

TWO PHASE AGGLOMERATIVE HIERARCHICAL CLUSTERING (AHC) DALAM PENENTUAN LOKASI

Tjutju Tarlih Dimyati¹⁾, Andi Firmansyah²⁾,

^{1&2)}Program Pascasarjana Magister Teknik Industri, Universitas Pasundan

Email²⁾: firmansyah0401@gmail.com

ABSTRAK

Industri Tekstil saat ini menjadi salah satu industri dengan jumlah perusahaan terbanyak di Indonesia, setelah industri makanan. Berdasarkan arahan dari Departemen Perindustrian Republik Indonesia, Kawasan Industri tekstil di Jawa Barat direncanakan akan berada di kabupaten Majalengka yang memiliki 26 kecamatan dengan karakteristik masing-masing.

Karena pemilihan lokasi suatu industri akan sangat berpengaruh bagi keberhasilan industri tersebut, maka persoalan ini harus diselesaikan dengan mempertimbangkan banyak faktor atau kriteria serta melibatkan sejumlah kandidat lokasi. Untuk memperoleh solusi yang handal, pada penelitian ini persoalan tersebut diselesaikan dengan menggunakan metode Agglomerative Hierarchical Clustering (AHC).

Penelitian dilakukan dalam dua phase. Phase pertama digunakan untuk mencari kombinasi metode terbaik antara metode pengukuran jarak ketidaksamaan dengan metode penggabungan atau klusterisasi. Ada delapan metode pengukuran jarak dan empat metode penggabungan yang diteliti. Hasil yang diperoleh dari kombinasi metode tersebut dievaluasi berdasarkan nilai simpangan baku yang terjadi. Kombinasi dengan nilai simpangan baku terkecil selanjutnya digunakan pada phase kedua yaitu menentukan klaster dan anggota masing-masing klaster.

Hasil penelitian menunjukkan bahwa kombinasi terbaik adalah metode City Block Distance dengan metode Average Linkage, dengan nilai simpangan baku sebesar 0.078. Adapun klaster yang terbentuk adalah: Klaster 1, terdiri dari Lemahsugih, Bantarujeg, Malausma, Cingambul, Talaga, Argapura, Maja, Majalengka, Sukahaji, Rajagaluh, Sindangwangi, Leuwimunding, Palasah, Dawuan, Kasokandel, Panyingkiran, Kadipaten, Jatitujuh, Ligung, dan Sumberjaya. Klaster 2 terdiri dari Cikijing, Banjaran, dan Cigasong. Klaster 3 terdiri dari Sindang, klaster 4 terdiri dari Jatiwangi, dan klaster 5 yang terdiri dari Kertajati.

Untuk menetapkan kecamatan mana yang akan terpilih sebagai lokasi industri tekstil ini, diperlukan informasi dan kajian lebih lanjut dengan melibatkan pemerintah daerah setempat. Karena itu penelitian yang dilakukan dibatasi hanya sampai penetapan klaster dengan anggota masing-masing.

Kata Kunci : Klustering, AHC, Tekstil, Dissimilarity, Pemilihan Lokasi, Industri

1. PENDAHULUAN

Pembangunan ekonomi pada suatu negara memiliki tujuan untuk dapat meningkatkan kesejahteraan serta meningkatkan taraf hidup masyarakatnya. Dengan kuatnya sektor ekonomi tentunya akan berpengaruh pada stabilitas nasional suatu negara, salah satu faktor dalam pembangunan ekonomi yang berperan sebagai penggerak utama adalah sektor industri yang lambat laun peranan dari sektor industri menjadi semakin penting dalam pembangunan Per-

ekonomian suatu negara [1].

Industri tekstil dan Produk tekstil (TPT) masih memberikan pengaruh yang signifikan dari industri tekstil dan produk tekstil (TPT) di Indonesia terhadap perekonomian nasional. Hal ini dapat terlihat pada Triwulan I – 2022 industri tekstil dan produk tekstil memberikan kontribusi sebesar 6,33% pada pembentukan Produk Domestik Bruto (PDB) sektor industri pengolahan non

migas, 5,67 terhadap total ekspor nasional pada 2021, pertumbuhan investasi pada sektor Industri Tekstil dan Produk Tekstil ini pada 2021 tercatat sebesar Rp6,5 Triliun dan pada Triwulan I – 2022 sebesar Rp2,4 Triliun [2].

Menurut Kementerian Perindustrian Republik Indonesia (2017) Industri Tekstil dan Produk Tekstil (TPT) serap 22,5% Tenaga kerja industri, disamping itu sektor industri juga akan merangsang kegiatan ekonomi sektor lain seperti sektor jasa, angkutan dan perdagangan. Pengembangan kawasan industri tekstil di Jawa Barat sendiri sudah direncanakan akan berada di Kabupaten Majalengka. Berdasarkan Rencana Induk Pembangunan Industri Nasional 2015-2035 [3], Majalengka juga menjadi salah satu Wilayah Pusat Pertumbuhan Industri (WPPI) Jawa Barat, hal ini sesuai dengan arahan pengembangan kegiatan ekonomi di Jawa Barat bagian timur yang meliputi Cirebon, Subang dan Majalengka dengan orientasi pergerakan ke arah Cirebon yang merupakan upaya untuk memperkuat dan memperluas sektor ekonomi di Jawa Barat bagian timur.

Kabupaten Majalengka merupakan salah satu Kabupaten yang terletak di Provinsi Jawa Barat dengan Luas wilayah Kabupaten Majalengka yaitu sebesar 1,204,24 km² yang terdiri dari 26 kecamatan yang terbagi menjadi 13 Kelurahan dan 323 Desa. Sektor Industri memiliki kontribusi tertinggi pada Produk Domestik Regional Bruto (PDRB) kabupaten Majalengka pada tahun 2021 dengan nilai sebesar 23,85% [4].

Industri Jeans merupakan salah satu subsektor industri Tekstil dan Produk Tekstil (TPT), jeans memang sangat populer diberbagai kalangan, baik pria maupun wanita, di Indonesia. Jeans termasuk kedalam jenis pakaian yang sering digunakan sehari-hari, hal tersebut karena jeans merupakan jenis pakaian yang sangat nyaman dan mudah digunakan un-

tuk aktivitas sehari-hari.

Pemilihan lokasi suatu organisasi (perusahaan) dapat mempengaruhi tingkat risiko dan tingkat keuntungan perusahaan secara keseluruhan, mengingat lokasi akan sangat berpengaruh pada biaya tetap ataupun biaya variabel, baik dalam jangka waktu menengah ataupun jangka panjang. Biaya dari transportasi dapat mencapai hingga dua puluh lima persen dari harga jual produk (tergantung pada produk dan tipe produksi atau jasa yang diberikan). Artinya dalam hal ini, seperempat total pendapatan perusahaan mungkin akan dibutuhkan untuk menutup biaya pengangkutan yang dilakukan untuk bahan mentah yang masuk dan produk jasa yang keluar dari perusahaan [5]. Adanya perbedaan tingkat sukses organisasi, perbedaan kekuatan dan juga kelemahan organisasi, sering diakibatkan karena faktor-faktor lokasi. Dalam situasi persaingan, faktor lokasi menjadi faktor kritis yang membuat faktor lokasi menjadi sangat penting. Agar usaha yang akan dijalankan dapat bersaing secara efektif, maka lokasi dari usaha yang dibuat haruslah strategis dan memiliki kemudahan untuk di jangkau [6].

Karena pada umumnya penentuan lokasi pabrik akan melibatkan sejumlah calon kandidat dan sejumlah kriteria. Maka, pada penelitian ini akan dilakukan penentuan lokasi pabrik dengan menggunakan metode *Agglomerative Hierarchical Clustering* (AHC), AHC digunakan karena data yang dimiliki dalam penentuan lokasi adalah klaster individu, yang kemudian dari hasil perhitungan kesamaan tiap klaster maka dua klaster yang memiliki kemiripan akan digabungkan menjadi klaster baru. Proses perhitungan ini akan terus berlanjut sampai diperoleh satu klaster tunggal.

Dalam metode *Agglomerative Hierarchical Clustering* (AHC) terdapat beberapa metode standarisasi data yang dapat digunakan yaitu *Z-Score* dan biner.

Standarisasi dengan *Z-Score* sering menggunakan metode klustering yaitu metode *Single Linkage (Nearest Neighbor)* yaitu jarak paling dekat, *Complete Linkage (Farthest Neighbor)* yaitu jarak paling jauh, *Centroid Linkage* yaitu jarak antar centroid pada dua klaster, *Average Linkage* yaitu dengan menghitung rata-rata jarak dari *pairwise distance*, *Ward's Minimum Variance (Ward's Linkage)* yaitu dengan menghitung nilai *Error sum of square (SSE)*. Yang selanjutnya akan dilakukan pengukuran kesamaan dan ketidaksamaan dengan kombinasi berbagai cara antara lain yaitu pengukuran jarak *Euclidean*, jarak *Square Euclidean*, jarak *Minkowski*, jarak *Chebyshev*, jarak *Mahalanobis*, jarak *City-Block* atau *Manhattan*, dan jarak *Canberra*. Sedangkan, standarisasi data dengan Biner biasanya dilakukan klusterisasi dengan berbagai *coefficient*. Saat ini, *coefficient* yang populer digunakan adalah *Jaccard Coefficient*, *Sokal – Michener Coefficient*, dan *Rogers – Tanimoto Coefficient* yang akan di kombinasikan dengan Metode *Single Linkage*.

Karena banyaknya metode pengelompokan dan pengukuran jarak dalam melakukan klusterisasi, maka dalam penelitian ini akan dicari metode standarisasi dan kombinasi metode mana yang terbaik dengan parameter nilai rata-rata simpangan baku dalam klaster (σ_w) dan simpangan baku antar klaster (σ_B) (Bunkers et al., 1996). Klaster yang baik merupakan klaster yang memiliki tingkat kesamaan atau homogenitas yang tinggi antar anggota dalam satu klaster yang sama dan memiliki ketidaksamaan atau heterogenitas yang tinggi antar klaster. Semakin kecil rasio simpangan baku dalam klaster dan simpangan baku antar klaster, maka semakin tinggi tingkat kesamaan atau homogenitasnya [7]. Selanjutnya, metode standarisasi dengan kombinasi terbaik dapat digunakan untuk pengolahan data klusterisasi yang dimiliki, sehingga dapat memberikan kemudahan bagi pengambil

keputusan dalam menentukan daerah mana yang tepat untuk didirikannya industri Tekstil dan Produk Tekstil (TPT) yaitu pabrik jeans di Majalengka. Proses pengklasteran akan dilakukan dengan cara *hybrid* yaitu menggunakan *software IBM SPSS Statistics 26, Andy's Agglomerative Hierarchical Clustering (AAHC)* dan secara manual dengan *Microsoft Office Excel*.

2. TUJUAN PEMBAHASAN

Pada makalah ini akan dibahas mengenai analisis kombinasi dari lima pengukuran jarak ketidaksamaan dengan tiga cara penggabungan pada data standarisasi *z-score* dan kombinasi tiga pengukuran jarak ketidaksamaan dengan satu cara untuk penggabungan pada data standarisasi Biner. Serta pengimplementasian dari metode klaster terbaik untuk memperoleh klaster terbaik dalam penentuan lokasi

3. METODOLOGI

Pada penelitian ini akan dilakukan dua phase penelitian yang akan dipaparkan pada poin-poin berikut:

3.1. Phase – 1

3.1.1. Menentukan Jumlah Kandidat /Objek dan Jumlah faktor atau kriteria

Langkah pertama yang dilakukan adalah menentukan banyaknya jumlah kriteria atau faktor pemilihan yang menjadi pertimbangan. Kandidat perlu ditentukan karena jumlah kandidat daerah tentunya akan berbeda tergantung pada tingkatan daerah yang diambil. Contohnya seperti jumlah desa dalam 1 kecamatan, jumlah kecamatan dalam 1 kabupaten, jumlah kabupaten atau kota dalam 1 provinsi, dan lain-lain.

Selain itu, faktor juga pasti akan berbeda-beda tergantung pada industri yang akan dibangun. Contohnya pemilihan lokasi pabrik akan mempertimbangkan kriteria atau faktor seperti lokasi pasar, sumber bahan baku, ketersediaan tenaga kerja,

listrik, air, akses transportasi, sikap masyarakat, serta peraturan pemerintah setempat. Tentunya kriteria atau faktor ini akan berbeda jika pemilihan lokasi akan dilakukan untuk memilih lokasi rumah sakit, tempat pembuangan sampah(TPS), lokasi peternakan, dan lain-lain.

3.1.2. Pengumpulan Data

Pada tahapan ini, dilakukan pengumpulan data dari kriteria dalam suatu kandidat. Data yang akan digunakan adalah data untuk menentukan kluster terbaik bagi pendirian lokasi industri manufaktur. Data yang digunakan merupakan data sekunder yang diperoleh dari *website* Badan Pusat Statistik.

Pada penelitian ini menggunakan data yang terdiri dari beberapa kriteria pemilihan yaitu:

1. Luas Wilayah, Data luas wilayah ini digunakan untuk dapat mengetahui luas masing-masing wilayah kecamatan yang ada pada kabupaten Majalengka dan juga digunakan untuk menghitung kepadatan penduduk dari setiap wilayah kecamatan. Luas wilayah ini memiliki satuan yaitu kilometer per segi (Km^2).
2. Kepadatan Penduduk, Luas wilayah yang besar bukan berarti memiliki ruang kosong yang besar juga. Oleh karena itu, data kepadatan penduduk juga digunakan pada penelitian ini, karena dalam salah satu aspek pemilihan lokasi adalah kemungkinan untuk melakukan perluasan. Kepadatan penduduk memiliki satuan yaitu orang per kilometer per segi (Orang/Km^2).
3. Angkatan Kerja, Angkatan kerja adalah penduduk usia kerja (15 tahun ke atas) yang bekerja, punya pekerjaan namun sementara tidak bekerja, dan pengangguran. Data ini dibutuhkan untuk dapat melihat daerah yang memiliki potensi untuk memudahkan dalam proses rekrutmen. Angkatan kerja memiliki satuan yaitu jumlah orang (orang).

4. Potensi Pasar, Potensi pasar dalam penelitian ini digunakan data jumlah UMK karena diasumsikan UMK dapat menjadi distributor hasil produksi atau konveksi. Potensi pasar memiliki satuan yaitu Unit UMK.

3.1.3. Standarisasi Data

3.1.3.1. Standarisasi dengan *Z-Score*

Dalam penelitian ini, standarisasi data dengan *z-score* akan menggunakan *software* IBM SPSS *Statistics* versi 26.

Adapun tahapan yang dilakukan untuk melakukan standarisasi data *Z-Score* dengan SPSS adalah sebagai berikut:

1. Masukan seluruh data yang dikumpulkan kedalam “IBM SPSS *Statistics Data Editor*”
2. Lakukan standarisasi data dengan memilih menu “*Analyze*” -> “*Descriptive Statistics*” -> “*Descriptives*”.
3. Kemudian masukan seluruh variabel instrumen penilai kedalam kotak *variables*, dan aktifkan bagian “*Save standardized values as variables*”. Lalu tekan OK.

3.1.3.2. Standarisasi dengan *Biner*

Standarisasi data dengan *Biner* akan menggunakan *software* *Andy's Agglomerative Hierarchical Clustering* (AAHC).

Tahapan untuk melakukan standarisasi data *Biner* dengan *software* AAHC adalah sebagai berikut:

1. Masuk kedalam *Software* AAHC
2. Masukkan data jumlah faktor atau kriteria yang menjadi pertimbangan serta masukkan data jumlah objek
3. Masukkan set data ke dalam program untuk memperoleh nilai masing-masing kelas yang akan digunakan untuk melakukan standarisasi data dengan melakukan klik hitung pada masing-masing kriteria – klik tabel *cluster* Pengukuran Jarak Ketidaksamaan

3.1.4. Pengukuran Jarak ketidaksamaan

Pengukuran jarak ketidaksamaan dengan z -score dilakukan dengan *software* SPSS, untuk melakukan pengukuran jarak dengan SPSS adalah sebagai berikut:

1. Masukan seluruh variabel yang telah dilakukan standarisasi data (Z -score) ke dalam bagian “*Variable(s)*”. kemudian isi pada bagian “*Label Cases by*” dengan variabel, sedangkan untuk bagian “*Cluster*” pilih “*Cases*”, pada bagian “*Display*” pilih keduanya yaitu *Statistics* dan *Plots*.
2. Pada menu kotak “*Statistics...*” aktifkan kotak “*Agglomeration schedule*” dan “*Proximity matrix*” untuk menampilkan jarak antar variabel. Kemudian pada bagian “*Cluster Membership*” klik pada pilihan “*Range of solutions*” dan tentukan nilai *Minimum number of clusters* dan *Maximum number of clusters*.
3. Pada menu kotak “*Plots...*” aktifkan pilihan “*Dendrogram*”. Kemudian pada bagian “*Icicle*” klik pada pilihan “*None*”.
4. Pada menu kotak “*Method...*” pada bagian “*Cluster Method*” pilih dengan metode pengklasteran yang akan dilakukan. Kemudian pada bagian “*Interval*” di kotak menu “*Measure*” pilih sesuai dengan cara pengukuran jarak yang akan dilakukan. Dan pada bagian “*Transform Values*” pilih “*Z-Score*”. Abaikan bagian yang lain lalu tekan tombol “*Continue*”.

Pengukuran jarak ketidaksamaan dengan biner dilakukan dengan *software* AAHC, untuk melakukan pengukuran jarak dengan AAHC adalah sebagai berikut:

1. Masukan seluruh variabel yang telah dilakukan standarisasi data Biner
2. Pilih persamaan. Persamaan yang akan digunakan dapat dipilih pada menu “Pengukuran Jarak Antar Objek” – pilih menu “D” untuk memilih *dissimilarity*

3.1.5. Proses klastering

Proses klastering dengan SPSS sudah di set pada tahapan sebelumnya, dan akan langsung memberikan hasil berupa urutan penggabungan, *Cluster membership*, dan dendogram .

Untuk proses klastering dengan AAHC, setiap objek yang bergabung (klaster yang terbentuk) dilakukan pencatatan, setelah semua objek bergabung, catatan yang dimiliki selanjutnya diolah menjadi sebuah dendogram,

3.1.6. Menentukan Kombinasi Terbaik dari Pengukuran Jarak Ketidaksamaan atau Dissimilarity dan Metode Penggabungan

Setelah klaster terbentuk pada tahapan sebelumnya, proses selanjutnya adalah melakukan perhitungan simpangan baku dalam klaster (σ_w) dan simpangan baku antar klaster (σ_B) untuk memperoleh nilai rasio simpangan baku (σ) [8]. Pada tahapan ini dilakukan beberapa langkah yaitu sebagai berikut:

1. Menghitung simpangan baku dalam klaster, langkah awal untuk melakukan perhitungan simpangan baku dalam klaster yaitu menentukan dahulu jumlah klaster yang akan dibentuk dari hasil klastering. Berikutnya, menghitung nilai rata-rata klaster ke- k , menghitung nilai simpangan baku klaster ke- k , lalu dapat dilakukan perhitungan simpangan baku dalam klaster. Untuk formulasi yang digunakan dalam menghitung simpangan baku dalam klaster adalah sebagai berikut:

- a. Formulasi Simpangan baku dalam *cluster* (σ_w)

$$\sigma_w = \frac{1}{K} \sum_{k=1}^K \sigma_k$$

Keterangan:

K adalah banyaknya *cluster* yang terbentuk

σ_k adalah simpangan baku *cluster* ke- k

- b. Formulasi Simpangan *cluster* ke- k (σ_k)

$$\sigma_k = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \mu_k)^2}$$

Keterangan:

N adalah jumlah anggota dari tiap *cluster*

μ_k adalah rata-rata *cluster* ke-k

X_i adalah anggota *cluster*, dari $I = 1, 2, \dots, N$

2. Menghitung simpangan baku antar kluster, dilakukan dengan terlebih dahulu menghitung rata-rata dari keseluruhan kluster yang terbentuk serta menghitung rata-rata dari kluster ke- k. Selanjutnya, perhitungan simpangan baku antar kluster dapat dilakukan. Untuk formulasi yang digunakan dalam menghitung simpangan baku antar kluster adalah sebagai berikut:

$$\sigma_B = \sqrt{\frac{1}{K-1} \sum_{k=1}^K (\mu_k - \mu)^2}$$

Keterangan:

μ_k adalah rata-rata *cluster* ke-k

μ adalah rata-rata keseluruhan *cluster*

3. Menghitung rasio simpangan baku, untuk memperoleh nilai rasio simpangan baku dapat dilakukan dengan cara membagi simpangan baku dalam kluster dengan simpangan baku antar kluster. Untuk formulasi yang digunakan dalam menghitung rasio simpangan baku adalah sebagai berikut:

$$\sigma = \frac{\sigma_W}{\sigma_B} \times 100\%$$

Keterangan:

σ_W adalah simpangan baku dalam *cluster*

σ_B adalah simpangan baku antar *cluster*

3.2. Phase – 2

3.3. Mendapatkan Hasil Klustering Terbaik Berdasarkan Parameter Rasio Simpangan Baku

Kluster yang baik adalah kluster yang mempunyai heterogenitas (perbedaan)

yang tinggi antar kluster yang satu dengan yang lain. Setelah semua kombinasi dihitung simpangan baku dalam kluster (σ_W) dan simpangan baku antar klusternya (σ_B) dan semua sudah diketahui nilai rasio simpangan bakunya (σ), metode klustering dengan kombinasi pengukuran jarak dan metode penggabungan terbaik dapat dilihat dari metode klustering dengan nilai rasio simpangan baku terkecil.

3.4. Implementasi Klustering dengan Kombinasi terbaik

Setelah diperoleh metode klustering dengan kombinasi terbaik, metode kluster dengan kombinasi terbaik tersebut akan diimplementasikan dalam klusterisasi data untuk memperoleh kluster yang paling baik atau memiliki kesamaan yang tinggi dan perbedaan yang tinggi antar kluster.

4. Hasil

4.1. Skenario Kasus

Skenario merupakan rancangan dari peristiwa yang dibuat sesuai dengan keinginan dari perancangannya. Agar proses dapat bekerja dan hasil dari kombinasi metode pengukuran jarak ketidaksamaan dan metode penggabungan atau pengelompokkan yang diusulkan sesuai dengan yang diharapkan, maka perlu dibuat suatu skenario kasus yang bertujuan untuk mencapai tujuan tersebut. Pada penelitian ini, data yang digunakan ialah data kabupaten Majalengka yang diperoleh dari *website* Badan Pusat Statistik Kabupaten Majalengka.

4.2. Data Numerik

Pada penelitian ini data yang digunakan adalah data kabupaten Majalengka. Dari data ini, ditentukan yang akan menjadi kandidat pemilihan adalah 26 kecamatan yang berada di kabupaten Majalengka, dan kriteria yang digunakan merupakan data kriteria atau faktor dalam memilih lokasi industri tekstil antara lain yaitu 4 faktor pemilihan dengan satuan yang berbeda

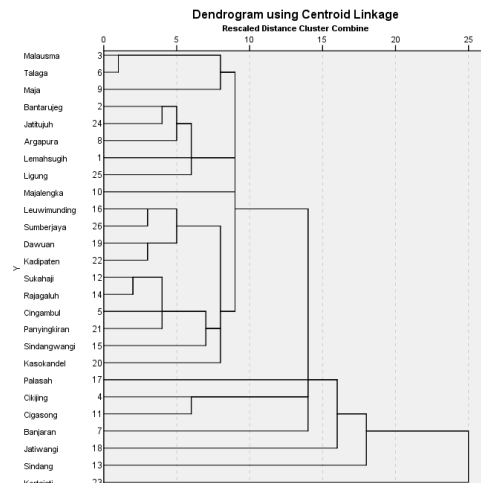
lima cara pengukuran jarak (*Euclidean Distance*, *Squared Euclidean Distance*, *Chebyshev Distance*, *City Block Distance*, dan *Minkowski distance*). Sedangkan data standarisasi biner dilakukan dengan tiga cara tiga cara pengukuran jarak yaitu *Sokal – Michener Coefficient*, *Russel Rao Coefficient*, dan *Jaccard Coefficient*.

4.3.3. Proses Klustering

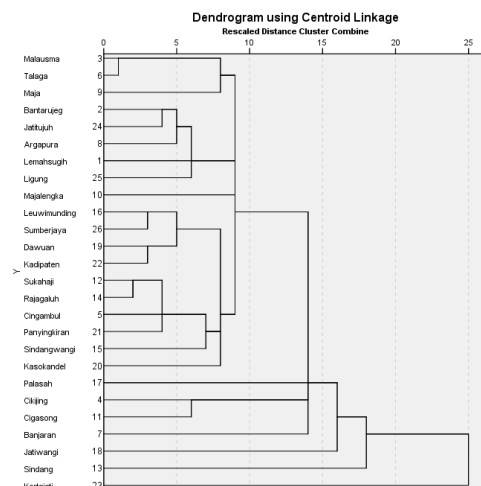
Proses agglomerasi (pengumpulan objek) dimulai dari dua objek yang memiliki jarak terdekat dari sejumlah kombinasi jarak di antara 26 objek yang ada menggunakan aturan dan perhitungan dari masing-masing metode penggabungan atau klustering. Proses klustering untuk data standarisasi *z-score* dilakukan dengan empat cara yaitu *Single Linkage*, *Centroid Linkage*, *Average Linkage*, dan *Ward's Linkage*. Sedangkan data standarisasi biner dilakukan dengan satu cara proses penggabungan yaitu *single linkage*.

4.3.4. Hasil Klustering

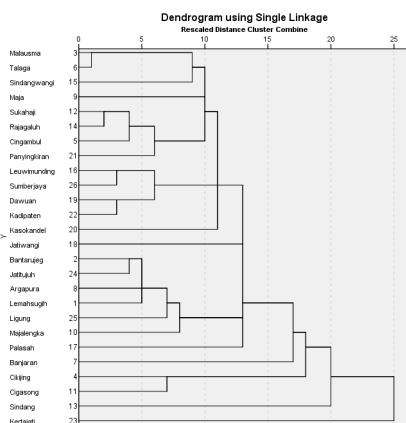
Hasil dari masing-masing kombinasi metode akan menghasilkan sebuah dendogram. Dendogram untuk masing-masing kombinasi dari pengukuran jarak ketidaksamaan dengan metode penggabungan akan di tampilkan pada gambar IV.1. sampai IV.23.



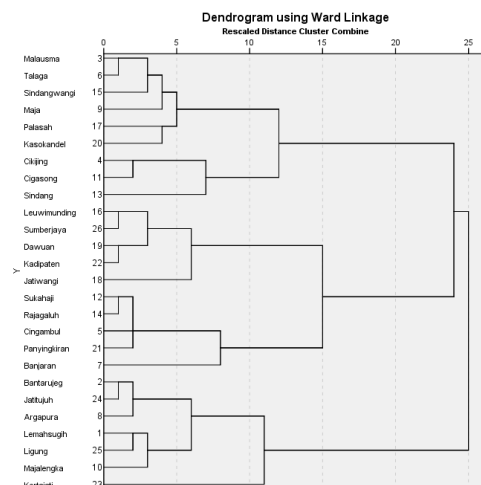
Gambar 4.2 *Euclidean Distance* dan *Centroid Linkage*



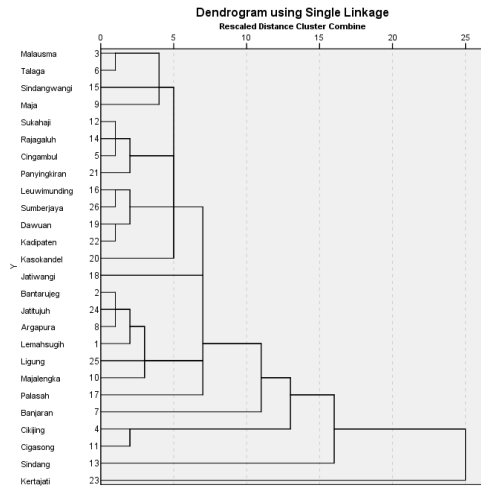
Gambar 4.3 *Euclidean Distance* dan *Average Linkage*



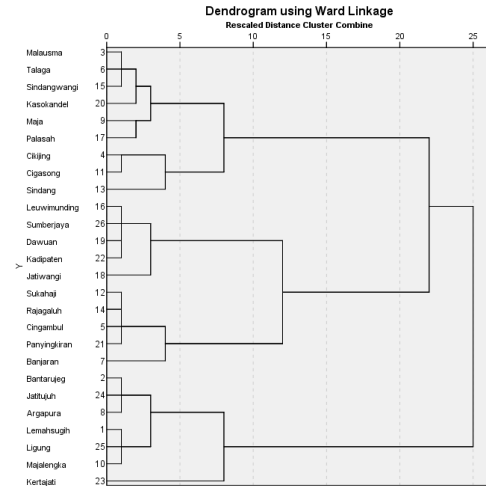
Gambar 4.1 *Euclidean Distance* dan *Single Linkage*



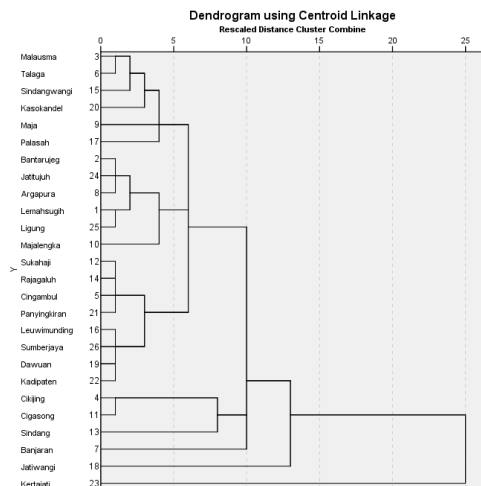
Gambar 4.4 *Euclidean Distance* dan *Ward's Method*



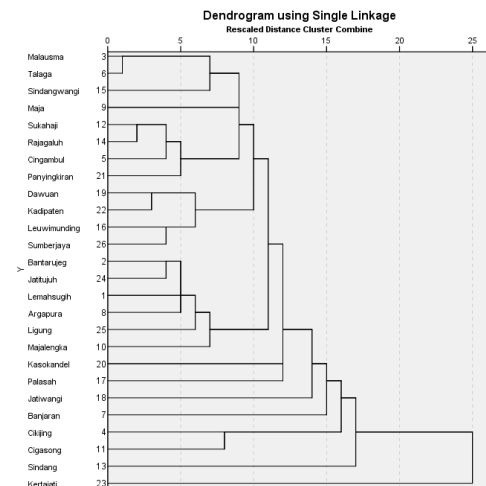
Gambar 4.5 Squared Euclidean Distance dan Single Linkage



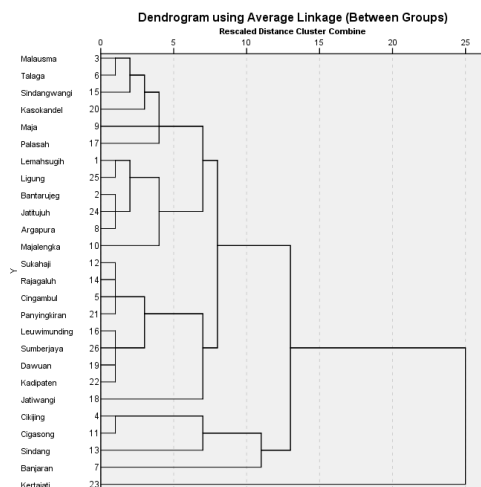
Gambar 4.8 Squared Euclidean Distance dan Ward's Method



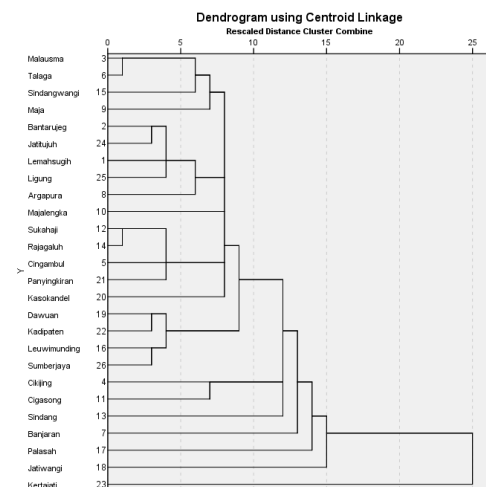
Gambar 4.6 Squared Euclidean Distance dan Centroid Linkage



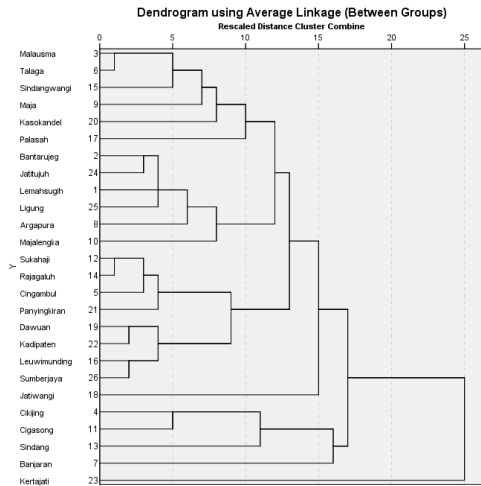
Gambar 4.9 Chebychev Distance dan Single Linkage



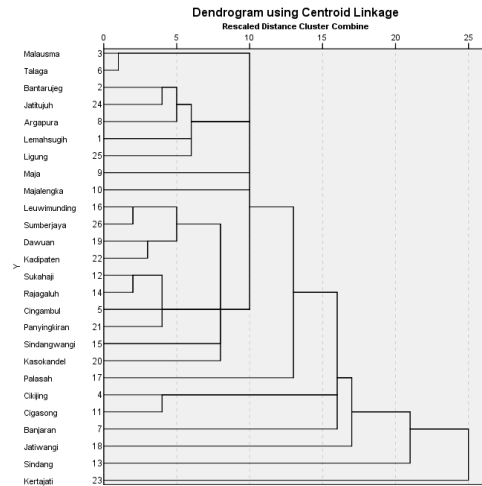
Gambar 4.7 Squared Euclidean Distance dan Average Linkage



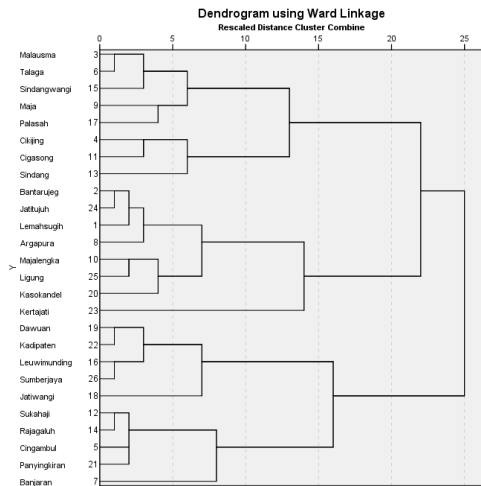
Gambar 4.10 Chebychev Distance dan Centroid Linkage



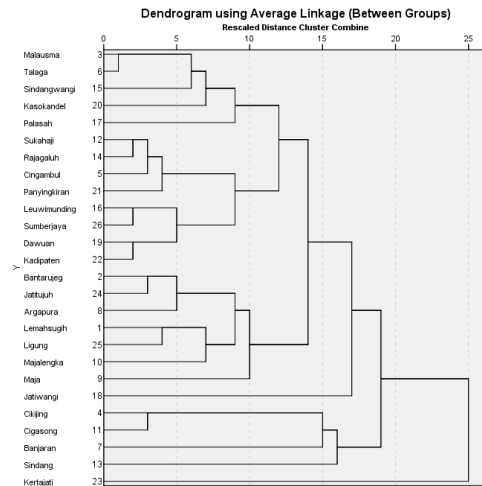
Gambar 4.11 *Chebychev Distance* dan *Average Linkage*



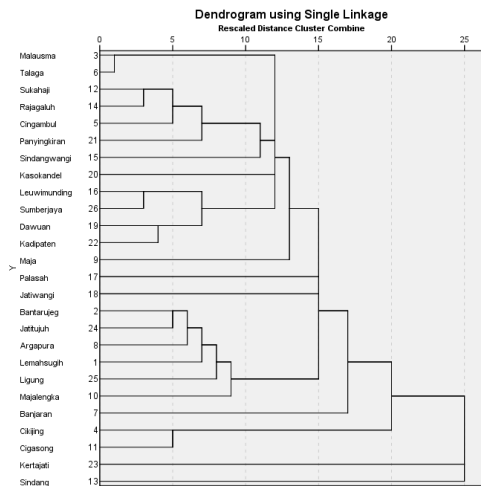
Gambar 4.14 *City Blok Distance* dan *Centroid Linkage*



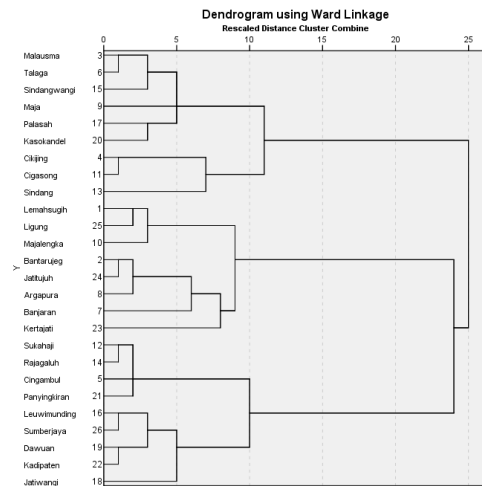
Gambar 4.12 *Chebychev Distance* dan *Ward's Method*



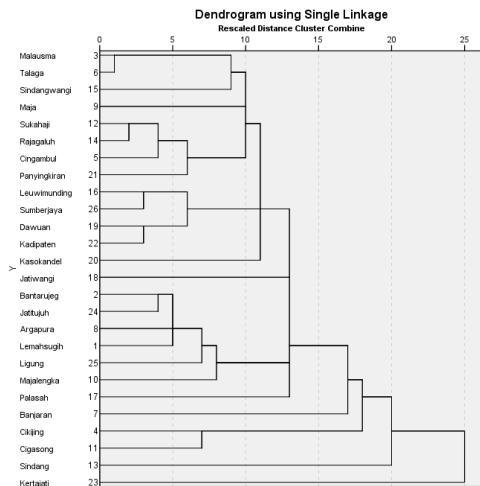
Gambar 4.15 *City Blok Distance* dan *Average Linkage*



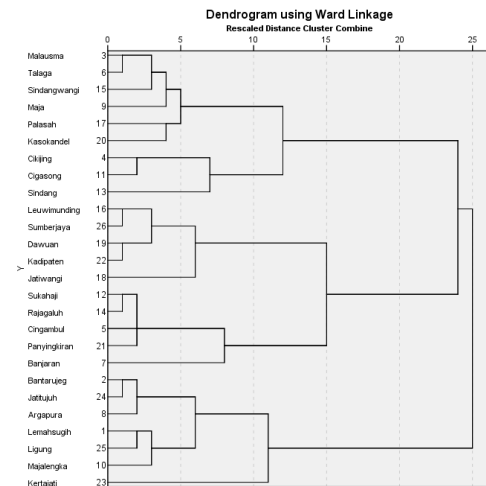
Gambar 4.13 *City Blok Distance* dan *Single Linkage*



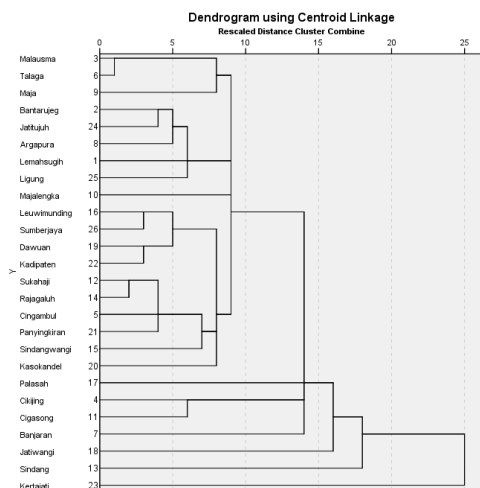
Gambar 4.16 *City Blok Distance* dan *Ward's Method*



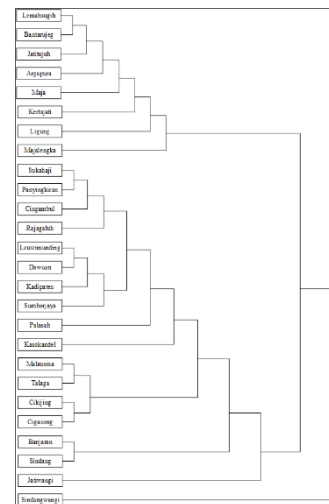
Gambar 4.17 *City Blok Distance* dan *Single Linkage*



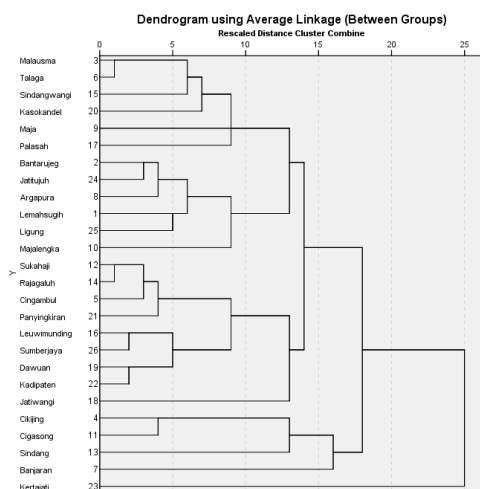
Gambar 4.20 *City Blok Distance* dan *Ward's Method*



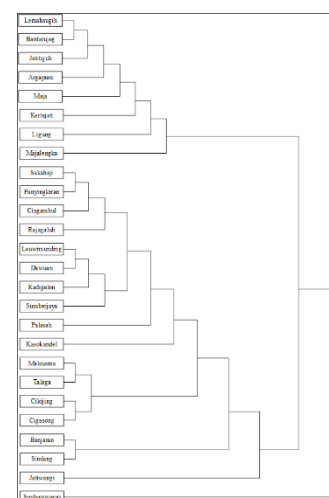
Gambar 4.18 *Minkowski Distance* dan *Centroid Linkage*



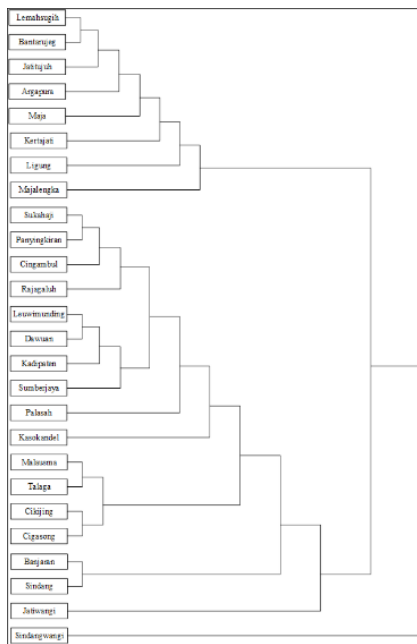
Gambar 4.21 *Sokal – Michener Coefficient* dan *Single Linkage*



Gambar 4.19 *Minkowski Distance* dan *Average Linkage*



Gambar 4.22 *Russel Rao Coefficient* dan *Single Linkage*



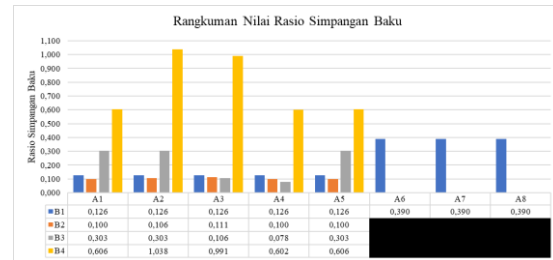
Gambar 4.23 Russel Rao Coefficient dan Single Linkage

4.4. Phase – 2

4.4.1. Perhitungan Rasio Simpangan Baku

Menurut Bunkers dkk (1996) untuk dapat mengetahui kombinasi dari cara pengukuran jarak dan metode penggabungan atau pengelompokan dengan kinerja yang terbaik dapat diketahui dengan menggunakan dua kriteria yaitu nilai dari rata-rata simpangan baku dalam kluster (σ_w) dan nilai dari simpangan baku antar kluster (σ_B). Kombinasi metode klustering terbaik dapat diketahui dari kombinasi yang memiliki nilai rasio dari nilai simpangan baku dalam kluster (σ_w) dan nilai dari simpangan baku antar kluster (σ_B) yang terkecil. Jadi, Jika nilai σ_w semakin kecil dan nilai σ_B semakin besar maka dapat diketahui bahwa metode klasterisasi tersebut adalah metode klasterisasi yang memiliki kinerja yang paling baik atau artinya memiliki tingkat kesamaan atau homogenitas yang tinggi antara anggota didalam satu kluster atau ketidaksamaan atau heterogenitas yang tinggi antara kluster satu dengan kluster yang lain.

Rangkuman hasil perhitungan rasio simpangan baku dari masing-masing kombinasi akan ditampilkan pada gambar 4.24.



Gambar 4.24 Grafik Rangkuman Nilai Rasio Simpangan Baku

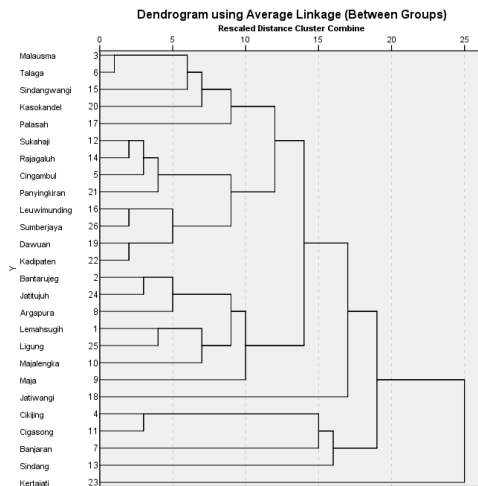
Keterangan

- A1 = Euclidean Distance
- A2 = Squared Euclidean Distance
- A3 = Chebychev Distance
- A4 = City Block Distance
- A5 = Minkowski Distance
- A6 = Koefisien Sokal – Michener
- A7 = Koefisien Russel Rao
- A8 = Koefisien Jaccard
- B1 = Single Linkage
- B2 = Centroid Linkage
- B3 = Average Linkage
- B4 = Ward's Linkage

Dari Grafik rangkuman hasil perhitungan rasio simpangan baku yang telah dibuat, dapat dilihat dan diketahui kombinasi metode terbaik yaitu metode dengan kombinasi dari cara pengukuran jarak ketidaksamaan *City Block Distance* dengan metode penggabungan *Average Linkage* yang memperoleh hasil nilai rasio simpangan baku terkecil yaitu 0,078.

4.4.2. Implementasi Klustering dengan Kombinasi terbaik

Setelah memperoleh metode terbaik yang sebelumnya dievaluasi melalui perolehan nilai simpangan baku terkecil, berikutnya kombinasi metode terbaik yaitu kombinasi dari cara pengukuran jarak ketidaksamaan *City Block Distance* dengan metode penggabungan *Average Linkage*. Tujuan di implementasikannya kombinasi terbaik ini yaitu untuk dapat memperoleh kluster terbaik dengan tingkat kesamaan yang tinggi antar anggota kluster dan memiliki tingkat perbedaan yang tinggi dengan kluster lain. Pada tahapan ini dibantu dengan *Software SPSS*, hasil dari tahapan ini akan berbentuk dendogram seperti yang di tampilkan pada gambar 4.25 berikut.



Gambar 4.25 Dendrogram Kluster kombinasi *City Blok Distance* dan *Average Linkage*

- Dari dendrogram tersebut dapat diketahui, jika digunakan kombinasi metode terbaik dan dipilih jumlah kluster yang akan terbentuk sebanyak 5 kluster maka masing-masing anggota kluster yang akan terbentuk dapat dilihat pada tabel IV.59.

Tabel 4.4 Anggota Kluster Kombinasi *City Blok Distance* dan *Average Linkage*

Kluster	Jumlah Anggota	Anggota Kluster
1	20	Lemahsugih, Bantarueg, Malausma, Cingambul, Talaga, Argapura, Maja, Majalengka, Sukahaji, Rajagaluh, Sindangwangi, Leuwimunding, Palasah, Dawuan, Kasokandel, Panyingkiran, Kadipaten, Jatituh, Ligung, Sumberjaya
2	3	Cikijing, Banjaran, Cigasong
3	1	Sindang
4	1	Jatiwangi
5	1	Kertajati

5. ANALISIS DAN KESIMPULAN

5.1. Analisis

Pada penelitian ini kombinasi pengukuran jarak dan proses pengklasteran akan dilakukan terhadap 26 kandidat berupa kecamatan yang mempertimbangkan 4 kriteria untuk dapat memperoleh kluster terbaik yang dapat digunakan untuk pemilihan lokasi.

Fase pertama pada penelitian ini dilakukan dengan langkah-langkah yaitu melakukan standarisasi terhadap data yang dimiliki, standarisasi data pada penelitian ini

digunakan 2 metode yaitu dengan *Z-Score* dan Biner. Standarisasi data ini berguna untuk menghilangkan keragaman atau perbedaan satuan dari faktor atau kriteria yang dipertimbangkan. Selanjutnya, dilakukan pengukuran jarak ketidaksamaan yaitu data standarisasi *Z-Score* dengan *Euclidean Distance*, *Squared Euclidean Distance*, *Chebyshev Distance*, *City Block Distance*, dan *Minkowski Distance*. Sedangkan, data standarisasi Biner dengan Koefisien *Sokal – Michener*, Koefisien *Russel Rao*, dan Koefisien *Jaccard*. Setelah itu, adalah proses pengklasteran dimana data standarisasi *Z-Score* dilakukan dengan *Single Linkage*, *Centroid Linkage*, *Average Linkage*, dan *Ward's Linkage*. Untuk data standarisasi biner dilakukan dengan *Single Linkage*. Setelah kluster terbentuk, selanjutnya adalah melakukan perhitungan nilai simpangan baku dalam kluster dan simpangan baku antar kluster yang nantinya kedua nilai tersebut akan digunakan untuk memperoleh nilai rasio simpangan baku.

Pada Fase kedua akan dilakukan pemilihan kombinasi metode terbaik, terlebih dahulu hasil dari proses perhitungan nilai rasio simpangan baku dikumpulkan dalam 1 tabel untuk dapat memudahkan dalam memberikan informasi kombinasi metode terbaik.

Pada fase kedua ini berdasarkan gambar 4.24 grafik rangkuman dapat dilihat bahwa kombinasi dari cara pengukuran jarak ketidaksamaan *City Block Distance* dengan metode penggabungan *Average Linkage* merupakan metode yang memiliki kinerja yang paling baik diantara kombinasi metode pengukuran jarak ketidaksamaan dengan metode penggabungan lainnya. hal ini dapat diketahui dari nilai rasio simpangan baku, kombinasi dari cara pengukuran jarak ketidaksamaan *City Block Distance* dengan metode penggabungan *Average Linkage* menghasilkan nilai rasio simpangan baku

terkecil. Pada *phase* kedua ini, selanjutnya metode terbaik diimplementasikan untuk dapat memberikan kluster terbaik dengan nilai rasio simpangan baku terkecil.

Dari penelitian ini dapat diketahui bahwa standarisasi *Z – Score* baik digunakan jika data yang dimiliki berdistribusi normal atau mendekati distribusi normal. Adapun kelebihan standarisasi *Z – Score* yaitu dapat mengidentifikasi *outlier* dalam dataset yang dapat dilihat dari nilai-nilai yang berada jauh dari rata-rata populasi, meningkatkan interpretasi data dengan melihat sejauh mana setiap titik data berbeda dari mean populasi (Nilai positif menunjukkan bahwa titik data berada di atas rata-rata, sedangkan nilai negatif menunjukkan bahwa titik data berada di bawah rata-rata), menghilangkan efek *outlier* (Standarisasi *Z – Score* mencakup pengurangan mean dan pembagian dengan standar deviasi, sehingga data yang berada jauh dari *mean* akan mendapatkan skor *Z – Score* yang tinggi atau rendah). Sedangkan kekurangan dari standarisasi *Z – Score* yaitu membutuhkan data yang berdistribusi normal atau mendekati distribusi normal, standarisasi *Z – Score* tidak tahan terhadap perbedaan pada data yang ekstrem (Jika terdapat perbedaan data ekstrem yang signifikan dalam dataset, standarisasi *Z – Score* mungkin tidak memberikan hasil yang baik. Data ekstrem dapat mempengaruhi nilai mean dan standar deviasi, sehingga mempengaruhi skor *Z – Score* dari data lainnya). Standarisasi biner baik digunakan untuk data kategorikal atau data yang dapat diukur dengan metrik (*Sokal – Michener Coefficient*, *Russel Rao Coefficient*, *Jaccard Coefficient*, *indeks Dice*, atau *Hamming distance*). Kelebihan dari standarisasi biner yaitu Sederhana dan mudah diimplementasikan (Binary-based dissimilarity mengubah data menjadi bentuk biner, di mana nilai 1 menunjukkan kehadiran suatu fitur atau atribut, sedangkan nilai 0 menunjukkan ketidakhadiran). Kekurangan metode ini

yaitu tidak memberikan informasi tentang intensitas perbedaan (Metode ini hanya memberikan informasi tentang adanya perbedaan atau kesamaan antara data poin tanpa memberikan informasi tentang seberapa besar atau kuat perbedaan tersebut) kehilangan informasi (informasi mengenai intensitas atau tingkat perbedaan antara fitur tidak diperhitungkan. Semua perbedaan dianggap sama, baik itu perbedaan kecil maupun perbedaan besar. Ini dapat mengakibatkan kehilangan informasi penting dalam analisis data).

Penelitian ini dilakukan hanya sampai diperoleh hasil kluster dengan sejumlah kluster yang ditetapkan dan tidak sampai kepada pemilihan lokasi terpilih. Hal ini dikarenakan untuk pemilihan lokasi terpilih memerlukan data dan informasi-informasi mengenai kebutuhan-kebutuhan lain serta peraturan pemerintah daerah setempat yang dapat menjadi pertimbangan dan dapat mendukung keputusan pemilihan calon lokasi atau kandidat. Jika tersedia data dan informasi kebutuhan lain yang dapat mendukung pengambilan keputusan akhir, proses yang selanjutnya dilakukan yaitu memilih kluster, pemilihan kluster ini dilakukan dengan evaluasi kluster berdasarkan kriteria tertentu. Jika kluster terpilih sudah didapatkan, selanjutnya yaitu melakukan evaluasi menggunakan kriteria atau faktor lain untuk dapat memilih satu anggota dari kluster terpilih.

Kluster yang dihasilkan dari kombinasi metode terbaik ini tidak memberikan informasi mengenai kluster mana yang lebih baik dari kluster yang lain. Tetapi, kluster ini dapat memberikan informasi berkaitan dengan lokasi yang memiliki kesamaan berdasarkan faktor-faktor atau kriteria-kriteria yang menjadi pertimbangan. Dari hasil pengimplementasian kombinasi metode terbaik didapatkan hasil berupa kluster tunggal.

Apabila klaster yang menjadi pilihan hanya memiliki satu anggota, maka anggota tersebut dapat langsung ditetapkan sebagai lokasi terpilih. Jika klaster yang dipilih terbentuk lebih dari satu anggota, maka untuk pemilihan lokasi sebagai keputusan akhir diperlukan pengkajian lebih lanjut untuk dapat mengetahui apa saja yang menjadikan kandidat atau calon lokasi pemilihan berada didalam klaster yang sama serta perlu dipelajari kelebihan-kelebihan dan juga kekurangan-kekurangan dari setiap kandidat pilihan. Selain itu, perlu diperhatikan juga informasi kebutuhan lain yang mendukung pendirian industri *jeans* ini.

5.2. Kesimpulan

Penelitian ini membahas penentuan klaster terbaik yang dapat digunakan untuk pemilihan lokasi menggunakan pendekatan metode kuantitatif. Berdasarkan hasil perhitungan, analisis serta pembahasan yang telah dilakukan sebelumnya berkaitan dengan tujuan penelitian ini, dapat disimpulkan bahwa:

1. Dari sejumlah kombinasi cara pengukuran jarak ketidaksamaan yaitu *Euclidean Distance*, *Squared Euclidean Distance*, *Chebyshev Distance*, *City Block Distance*, *Minkowski Distance*, Koefisien *Sokal – Michener*, Koefisien *Russel Rao*, dan Koefisien *Jaccard* dengan metode penggabungan atau klasterisasi yaitu *Single Linkage*, *Centroid Linkage*, *Average Linkage*, dan *Ward's Linkage*, diperoleh kombinasi terbaik (memberikan nilai simpangan baku terkecil yaitu 0,078) pada saat digunakan metode *City Block Distance* dan *Average Linkage*.
2. Hasil implementasi dari kombinasi metode terbaik adalah sebuah klaster besar, yang apabila klaster besar ini ditentukan terdiri dari 5 klaster akan beranggotakan masing-masing klasternya yaitu sebagai berikut:

- a. Klaster 1 : Lemahsugih, Bantarujeg, Malausma, Cingambul, Talaga, Argapura, Maja, Majalengka, Sukahaji, Rajagaluh, Sindangwangi, Leuwimunding, Palasah, Dawuan, Kasokandel, Panyingkiran, Kadipaten, Jatitujuh, Ligung, Sumberjaya
 - b. Klaster 2 : Cikijing, Banjaran, Cigasong
 - c. Klaster 3 : Sindang
 - d. Klaster 4 : Jatiwangi
 - e. Klaster 5 : Kertajati
3. Untuk memilih klaster terbaik dari kelima klaster diatas, maka diperlukan kriteria tertentu yang harus ditentukan oleh pengambil keputusan didaerah tersebut.

Daftar Pustaka

- [1] Todaro Michael Paul and Smith Stephen Charles, *Economic Development*, 11th ed. Pearson Addison Wesley, 2011.
- [2] Badan Pusat Statistik Indonesia, "Statistik Indonesia 2022," 2022.
- [3] Kementerian Perindustrian Republik Indonesia, "RENCANA INDUK PEMBANGUNAN INDUSTRI NASIONAL REPUBLIK INDONESIA," 2015. Accessed: Feb. 02, 2023. [Online]. Available: www.kemenperin.go.id
- [4] Badan Pusat Statistik Kabupaten Majalengka, "Kabupaten Majalengka Dalam Angka 2022," Majalengka, Feb. 2022.
- [5] J. Heizer and B. Render, *Operations Management, 7th Edition*, 7th ed. New Jersey: Pearson Education. Inc., Upper Saddle River, 2004.
- [6] T. H. Handoko, *Manajemen Personalia dan Sumber Daya Manusia*, Edisi ke 2. Yogyakarta: BPFE, 2000.
- [7] F. R. Maulana, "Analisis Hasil Klastering Menggunakan Lima Metode Pengelompokan dengan Tiga

Cara Pengukuran Jarak pada Hierarchical Clustering,” Universitas Pasundan, Bandung, 2022.

- [8] M. J. Bunkers, J. R. Miller, and A. T. DeGaetano, “Definition of climate regions in the northern plains using an objective cluster modification technique,” *J. Clim.*, vol. 9, no. 1, pp. 130–146, 1996, doi: 10.1175/1520-0442(1996)009<0130:DOCRIT>2.0.CO;2.