplorasi Metode Clustering Lain: Metode clustering seperti DBSCAN atau Hierarchical Clustering dapat dieksplorasi untuk melihat apakah metode ini lebih

efektif dalam mendeteksi struktur atau pola dalam data yang tidak dapat ditangkap dengan baik oleh K-Means, terutama jika terdapat outlier atau distribusi yang tidak seragam.

- Evaluasi Outlier: Penting untuk melakukan evaluasi lebih mendalam terhadap outlier yang ada. Outlier tersebut mungkin memberikan informasi penting atau justru mengganggu proses clustering. Penghapusan atau penanganan outlier dengan lebih baik bisa meningkatkan hasil analisis.
- Pengurangan Dimensi Lanjutan: Selain PCA, metode pengurangan dimensi lain seperti t-SNE atau UMAP dapat dicoba untuk mendapatkan visualisasi yang lebih baik dan memperlihatkan pola yang mungkin tidak terdeteksi pada PCA 2D.

Daftar Pustaka

- [1] Kumparan. Isi pasal 1 ayat 2 uud1945 sebelum dan sesudah amandemen. [Online]. Available: <https://kumparan.com/berita-terkini/isi-pasal-1-ayat-2-uud1945-sebelum-dan-sesudah-amandemen-1weOIUtUKWB>
- [2] KPU. Dptpilkada. [Online]. Available: <https://www.kpu.go.id/berita/baca/11702/dpt-pemilu-2024-nasional-2048-juta-pemilih>
- [3] K. P. Umum. Pentingnya daftar pemilih tetap yang valid untuk pemilu yang demokratis. [Online]. Available: kpu.go.id
- [4] L. Ointu, V. Rotty, and F. Mamonto, “Implementasi program pemutakhiran data pemilih berkelanjutan di kota manado,” *SIBATIK JURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*, vol. 1, pp. 2969–2976, 11 2022.
- [5] Y. Li and H. Wu, “A clustering method based on k-means algorithm,” *Physics Procedia*, vol. 25, pp. 1104–1109, 2012, international Conference on Solid State Devices and Materials Science, April 1-2, 2012, Macao. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1875389212006220>
- [6] F. Nie, Z. Li, R. Wang, and X. Li, “An effective and efficient algorithm for k-means clustering with new formulation,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 4, pp. 3433–3443, 2023.
- [7] J. Han and M. Kamber, “Data mining: concepts and techniques morgan kaufmann,” vol. 54, 01 2006.
- [8] A. S. Ahmar, D. Napitupulu, R. Rahim, R. Hidayat, Y. Sonatha, and M. Azmi, “Using k-means clustering to cluster provinces in indonesia,” *Journal of Physics: Conference Series*, vol. 1028, no. 1, p. 012006, jun 2018. [Online]. Available: <https://dx.doi.org/10.1088/1742-6596/1028/1/012006>
- [9] G. Fajriansyah, “Analisis daftar pemilih tetap pada hasil rekapitulasi kpu berdasarkan usia menggunakan algoritma k-means (studi kasus : Kota bandar lampung),” *Electrician*, vol. 15, pp. 39–53, 01 2021.
- [10] J. Partheymüller, W. C. Mueller, A. Rabitsch, M. Lidauer, and P. Grohma, “Participation in the administration of elections and perceptions of electoral integrity,” *Electoral Studies*, vol. 77, 05 2022.
- [11] Annisa. Komisi pemilihan umum (kpu), tugas dan wewenangnya. [Online]. Available: [KomisiPemilihanUmum\(KPU\),TugasdanWewenangnya](#)
- [12] KPU. Visi dan misi kpu. [Online]. Available: <https://www.kpu.go.id/page/read/4/visi-dan-misi>
- [13] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: a review,” *ACM Comput. Surv.*, vol. 31, no. 3, p. 264–323, Sep. 1999. [Online]. Available: <https://doi.org/10.1145/331499.331504>

- [14] G. J. Oyewole and G. A. Thopil, “Data clustering: application and trends,” *Artificial Intelligence Review*, vol. 56, no. 7, pp. 6439–6475, 2023. [Online]. Available: <https://doi.org/10.1007/s10462-022-10325-y>
- [15] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” 1967. [Online]. Available: <https://api.semanticscholar.org/CorpusID:6278891>
- [16] S. C. Johnson, “Hierarchical clustering schemes,” *Psychometrika*, vol. 32, no. 3, pp. 241–254, 1967.
- [17] S. Patel, S. Sihmar, and A. Jatain, “A study of hierarchical clustering algorithms,” in *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACoM)*, 2015, pp. 537–541.
- [18] D. Deng, “Dbscan clustering algorithm based on density,” in *2020 7th International Forum on Electrical Engineering and Automation (IFEEA)*, 2020, pp. 949–953.
- [19] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: A review,” *ACM Computing Surveys*, vol. 31, no. 3, pp. 264–323, 1999.
- [20] C. C. Aggarwal and C. Procopiuc, “Fast algorithms for projected clustering,” in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 2001, pp. 61–72.
- [21] A. Huang, “Similarity measures for text document clustering,” in *Proceedings of the New Zealand Computer Science Research Student Conference*, 2008.
- [22] E. Umargono, J. Suseno, and S. Gunawan, “K-means clustering optimization using the elbow method and early centroid determination based on mean and median formula,” 01 2020.
- [23] T. M. Kodinariya and P. R. Makwana, “Review on determining number of cluster in k-means clustering,” *International Journal of Advanced Research in Computer Science and Management Studies*, vol. 1, no. 6, pp. 90–95, 2013.
- [24] Y. Januzaj, E. Beqiri, and A. Luma, “Determining the optimal number of clusters using silhouette score as a data mining technique,” *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 19, no. 04, p. pp. 174–182, Apr. 2023. [Online]. Available: <https://online-journals.org/index.php/i-joe/article/view/37059>
- [25] K. R. Shahapure and C. Nicholas, “Cluster quality analysis using silhouette score,” in *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, 2020, pp. 747–748.
- [26] H. Henderi, T. Wahyuningsih, and E. Rahwanto, “Comparison of min-max normalization and z-score normalization in the k-nearest neighbor (knn) algorithm to test the accuracy of types of breast cancer,” *International Journal of Informatics and Information Systems*, vol. 4, no. 1, pp. 13–20, 2021. [Online]. Available: <https://ijiiis.org/index.php/IJIIIS/article/view/73>

- [27] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Waltham, MA, USA: Morgan Kaufmann, 2011.
- [28] S. García, J. Luengo, and F. Herrera, *Data Mining and Knowledge Discovery Handbook*. Springer, 2015.
- [29] M. Greenacre, P. J. F. Groenen, T. Hastie, A. I. D'Enza, A. Markos, and E. Tuzhilina, "Principal component analysis," *Nature Reviews Methods Primers*, vol. 2, no. 1, p. 100, 2022. [Online]. Available: <https://doi.org/10.1038/s43586-022-00184-w>
- [30] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, 2010.

Lampiran A. TOR (Term of Reference)

A.1 Latar Belakang

Pilkada atau Pemilihan Kepala Daerah di Indonesia merupakan momen penting dalam proses demokrasi, di mana pemimpin di tingkat lokal, seperti gubernur, bupati, dan walikota, dipilih secara langsung oleh masyarakat. Untuk menjamin keberlangsungan pemilihan yang demokratis dan adil, Daftar Pemilih Tetap (DPT) disusun sebagai instrumen resmi yang mencatat warga negara yang memenuhi syarat untuk memberikan suara. Validitas data pemilih memainkan peran penting dalam menghindari permasalahan seperti pemilih ganda atau pemilih tidak sah, yang bisa mengganggu kredibilitas hasil pilkada.

Pengelompokan pemilih berdasarkan karakteristik demografis dapat memberikan wawasan lebih dalam bagi penyelenggara pilkada dan kandidat mengenai perilaku pemilih, distribusi geografis, serta potensi keterlibatan politik di setiap daerah. Namun, tantangan utama dalam menganalisis data pemilih adalah kompleksitas serta volume data yang besar. Oleh karena itu, data mining dan metode analisis data, seperti klasterisasi, memainkan peran penting dalam menganalisis data pemilih dengan lebih efektif.

Melalui penelitian ini, metode klasterisasi diterapkan pada data pemilih di beberapa kecamatan dengan tujuan untuk mengevaluasi bagaimana data ini dapat diolah secara lebih efektif. Analisis yang dilakukan diharapkan mampu mengungkap pola-pola penting yang dapat berkontribusi dalam perencanaan logistik pilkada serta dalam memahami karakteristik demografis pemilih di berbagai wilayah.

A.2 Tujuan Pekerjaan

Tujuan Kerja Praktek (KP) di KPU Provinsi Lampung adalah sebagai berikut:

1. Melakukan penelitian ini adalah bagaimana penerapan metode clustering (K-Means) pada data pemilih.
2. Membantu penyelesaian tugas harian yang diberikan dibidang oleh bagian divisi datin dan informasi KPU Provinsi Lampung
3. Belajar dalam memahami lingkungan kerja dan implementasi bidang keahlian program studi pada lingkungan kerja

A.3 Lingkup Pekerjaan

Lingkup pekerjaan Kerja Praktek di KPU adalah sebagai berikut:

1. Kerja Praktek dilaksanakan dalam waktu 35 hari dimulai dari tanggal 19 Juni– 05 Agustus 2024.

2. Lingkup pekerjaan di KPU Provinsi Lampung Subbag Data dan Informasi yang dilakukan oleh mahasiswa selama menjadi KP, adalah membantu kegiatan administrasi dan surat-menurat.
3. Penelitian mencakup hanya data rekapitulasi daftar pemilih dan data tambahan dari Badan Pusat Statistik (BPS)

A.4 Metodologi

Metodologi yang digunakan dalam penelitian ini adalah dengan menggunakan K-Means sebagai metode clustering dan dataset yang digunakan sebagai bahan adalah data rekapitulasi daftar pemilih dan data pemilih tambahan dari Badan Pusat Statistik.

A.5 Hasil Pekerjaan

Hasil pekerjaan penulis di KPU Provinsi Lampung di bagian Data dan Informasi sebagai berikut:

1. Melakukan penginputan dan update desain progress Coklit KPU Provinsi Lampung
2. Melakukan kegiatan administrasi seperti membuat Surat Tugas, Surat Dinas yang digunakan oleh divisi Data dan Informasi.
3. Melakukan penelitian dan pengumpulan data clustering data daftar pemilih.
4. Membuat model clustering untuk mengelompokan data pemilih.

A.6 Jadwal Kerja

Jadwal Kerja pada Kerja Praktek (KP) di KPU Provinsi Lampung bagian divisi Data dan Informasi adalah:

- Hari Kerja: Senin - Jum'at
- Jam Kerja : 08:00 - 16:00
- Tanggal Mulai : 19 Juni 2024
- Tanggal Selesai : 05 Agustus 2024

Terms of reference ini telah dibaca dan disetujui oleh:

Pihak Mahasiswa



Arsyadana Estu Aziz

121140068

Pihak Instansi



Nurwafa Finanda, S.E., M.M. / Staff Bagian Data dan Informasi

Nomor Pegawai -

Pihak Instansi



Ressy Silvia Dwi, S.E., M.Kn / Kepala Subbagian Data dan Informasi

NIP. 198306092009022005

Lampiran B. Log Sheet



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN,
RISET, DAN TEKNOLOGI
INSTITUT TEKNOLOGI SUMATERA
FAKULTAS TEKNOLOGI INDUSTRI
Jalan Terusan Ryacudu Way Hui, Kecamatan Jati Agung, Lampung Selatan 35365
Telepon: (0721) 8030188
Email : itpi@itera.ac.id Website : <http://itera.ac.id>

PROGRAM STUDI TEKNIK INFORMATIKA

Log Sheet Praktek Kerja Lapangan

Saya yang bertanda tangan di bawah ini,

Nama : Ressy Silvia Dewi, S.E., M.M.
Jabatan : Kasubbag Data dan Informasi KPU Provinsi Lampung.
Perusahaan/Instansi : KPU Provinsi Lampung

dengan ini menyatakan mahasiswa berikut:

Nama : Arsyadana Estu Aziz
NIM : 121140068
Topik PKL : Analisis Rekapitulasi Data Pemilih Tetap Pilkada 2024
Provinsi Lampung Menggunakan K-Means Clustering
Pembimbing Lapangan : Ressy Silvia Dewi, S.E., M.M./ Kasubbag Data dan Informasi
/Jabatan

Telah melaksanakan Praktek Kerja Lapangan di instansi terkait selama 35 hari kerja, dengan daftar kehadiran sebagai berikut :

Hari	Hari / Tanggal	Kegiatan	Hari	Hari / Tanggal	Kegiatan
1	Rabu, 19 Juni 2024	Agenda Perkenalan dengan Lingkungan Kerja	18	Jumat, 12 Juli 2024	Senam Pagi, Update Progress Coklit
2	Kamis, 20 Juni 2024	Design Nametag, membantu administrasi Divisi Datin	19	Senin, 15 Juli 2024	Izin Sakit
3	Jumat, 21 Juni 2024	Agenda Senam Pagi, Membantu Persiapan dan Dokumentasi Rapat Koordinasi	20	Selasa, 16 Juli 2024	Bantu Administrasi
4	Senin, 24 Juni 2024	Apel Rutin Mingguan, Diskusi Mengenai Project.	21	Rabu, 17 Juli 2024	Bantu Administrasi, Improving Clean Code
5	Selasa, 25 Juni 2024	Diskusi Terkait Metode Pengumpulan Project	22	Kamis, 18 Juli 2024	Bantu Administrasi, Finalisasi Progress Coklit
6	Rabu, 26 Juni 2024	Membantu Administrasi di Divisi Datin	23	Jumat, 19 Juli 2024	Bantu Agenda Rapat Koordinasi Penyusunan DPHP di Hotel Emersia
7	Kamis, 27 Juni 2024	Membantu Administrasi di Divisi Datin	24	Senin, 22 Juli 2024	Bantu Adminstrasi
8	Jumat, 28 Juni 2024	Pengumpulan Data Rekapitulasi Daftar Pemilih	25	Selasa, 23 Juli 2024	Bantu Administrasi, Metode Preprocessing Data
9	Senin, 1 Juli 2024	Apel Rutin Mingguan, Membantu	26	Rabu, 24 Juli 2024	Bantu Administrasi



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN,

RISET, DAN TEKNOLOGI

INSTITUT TEKNOLOGI SUMATERA

FAKULTAS TEKNOLOGI INDUSTRI

Jalan Terusan Ryacudu Way Hui, Kecamatan Jati Agung, Lampung Selatan 35365

Telepon: (0721) 8030188

Email : jtpi@itera.ac.id Website : <http://itera.ac.id>

		Administrasi Divisi Datin			
10	Selasa, 2 Juli 2024	Izin Cuti, Agenda Bangkit (MBKM)	27	Kamis, 25 Juli 2024	Bantu Administrasi, Pencatatan Sertifikat Pantarlih.
11	Rabu, 3 Juli 2024	Desain Infografis Update Coklit	28	Jumat, 26 Juli 2024	Pembantuan Pembagian dan Pencatatan Sertifikat Pantarlih
12	Kamis, 4 Juli 2024	Desain Infografis Update Coklit, Bantu Administrasi di Divisi Datin	29	Senin, 29 Juli 2024	Membantu Administrasi, Penginputan Data Datin
13	Jumat, 5 Juli 2024	Agenda Senam Pagi, Membantu Persiapan Stand di Acara K-Fest 2024	30	Selasa, 30 Juli 2024	Mengerjakan Laporan KP
14	Senin, 8 Juli 2024	Metode Pengumpulan Data (DATA BPS)	31	Rabu, 31 Juli 2024	Mengerjakan Laporan KP
15	Selasa, 9 Juli 2024	Membantu Administrasi di Divisi Datin	32	Kamis, 1 Agustus 2024	Mengerjakan Laporan KP
16	Rabu, 10 Juli 2024	Desain Infografis Update Coklit, Bantu Administrasi.	33	Jumat, 2 Agustus 2024	Membantu Administrasi Pelaporan SPJ Rapat Koordinasi
17	Kamis, 11 Juli 2024	Bantu Administrasi Divisi Datin	34	Senin, 5 Agustus 2024	Agenda Perpisahan dengan Divisi Data dan Informasi

Bandar Lampung, 5 Agustus 2024

Kasubbag Data dan Informasi KPU Provinsi Lampung



Lampiran C. Dokumen Teknik

Untuk *source code*, baik dari analisis program dan *dataset* bisa dilihat melalui situs ini.

- <https://github.com/archiseino/Vox-KMeans-Cluster>

Lampiran D. Dokumentasi Kegiatan

1. Minggu 1 (19 – 28 Juni 2024) Kegiatan Analisis Permasalahan



2. Minggu 2 (1 – 5 Juli 2024) Kegiatan Memperbaiki dan Belajar Struktur Data di Server



3. Minggu 3 (8 – 12 Juli 2024) Kegiatan Sosialisasi Edukasi dan Observasi Data Pemilih di Event Krakatau Fest.



4. Minggu 4 (15 – 19 Juli 2024) Kegiatan melihat preprocessing data rekapitulasi daftar pemilih dalam Rapat Koordinasi



5. Minggu 5 (22 – 26 Juli 2024) Kegiatan Optimalisasi Model Clustering.



6. Minggu 6 (29 Juli – 5 Agustus 2024) Kegiatan Presentasi Project dan Perpisahan dengan Bagian Datin dan KPU Provinsi Lampung

