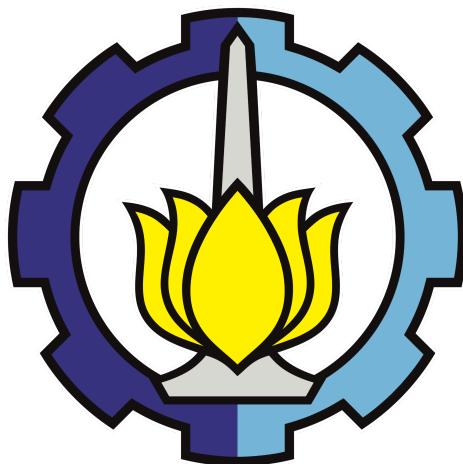


**ANALISIS ZONASI DAN PERSEBARAN SMA NEGERI DI KOTA
DAN KABUPATEN MALANG DENGAN METODE
AGGLOMERATIVE CLUSTERING**

ANALITIKA DATA DAN DIAGNOSTIK



DOSEN PENGAMPU :

Retno Aulia Vinarti, S.Kom., M.Kom., Ph.D.

KELOMPOK FINAL PROJECT - ADD Kelas D:

- | | |
|----------------------------------|--------------|
| 1. Farhan Bilawa Gita Maulana | (5026211158) |
| 2. Alif Faturrohman | (5026221040) |
| 3. Dewi Maharani | (5026221046) |
| 4. Muhammad Iqbal Baiduri Yamani | (5026221103) |
| 5. Keysha Amelia | (5026221122) |

**DEPARTEMEN SISTEM INFORMASI
FAKULTAS TEKNOLOGI ELEKTRO DAN INFORMATIKA CERDAS
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA**

2023

DAFTAR ISI

| | |
|--|-----------|
| DAFTAR ISI..... | 2 |
| KATA PENGANTAR..... | 3 |
| BAB I PENDAHULUAN..... | 4 |
| 1.1. Latar Belakang..... | 4 |
| 1.2. Rumusan Masalah..... | 5 |
| 1.3. Tujuan dan Manfaat..... | 5 |
| BAB II METODOLOGI PENELITIAN..... | 7 |
| 2.1. Data..... | 7 |
| 2.2. Agglomerative Hierarchical Clustering..... | 16 |
| 2.2.1. Clustering..... | 16 |
| 2.2.2. Agglomerative Hierarchical Clustering..... | 16 |
| 2.3. Sistem Zonasi di Kota dan Kabupaten Malang..... | 16 |
| 2.4 Outliers..... | 18 |
| 2.5 Tools untuk Mengolah Data..... | 19 |
| 2.6 Langkah-langkah Agglomerative Hierarchical Clustering..... | 20 |
| BAB III PEMBAHASAN..... | 25 |
| 3.1 Pengolahan Data..... | 25 |
| 3.1.1 Data SMA Negeri di Kota dan Kabupaten Malang..... | 25 |
| 3.1.2 Data Kecamatan dengan Kepadatan Penduduk Tertinggi di Kota dan Kabupaten Malang..... | 26 |
| 3.1.3 Data Kantor Kecamatan di Kota dan Kabupaten Malang..... | 27 |
| 3.2. Analisis Hasil Agglomerative Hierarchical Clustering..... | 27 |
| 3.2.1 AHC Metode Single Linkage:..... | 28 |
| 3.2.2 AHC Metode Average Linkage:..... | 28 |
| 3.2.3 AHC Metode Complete Linkage:..... | 29 |
| BAB IV PENUTUP..... | 30 |
| 4.1. Simpulan..... | 30 |
| 4.2. Saran..... | 30 |
| LAMPIRAN..... | 31 |
| DAFTAR PUSTAKA..... | 32 |

KATA PENGANTAR

Puji syukur kami panjatkan kepada Tuhan Yang Maha Esa atas berkat dan rahmat-Nya, sehingga kami dapat menyelesaikan laporan ini. Laporan ini berjudul "Analisis Zonasi dan Persebaran SMA Negeri di Kota dan Kabupaten Malang Tahun 2023 Dengan Metode Agglomerative Clustering". Laporan penelitian ini disusun berdasarkan hasil pengolahan data yang telah dilakukan oleh penulis. Penelitian ini menggunakan data seluruh SMA Negeri yang berada pada Kota dan Kabupaten Malang. Penulis menyadari bahwa laporan analisis data outlier ini masih memiliki kekurangan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun dari pembaca. Demikian laporan analisis ini, semoga bermanfaat.

Surabaya, 21 Desember 2023

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pendidikan merupakan salah satu aspek penting dalam pembangunan suatu daerah. Salah satu indikator kemajuan pendidikan di suatu wilayah dapat dilihat dari distribusi dan persebaran Sekolah Menengah Atas (SMA) negeri. Oleh karena itu, penelitian ini bertujuan untuk menganalisis zonasi dan persebaran SMA Negeri di Kota dan Kabupaten Malang menggunakan metode Agglomerative Clustering.

Makalah ini disusun sebagai tugas perkuliahan guna memenuhi persyaratan akademis. Tujuan utama penelitian ini adalah untuk menyelidiki dan menggambarkan distribusi serta persebaran SMA Negeri di Kota dan Kabupaten Malang melalui analisis metode Agglomerative Clustering. Penelitian ini menitikberatkan pada penggunaan data koordinat lokasi seluruh SMA Negeri di wilayah Kota dan Kabupaten Malang sebagai titik awal analisis. Selain itu, koordinat lokasi kantor kecamatan di wilayah tersebut juga diambil sebagai parameter untuk menentukan clustering data. Metode yang digunakan dalam penelitian ini adalah Agglomerative Clustering, sebuah metode clustering hirarkis yang memungkinkan pengelompokan berdasarkan kesamaan jarak antar elemen.

Dengan mendasarkan temuan pada data koordinat dan metode Agglomerative Clustering, penelitian ini berpotensi memberikan kontribusi signifikan dalam pemahaman tentang zonasi dan distribusi SMA Negeri di wilayah tersebut. Hasil penelitian ini diharapkan dapat menjadi dasar bagi pembuat kebijakan pendidikan untuk merancang kebijakan yang lebih efektif dan efisien dalam memajukan sistem pendidikan di Kota dan Kabupaten Malang. Dengan menggali informasi ini, diharapkan pula dapat mendorong langkah-langkah strategis yang lebih terarah untuk meningkatkan kualitas pendidikan di tingkat SMA Negeri, yang pada gilirannya dapat mendukung pembangunan daerah secara keseluruhan.

1.2. Rumusan Masalah

Dalam upaya mencapai tujuan pokok penelitian ini, kami mengajukan beberapa rumusan masalah yang sangat berkaitan dengan fokus analisis data distribusi dan zonasi SMA Negeri di Kota dan Kabupaten Malang. Berikut adalah rumusan masalah yang menjadi landasan utama dalam eksplorasi dan analisis kami:

1. Bagaimana distribusi dan persebaran SMA Negeri di Kota dan Kabupaten Malang?
2. Bagaimana pengaruh area pemukiman terhadap jarak antar SMA Negeri di Kota dan Kabupaten Malang?
3. Bagaimana aplikasi konsep agglomerative hierarchical clustering dalam menganalisis zonasi SMA Negeri di Kota dan Kabupaten Malang?

1.3. Tujuan dan Manfaat

Seiring dengan rumusan masalah yang telah diajukan, penelitian ini bertujuan untuk mencapai serangkaian tujuan untuk menghasilkan pemahaman yang lebih mendalam tentang distribusi dan zonasi SMA Negeri di Kota dan Kabupaten Malang.

Berikut adalah tujuan yang ingin dicapai dalam penelitian ini:

1. Menganalisis distribusi dan persebaran SMA Negeri di Kota dan Kabupaten Malang
2. Menentukan pengaruh area pemukiman terhadap jarak antar SMA Negeri di Kota dan Kabupaten Malang
3. Menerapkan konsep agglomerative hierarchical clustering sebagai dasar untuk analisis zonasi SMA Negeri di Kota dan Kabupaten Malang
4. Mengaplikasikan metode agglomerative clustering untuk mengelompokkan SMA Negeri berdasarkan karakteristik tertentu

Selain tujuan yang telah dijabarkan, penelitian ini juga diharapkan dapat memberikan manfaat baik secara akademis maupun praktis dan memberikan kontribusi positif terhadap pemahaman tentang zonasi SMA Negeri di Kota dan Kabupaten Malang, serta memberikan pandangan yang lebih holistik terhadap distribusi sekolah di wilayah tersebut. Adapun manfaat dari penelitian ini melibatkan beberapa aspek, antara lain:

1. Memberikan informasi mengenai distribusi dan persebaran SMA Negeri di Kota dan Kabupaten Malang

2. Memberikan pemahaman tentang pengaruh area pemukiman terhadap jarak antar SMA Negeri di Kota dan Kabupaten Malang yang dapat digunakan untuk perencanaan ruang dan infrastruktur pendidikan
3. Memberikan dasar pengembangan strategi zonasi pendidikan yang lebih efisien dan efektif di Kota dan Kabupaten Malang
4. Menyediakan pandangan yang jelas tentang karakteristik klaster SMA Negeri, memudahkan pengambilan keputusan terkait penyusunan kebijakan pendidikan.

BAB II

METODOLOGI PENELITIAN

2.1. Data

Data yang digunakan dalam penelitian ini yang akan diolah menggunakan Agglomerative Hierarchical Clustering diperoleh dari proses data collection atau web data gathering dengan teknik pengambilan dataset data aggregation, data yang digunakan bersumber dari integrasi data beberapa sumber utama, meliputi Data Pokok Pendidikan (Dapo) Kementerian Pendidikan dan Kebudayaan (Kemendikbud), Badan Pusat Statistik (BPS) Kota Malang, BPS Kabupaten Malang, serta situs resmi pemerintahan Kota dan Kabupaten Malang. Dataset mencakup informasi seputar SMA Negeri, wilayah, data penduduk, dan informasi kantor kecamatan. Pemilihan kantor kecamatan sebagai representasi dari wilayah pemukiman penduduk dikarenakan kantor kecamatan dapat mewakili keberagaman dan keunikan pemukiman penduduk di suatu wilayah yang mungkin mencakup berbagai kelurahan dan desa dan dapat dipastikan bahwa lokasi kantor kecamatan mudah dijangkau oleh masyarakat, sehingga memudahkan akses untuk kepentingan administratif dan pelayanan publik. Hasil data yang dikumpulkan diubah ke dalam 6 tabel yang dijelaskan sebagai berikut:

Tabel 1. Data SMA Negeri di Kota Malang

| No. | Nama SMA | Alamat | X (Latitude) | Y (Longitude) |
|-----|---------------|--|--------------|---------------|
| 1 | SMAN 1 Malang | Jl. Tugu No.1, Kiduldalem, Kec. Klojen, Kota Malang, Jawa Timur 65111 | -7,976,760 | 112,634,895 |
| 2 | SMAN 2 Malang | Jl. Laksamana Martadinata No.84, Sukoharjo, Kec. Klojen, Kota Malang, Jawa Timur 65118 | -7,991,366 | 112,633,858 |
| 3 | SMAN 3 Malang | Jl. Sultan Agung No.7, Klojen, Kec. Klojen, Kota Malang, Jawa Timur 65144 | -7,976,143 | 112,635,087 |
| 4 | SMAN 4 Malang | Jl. Tugu No.1, Klojen, Kec. Klojen, Kota Malang, Jawa | -7,976,357 | 112,634,413 |

| | | | | |
|----|------------------|--|------------|-------------|
| | | Timur 65111 | | |
| 5 | SMAN 5 Malang | Jl. Tanimbar No.24, Kasin, Kec. Klojen, Kota Malang, Jawa Timur 65117 | -7,989,148 | 112,626,048 |
| 6 | SMAN 6 Malang | Jl. Mayjen Sungkono No.58, Buring, Kec. Kedungkandang, Kota Malang, Jawa Timur 65136 | -8,010,425 | 112,643,253 |
| 7 | SMAN 7 Malang | Jl. Cengger Ayam I No.14, Tulusrejo, Kec. Lowokwaru, Kota Malang, Jawa Timur 65141 | -7,944,860 | 112,628,498 |
| 8 | SMAN 8 Malang | Jl. Veteran No.37, Sumbersari, Kec. Lowokwaru, Kota Malang, Jawa Timur 65145 | -7,956,918 | 112,616,872 |
| 9 | SMAN 9 Malang | Jl. Puncak Borobudur No.1, Mojolangu, Kec. Lowokwaru, Kota Malang, Jawa Timur 65142 | -7,936,135 | 112,625,221 |
| 10 | SMAN 10 Malang | Jl. Danau Grati No.1, Sawojajar, Kec. Kedungkandang, Kota Malang, Jawa Timur 65139 | -7,974,212 | 112,660,270 |
| 11 | SMAN Taruna Nala | Jl. Raya Tlogowaru, Tlogowaru, Kec. Kedungkandang, Kota Malang, Jawa Timur 65133 | -8.037968 | 112.656192 |

Tabel 1 menunjukkan data yang berisi informasi tentang alamat dan geoposition (latitude, longitude) dari setiap SMA Negeri di Kota Malang.

Tabel 2. Data SMA Negeri di Kabupaten Malang

| No. | Nama SMA | Alamat | X (Latitude) | Y (Longitude) |
|-----|-------------------|---|--------------|---------------|
| 1 | SMAN 1 Bantur | Jl. Raya Rejosari No.15, Balewarti, Rejosari, Kec. Bantur, Kabupaten Malang, Jawa Timur 65179 | -8,274,788 | 112,598,452 |
| 2 | SMAN 1 Bululawang | Jl. Raya Bululawang No.17, Bululawang, Kec. Bululawang, | -8,072,094 | 112,641,945 |

| | | | | |
|----|-------------------------|--|------------|-------------|
| | | Kabupaten Malang, Jawa Timur 65171 | | |
| 3 | SMAN 1 Dampit | Jl. Gn. Jati, Sumberkembar, Dampit, Kec. Dampit, Kabupaten Malang, Jawa Timur 65181 | -8,210,020 | 112,766,054 |
| 4 | SMAN 1 Gondanglegi | Jl. Raya Ketawang No.02, Krajan, Ketawang, Kec. Gondanglegi, Kabupaten Malang, Jawa Timur 65174 | -8,136,982 | 112,644,245 |
| 5 | SMAN 1 Kepanjen | Jl. Ahmad Yani No.48, Ardirejo, Kec. Kepanjen, Kabupaten Malang, Jawa Timur 65163 | -8,124,450 | 112,572,355 |
| 6 | SMAN 1 Lawang | Jl. Pramuka No.152, Krajan, Kalirejo, Kec. Lawang, Kabupaten Malang, Jawa Timur 65216 | -7,841,512 | 112,699,988 |
| 7 | SMAN 1 Ngantang | Jl. Raya Ngantang N0.253, Mulyorejo, Kec. Ngantang, Kabupaten Malang, Jawa Timur, 65392 | -8,160,868 | 112,732,417 |
| 8 | SMAN 1 Pagak | Jl. Kahuripan No.4, Kulon Kali, Sumbermanjing Kulon, Kec. Pagak, Kabupaten Malang, Jawa Timur 65168 | -8,299,092 | 112,493,764 |
| 9 | SMAN 1 Singosari | Jl. Ki Hajar Dewantara Desa No.1, Tanjung, Banjararum, Kec. Singosari, Kabupaten Malang, Jawa Timur 65153 | -7,918,600 | 112,673,066 |
| 10 | SMAN 1 Sumbermanjing | Jl. Raya, Sidomulyo Kulon, Tambakasri, Kec. Sumbermanjing Wetan, Kabupaten Malang, Jawa Timur 65176 | -8,351,517 | 112,752,721 |

| | | | | |
|----|------------------------|---|------------|-------------|
| 11 | SMAN 1 Sumberpucung | Jl. Raya No.58, Mentaraman, Jatiguwi, Kec. Sumberpucung, Kabupaten Malang, Jawa Timur 65165 | -8,149,623 | 112,495,908 |
| 12 | SMAN 1 Tumpang | Jl. Kamboja No.10, 65156, Malangsuko, Kec. Tumpang, Kabupaten Malang, Jawa Timur 65313 | -8,000,145 | 112,763,455 |
| 13 | SMAN 1 Turen | Jl. Mayjen Panjaitan No.65, Dusun Sedayu, Sedayu, Kec. Turen, Kabupaten Malang, Jawa Timur 65175 | -8,183,645 | 112,679,743 |

Tabel 2 menunjukkan data yang berisi informasi tentang alamat dan geoposition (latitude, longitude) dari setiap SMA Negeri di Kabupaten Malang.

Tabel 3. Data Wilayah, Penduduk, dan Jumlah SMA Negeri Per Kecamatan di Kota Malang
(Tahun 2020 - 2022)

| No. | Kecamatan | Luas Wilayah (Km ²) | Jumlah Penduduk | Kepadatan Penduduk | Jumlah SMA Negeri |
|-------|---------------|------------------------------------|--------------------|-----------------------|----------------------|
| 1 | Blimbing | 17.73 | 182693 | 10303.59 | 5 |
| 2 | Kedungkandang | 39.85 | 208075 | 5221.19 | 3 |
| 3 | Klojen | 8.83 | 94072 | 10653.68 | 3 |
| 4 | Lowokwaru | 22.60 | 163795 | 7247.57 | 0 |
| 5 | Sukun | 20.97 | 196487 | 9369.91 | 0 |
| Total | | 109.98 | 845122 | 7684.114818 | 11 |

Tabel 3 menunjukkan data yang berisi informasi tentang luas wilayah, jumlah penduduk, dan kepadatan penduduk di setiap Kecamatan di Kota Malang. Pada data ini juga berisi informasi jumlah SMA Negeri yang berada di masing-masing kecamatan Kota Malang.

Tabel 4. Data Wilayah, Penduduk, dan Jumlah SMA Negeri Per Kecamatan di Kabupaten Malang
 (Tahun 2015 - 2020)

| No. | Kecamatan | Luas Wilayah (Km ²) | Jumlah Penduduk | Kepadatan Penduduk | Jumlah SMA Negeri |
|-----|---------------|---------------------------------|-----------------|--------------------|-------------------|
| 1 | Ampelgading | 79.60 | 52000 | 653.27 | 0 |
| 2 | Bantur | 159.15 | 68824 | 432.45 | 1 |
| 3 | Bululawang | 49.36 | 72917 | 1477.25 | 1 |
| 4 | Dampit | 135.31 | 118479 | 875.61 | 1 |
| 5 | Dau | 41.96 | 82220 | 1959.49 | 0 |
| 6 | Donomulyo | 192.60 | 62585 | 324.95 | 0 |
| 7 | Gedangan | 130.55 | 53289 | 408.19 | 0 |
| 8 | Gondanglegi | 79.74 | 86796 | 1088.49 | 1 |
| 9 | Jabung | 135.89 | 75365 | 554.6 | 0 |
| 10 | Kalipare | 105.39 | 59545 | 565 | 0 |
| 11 | Karangploso | 58.74 | 89032 | 1515.7 | 0 |
| 12 | Kasembon | 55.67 | 31595 | 567.54 | 0 |
| 13 | Kepanjen | 46.25 | 109634 | 2370.46 | 1 |
| 14 | Kromengan | 38.63 | 38033 | 984.55 | 0 |
| 15 | Lawang | 68.23 | 114928 | 1684.42 | 1 |
| 16 | Ngajum | 60.12 | 49504 | 823.42 | 0 |
| 17 | Ngantang | 147.70 | 56376 | 381.69 | 1 |
| 18 | Pagak | 90.08 | 45597 | 506.18 | 1 |
| 19 | Pagelaran | 45.83 | 68147 | 1486.95 | 0 |
| 20 | Pakis | 53.62 | 171657 | 3201.36 | 0 |
| 21 | Pakisaji | 38.41 | 93157 | 2425.33 | 0 |
| 22 | Poncokusumo | 102.99 | 92648 | 899.58 | 0 |
| 23 | Pujon | 130.75 | 69040 | 528.03 | 0 |
| 24 | Singosari | 118.51 | 190487 | 1607.35 | 1 |
| 25 | Sumbermanjing | 239.49 | 89928 | 375.5 | 1 |

| | | | | | |
|-------|--------------|----------|---------|-------------|----|
| | Wetan | | | | |
| 26 | Sumberpubung | 35.90 | 55460 | 1544.85 | 1 |
| 27 | Tajinan | 40.11 | 55119 | 1374.2 | 0 |
| 28 | Tirtoyudo | 141.96 | 60928 | 429.19 | 0 |
| 29 | Tumpang | 72.09 | 75657 | 1049.48 | 1 |
| 30 | Turen | 63.90 | 115290 | 1804.23 | 1 |
| 31 | Wagir | 75.43 | 93211 | 1235.73 | 0 |
| 32 | Wajak | 94.56 | 81170 | 858.4 | 0 |
| 33 | Wonosari | 48.53 | 41357 | 852.19 | 0 |
| Total | | 2,977.05 | 2619975 | 880.0574394 | 13 |

Tabel 4 menunjukkan data yang berisi informasi tentang luas wilayah, jumlah penduduk, dan kepadatan penduduk di setiap Kecamatan di Kabupaten Malang. Pada data ini juga berisi informasi jumlah SMA Negeri yang berada di masing-masing kecamatan Kabupaten Malang.

Tabel 5. Data Kantor Kecamatan di Kota Malang

| No. | Kecamatan | Alamat | X (Latitude) | Y (Longitude) |
|-----|---------------|--|--------------|---------------|
| 1 | Blimbing | Jl. Raden Intan No.14, Polowijen, Kec. Blimbing, Kota Malang, Jawa Timur 65125 | -7.9303969 | 112.651949 |
| 2 | Kedungkandang | Jl. Mayjen Sungkono No.59, Buring, Kec. Kedungkandang, Kota Malang, Jawa Timur 65136 | -8.0118966 | 112.643807 |
| 3 | Klojen | Jl. Surabaya No.3, 65112, Sumbersari, Kec. Lowokwaru, Kota Malang, Jawa Timur 65115 | -7.9663394 | 112.6179851 |
| 4 | Lowokwaru | Jl. Cengger Ayam I No.12, Tulusrejo, Kec. Lowokwaru, Kota Malang, Jawa Timur 65141 | -7.946127 | 112.630495 |
| 5 | Sukun | Jl. Keben II, Bandungrejosari, Kec. Sukun, Kota Malang, Jawa Timur 65148 | -8.0050679 | 112.6186244 |

Tabel 5 menunjukkan data yang berisi informasi tentang alamat dan geoposition (latitude, longitude) dari setiap kantor kecamatan di Kota Malang.

Tabel 6. Data Kantor Kecamatan di Kabupaten Malang

| No. | Kecamatan | Alamat | X (Latitude) | Y (Longitude) |
|-----|-------------|--|--------------|---------------|
| 1 | Ampelgading | Jl. Raya, Bahroto, Tirtomarto, Kec. Ampelgading, Kabupaten Malang, Jawa Timur 65183 | -6.9233033 | 109.5105354 |
| 2 | Bantur | Banturkrajan, Bantur, Kec. Bantur, Kabupaten Malang, Jawa Timur 65179 | -8.3155405 | 112.5804951 |
| 3 | Bululawang | Jl. Suropati Raya No.6, Bululawang, Blayu, Kec. Bululawang, Kabupaten Malang, Jawa Timur 65173 | -8.0779069 | 112.641046 |
| 4 | Dampit | Jl. Semeru Sel. No.23, Dampit, Kec. Dampit, Kabupaten Malang, Jawa Timur 65181 | -8.2117469 | 112.74857 |
| 5 | Dau | Jl. Raya Sengkaling No.200, Sengkaling, Mulyoagung, Kec. Dau, Kabupaten Malang, Jawa Timur 65151 | -7.91433 | 112.5858887 |
| 6 | Donomulyo | Mulyosari, Dono Mulyo, Kec. Donomulyo, Kabupaten Malang, Jawa Timur 65167 | -8.2854299 | 112.4278789 |
| 7 | Gedangan | Jl. Raya Hassanudin No.180, Sumbernanas, Gedangan, Malang, Kabupaten Malang, Jawa Timur 65178 | -7.389735 | 112.726687 |
| 8 | Gondanglegi | Jl. Diponegoro No.76, Krajan Satu, Gondanglegi Kulon, Kec. Gondanglegi, Kabupaten Malang, Jawa Timur 65174 | -8.1762581 | 112.6356317 |
| 9 | Jabung | Jl. Raya Jabung No.22, Dusun Gandon Timur, Sukolilo, Kec. Jabung, Kabupaten Malang, Jawa Timur 65155 | -7.9451316 | 112.734142 |
| 10 | Kalipare | Kalipare, Kabupaten Malang, Jawa Timur 65166 | -8.1829963 | 112.4239395 |
| 11 | Karangploso | Jl. Diponegoro No.70, Girimoyo, Karangploso, Ngambon, Girimoyo, | -7.891783 | 112.594125 |

| | | | | |
|----|------------|--|------------|-------------|
| | | Kec. Karang Ploso, Kabupaten Malang, Jawa Timur 65152 | | |
| 12 | Kasembon | Kantor Kecamatan Kasembon no 7 Kasembon Malang, Sanggrahan Kidul, Kasembon, Kec. Kasembon, Kabupaten Malang, Jawa Timur 65393 | -7.7838423 | 112.3110636 |
| 13 | Kepanjen | Jl. Kawi No.50, Banurejo, Kepanjen, Kec. Kepanjen, Kabupaten Malang, Jawa Timur 65163 | -8.1300445 | 112.5661771 |
| 14 | Kromengan | Jl. Nailun Sel., Krajan, Kromengan, Malang, Kabupaten Malang, Jawa Timur 65165 | -8.130466 | 112.4900433 |
| 15 | Lawang | Jl. Thamrin No.2, Lawang, Kec. Lawang, Kabupaten Malang, Jawa Timur 65211 | -7.830759 | 112.697098 |
| 16 | Ngajum | Jl. Ahmad Yani No.76, Ngajum, Kec. Ngajum, Kabupaten Malang, Jawa Timur 65164 | -8.0965217 | 112.5408668 |
| 17 | Ngantang | Jl. Raya Ngantang No.68, Kaumrejo, Ngantang, Prabon I, Kaumrejo, Kec. Ngantang, Kabupaten Malang, Jawa Timur 65392 | -7.8468018 | 112.3718771 |
| 18 | Pagak | Jl. Abdul Razak No.1, Krajan, Pagak, Kec. Pagak, Kabupaten Malang, Jawa Timur 35168 | -8.2256649 | 112.5273673 |
| 19 | Pagelaran | Jl. Panglima Sudirman No.21, Pagelaran, Kantor, Kec. Pagelaran, Kabupaten Malang, Jawa Timur 65174 | -8.1836439 | 112.5796201 |
| 20 | Pakis | Jl. Raya Pakis No.69, Krajan, Pakisjajar, Kec. Pakis, Kabupaten Malang, Jawa Timur 65154 | -7.9592009 | 112.7168389 |
| 21 | Pakisaji | Jl. Raya Pakisaji No.9, Jatirejo, Pakisaji, Kec. Pakisaji, Kabupaten Malang, Jawa Timur 65162 | -8.0610009 | 112.601334 |
| 22 | Ponokusumo | Jl. Raya Wonorejo No.4, Wonorejo, Kec. Ponokusumo, Kabupaten | -8.0423999 | 112.772821 |

| | | | | |
|----|------------------------|---|------------|-------------|
| | | Malang, Jawa Timur 65157 | | |
| 23 | Pujon | Jl. Brigjen Abdul Manan No.08, Jurangrejo, Pandesari, Kec. Pujon, Kabupaten Malang, Jawa Timur 65391 | -7.8496549 | 112.4803817 |
| 24 | Singosari | Jl. Raya Tumapel No.38, Pangetan, Pagantan, Kec. Singosari, Kabupaten Malang, Jawa Timur 65153 | -7.8939972 | 112.662854 |
| 25 | Sumbermanjing Wetan | Jl. Raya Sumbermanjing Wetan, Sumbermanjing Wetan, Argotirto, Kec. Sumbermanjing Wetan, Kabupaten Malang, Jawa Timur 65176 | -8.2792197 | 112.6900655 |
| 26 | Sumberpucung | Jl. Jend. Sudirman No.277, Dusun Suko, Sumberpucung, Kec. Sumberpucung, Kabupaten Malang, Jawa Timur 65165 | -8.1568074 | 112.474058 |
| 27 | Tajinan | Jl. Raya Tajinan No.18, Tajinan, Kec. Tajinan, Kabupaten Malang, Jawa Timur 65172 | -8.0497961 | 112.6817364 |
| 28 | Tirtoyudo | Jl. Tirtoyudo, Tlogosari, Kec. Tirtoyudo, Kabupaten Malang, Jawa Timur 65182 | -8.2272887 | 112.8331393 |
| 29 | Tumpang | Jl. Raya Malangsuko No.1A, Malangsuko, Kec. Tumpang, Kabupaten Malang, Jawa Timur 65156 | -8.002422 | 112.761582 |
| 30 | Turen | Jl. Panglima Sudirman No.102, Turen, Kec. Turen, Kabupaten Malang, Jawa Timur 65175 | -8.1665813 | 112.6918728 |
| 31 | Wagir | Jl. Raya Gondowangi No.3, Gedangan, Gondowangi, Kec. Wagir, Kabupaten Malang, Jawa Timur 65158 | -8.0089501 | 112.5724535 |
| 32 | Wajak | Jl. Panglima Sudirman No.96, Wajak, Kec. Wajak, Kabupaten Malang, Jawa Timur 65173 | -8.1024549 | 112.730147 |
| 33 | Wonosari | Jl. Raya Kebobang No.12, Wonosari, Kluwut, Kec. Wonosari, Kabupaten Malang, Jawa Timur 65164 | -8.033319 | 112.497693 |

Tabel 6 menunjukkan data yang berisi informasi tentang alamat dan geoposition (latitude, longitude) dari setiap kantor kecamatan di Kabupaten Malang.

2.2. Agglomerative Hierarchical Clustering

2.2.1. Clustering

Analisis *cluster* adalah teknik analisis *multivariat* dengan tujuan mengelompokkan data observasi atau variabel pengamatan ke dalam *cluster* sedemikian sehingga masing-masing bersifat homogen. Proses pembentukan *cluster* dapat dilakukan dengan dua cara, yaitu dengan metode *hierarchical* dan *non-hierarchical*. Pada metode hirarki terdiri dari metode *agglomerative* dan metode *deficit*. Metode *agglomerative* sendiri terdiri dari 3 metode, yaitu metode *linkage*, metode *variance*, dan metode *centroid*, dimana *linkage* terdiri dari metode *single linkage*, *complete linkage*, dan *average linkage*.

[1]

2.2.2. Agglomerative Hierarchical Clustering

Agglomerative hierarchical clustering merupakan suatu pendekatan pengelompokan hierarki yang menggunakan metode dari bawah ke atas. Proses ini dimulai dengan menganggap setiap data sebagai sebuah kelompok, lalu mencari kelompok potensial berdasarkan jarak antara mereka untuk kemudian digabungkan menjadi kelompok yang lebih besar (*bottom-up*). Proses ini terus berulang, secara bertahap bergerak ke atas (aglomeratif), dan pada akhirnya membentuk struktur hierarki.[2]

2.3. Sistem Zonasi di Kota dan Kabupaten Malang

Penerimaan Peserta Didik Baru (PPDB) di Kota dan Kabupaten Malang, Provinsi Jawa Timur, merupakan proses yang melibatkan beberapa jalur untuk memberikan kesempatan akses pendidikan kepada calon siswa. Sistem ini didesain untuk mencakup beragam kondisi dan kebutuhan calon siswa sehingga dapat menciptakan penerimaan yang adil dan merata. Jalur-jalur yang tersedia dalam PPDB mencakup zonasi, afirmasi, perpindahan tugas orang tua/wali, dan prestasi.

Sistem zonasi adalah pembagian wilayah kedalam beberapa zona. Zonasi sendiri berasal dari kata zona yaitu kawasan atau area yang memiliki fungsi dan karakteristik lingkungan yang spesifik 11. Zonasi dimaknai sebagai pembagian atau pemecahan suatu areal menjadi beberapa bagian, sesuai dengan fungsi dan tujuan

pengelolaan.¹² Dengan kata lain kebijakan zonasi dipandang sebagai solusi untuk menyelesaikan dua masalah pokok pendidikan, yaitu pemerataan akses dan kualitas pendidikan.¹³ Kebijakan sistem zonasi merupakan kebijakan dalam rangka manajemen peserta didik yang mulai diberlakukan pada tahun ajaran 2017/2018. Dengan sistem zonasi semua (khususnya sekolah negeri) disiapkan untuk memberikan layanan pendidikan yang bermutu secara merata bagi warga anggota masyarakat pada suatu areal atau kawasan tertentu sehingga anak- “anak terbaik” tidak perlu mencari “sekolah terbaik” yang lokasinya jauh dari tempat tinggalnya. Berdasarkan Permendikbud Nomor 17 Tahun 2017, dengan menerapkan sistem zonasi, sekolah yang diselenggarakan oleh

Berdasarkan situs resmi PPDB Jawa Timur 2023, PPDB SMAN di Kota Malang terbagi atas 3 zona. Zona 1 yang terdiri dari SMA Negeri 4 Malang, SMA Negeri 5 Malang, dan SMA Negeri 7 Malang meliputi Kecamatan Blimbingsari, Klojen, Lowokwaru, Sukun dan Wagir (Kab.Malang). Zona 2 terdiri dari SMA Negeri 1 Malang, SMA Negeri 8 Malang, dan SMA Negeri 9 Malang meliputi Kecamatan Blimbingsari, Kedungkandang, Lowokwaru, Dau (Kab.Malang), Karangploso (Kab. Malang), Singosari (Kab.Malang). Dan Zona 3 terdiri dari SMA Negeri 2 Malang, SMA Negeri 3 Malang, SMA Negeri 6 Malang, dan SMA Negeri 10 Malang meliputi kecamatan Kedungkandang, Klojen, Sukun, Tajinan (Kab.Malang), dan Pakis (Kab.Malang)

Sementara PPDB SMAN di Kabupaten Malang terbagi menjadi 5 zona. Zona 1 meliputi Dau, Karangploso, Kasembon, Lawang, Ngantang, Pujon dan Singosari terdiri dari SMA Negeri 1 Ngantang, SMA Negeri 1 Lawang, dan SMA Negeri 1 Singosari. Di Zona 2 meliputi Kecamatan Bantur, Bululawang, Donomulyo, Gondanglegi, Jabung, Kalipare, Pagak, Pagelaran, Pakis, Poncokusumo, Tajinan, Tumpang, dan Wajak meliputi SMA Negeri 1 Bantur, SMA Negeri 1 Bululawang, SMA Negeri 1 Gondanglegi, SMA Negeri 1 Pagak, SMA Negeri 1 Tumpang. Zona 3 meliputi Ampelgading, Dampit, Sumbermanjing Wetan, Turen, dan Tirtoyudo yang terdiri dari SMA Negeri 1 Turen, SMA Negeri Dampit, dan SMA Negeri Sumbermanjing Wetan. Sementara Zona 4 meliputi Kromengan, Nganjum, Pakisaji, Sumberpucung, Wagir, dan Wonosari yang terdiri dari SMA Negeri 1 Kepanjen, dan SMA Negeri 1 Sumberpucung

2.4 Outliers

Outliers, atau yang sering disebut sebagai nilai data ekstrem, merupakan konsep yang penting dalam analisis statistik. Outliers dapat didefinisikan sebagai pengamatan yang "terlihat" tidak konsisten dengan pengamatan lain dalam kumpulan data (Aggarwal & Aggarwal, 2017) (Iglewicz & Hoaglin, 2017). Pengertian outlier ini lebih lanjut menjelaskan bahwa outlier memiliki probabilitas rendah bahwa ia berasal dari distribusi statistik yang sama dengan pengamatan lain dalam kumpulan data tersebut. Dengan kata lain, outlier adalah data yang secara statistik signifikan berbeda dari mayoritas data lainnya dalam kumpulan data.

Outliers dapat memberikan informasi yang berharga tentang suatu proses. Mereka dapat muncul akibat perubahan dalam lokasi (rerata) atau skala (variabilitas) dari proses. Beberapa kali, nilai data yang mencurigakan ini mungkin disebabkan oleh kesalahan pengukuran, kesalahan pemasukan data, ataupun memang karena terdapat nilai yang benar-benar berbeda. Kesalahan pengukuran dapat terjadi karena berbagai faktor, seperti kesalahan alat ukur, kesalahan operator, atau kondisi lingkungan yang tidak sesuai.. Kemudian kesalahan pemasukan data dapat terjadi karena human error, seperti salah ketik atau salah memasukkan data. Kesalahan selanjutnya yaitu nilai outlier juga dapat terjadi karena nilai tersebut memang benar-benar berbeda dari nilai data lainnya. Misalnya, dalam kumpulan data hasil pengukuran berat badan balita, nilai outlier dapat berupa nilai berat badan orang dewasa yang secara logis tidak mungkin terjadi.

Pertanyaan yang muncul adalah bagaimana seharusnya outliers diatasi? Setelah identifikasi awal sebuah pengamatan sebagai outlier, kemudian dilakukan analisis grafik secara visual. Analisis tersebut dilakukan untuk menentukan apakah ada penyebab yang dapat ditemukan untuk hasil yang mencurigakan tersebut. Jika memang terdapat kesalahan input/pengukuran dan pengulangan tes data yang outlier tersebut dapat dilaksanakan kembali, maka data yang outlier tersebut dapat diperbarui dengan data yang akan diukur lagi. Seringkali, nilai yang tampak sebagai outliers adalah ekstrem di bagian ekor kanan dari distribusi yang condong. (Aggarwal & Aggarwal, 2017).

Terdapat berbagai pendekatan untuk mendeteksi outliers tergantung pada aplikasi dan jumlah pengamatan dalam kumpulan data. Penting untuk diingat bahwa inspeksi visual saja tidak selalu dapat mengidentifikasi outlier dan dapat

menyebabkan kesalahan dalam memberi label sebuah pengamatan sebagai outlier. Oleh karena itu, menggunakan fungsi spesifik dari pengamatan dapat menghasilkan aturan penandaan outlier yang lebih baik. Karena data digunakan dalam estimasi dengan metode klasik seperti rerata sangat sensitif terhadap outliers, metode statistik dikembangkan untuk mengakomodasi outliers dan mengurangi dampaknya terhadap analisis. Beberapa metode identifikasi yang umum digunakan dibahas dalam laporan ini.

2.5 Tools untuk Mengolah Data

Tools yang digunakan dalam melakukan analisis zonasi dan persebaran SMA Negeri di Kota dan Kabupaten Malang dengan menggunakan metode Agglomerative Clustering adalah Google Maps. Google Maps adalah layanan pemetaan web yang dikembangkan oleh Google. Google maps dipilih sebagai tools dalam penelitian ini sebagai alat untuk mengakses dan koordinat antar SMA Negeri dengan kecamatan yang merupakan titik clustering. Selanjutnya, koordinat SMA dan Kecamatan akan dicatat dengan *tools* kedua yaitu Excel. Digunakan Excel dalam pencatatan data karena kemampuan mendukung berbagai format data, memudahkan penggunaan dan pengelolaan informasi. Fungsi kolaboratifnya memungkinkan beberapa orang bekerja secara bersamaan (kolaborasi tim) untuk analisis data yang efisien, sementara kemampuannya untuk menyimpan riwayat perubahan dokumentasi dan pelacakan revisi data (audit dan manajemen perubahan).

Untuk mengolah data, digunakan *tools* untuk membantu perhitungan dan membuat visualisasi dari hasil *Agglomerative Hierarchical Clustering* ini sendiri. *Tools* untuk melakukan ini yaitu R Studio karena aplikasi ini memiliki berbagai fungsi dan pustaka yang mendukung analisis data statistik, termasuk analisis cluster. R studio menyediakan fungsi `hclust()` yang memiliki berbagai parameter yang dapat digunakan untuk menyesuaikan algoritma agglomerative hierarchical clustering, seperti jarak yang digunakan dan metode linkage yang digunakan. R studio juga memiliki fungsi `plot()` untuk memvisualisasikan dendogram yang menunjukkan bagaimana cluster terbentuk dari objek-objek data hasil clustering.

2.6 Langkah-langkah Agglomerative Hierarchical Clustering

Untuk melakukan analisis zonasi dan persebaran Sekolah Menengah Atas (SMA) Negeri di Kota dan Kabupaten Malang dengan metode Agglomerative Clustering menggunakan R Studio, langkah awal yang harus dilakukan adalah menggunakan library dendextend. Library ini berguna untuk melakukan manipulasi dan visualisasi dendrogram, yang merupakan representasi diagram hirarki. Selain itu, digunakan juga library readxl untuk membaca data dari file Excel. Pada kasus ini, fungsi readxl digunakan untuk membaca file excel dari ‘Data_Malang’, sehingga informasi mengenai lokasi dan atribut SMA Negeri di wilayah tersebut dapat diakses dan dianalisis lebih lanjut

```
library(dendextend)
# Membaca data dari file Excel
library(readxl)
FP <- read_excel("Data_Malang.xlsx")
```

Gambar 1 Langkah-langkah Agglomerative Hierarchical Clustering

Setelah membaca data, langkah selanjutnya adalah menyimpan keseluruhan dataset ke dalam variabel yang dinamakan ‘Data’

```
# Menyimpan data ke dalam variabel Data
Data <- FP
```

Gambar 2 Langkah-langkah Agglomerative Hierarchical Clustering

Langkah selanjutnya adalah memahami struktur dan konten dari dataset tersebut. Hal yang perlu dilakukan adalah pengecekan struktur data dengan menggunakan fungsi ‘str’, yang memberikan gambaran menyeluruh mengenai jenis dan tipe data dari setiap kolom dalam dataset. Setelah itu dilakukan evaluasi beberapa baris pertama dari dataset dengan menggunakan fungsi ‘head’, hal ini memberikan preview awal dari data.

```
# Menampilkan struktur data dan beberapa baris pertama
str(Data)
head(Data)
```

Gambar 3 Langkah-langkah Agglomerative Hierarchical Clustering

Setelah itu langkah berikutnya yang dilakukan adalah memvisualisasikan informasi geografis dengan membuat scatter plot Latitude vs Longitude. Ini dilakukan menggunakan fungsi ‘plot’, di mana sumbu x dan y merepresentasikan latitude dan longitude. Scatter plot ini memberikan gambar visual yang berguna terkait distribusi geografis dari titik-titik data. Untuk meningkatkan interpretasi, label berisi informasi Nama ditambahkan pada setiap titik plot.

```
# Membuat scatter plot
plot(Data$Y (Longitude) ~ Data$X (Latitude),
     main = "Scatter Plot of Latitude vs Longitude",
     xlab = "Latitude", ylab = "Longitude")
text(Data$Y (Longitude), Data$X (Latitude),
     labels = Data$Nama, pos = 3, cex = 0.8)
```

Gambar 4 Langkah-langkah Agglomerative Hierarchical Clustering

Parameter yang terdapat dalam kode diatas seperti ‘main’, ‘xlab’, dan ‘ylab’ memberikan judul dan label pada plot. Fungsi ‘text’ digunakan untuk menambahkan label pada setiap titik plot dengan informasi Nama dari dataset, posisi teks diatur di atas setiap titik (pos = 3), dan ukuran teks dikontrol dengan parameter ‘cex’. Selanjutnya dilakukan normalisasi data pada dataset yang telah dibaca dan disimpan dalam variabel ‘Data’. Pertama, kolom pertama dan kedua dari dataset, yang berisi informasi "Nama" dan "ID", dihapus menggunakan operator ‘-c(1, 2)’ dan hasilnya disimpan dalam variabel ‘z’. Langkah ini bertujuan untuk menghilangkan atribut yang tidak diperlukan dalam analisis clustering geografis. Selanjutnya, untuk setiap kolom dalam ‘z’, rata-rata ‘mean’ dan deviasi standar ‘sd’ dihitung menggunakan fungsi ‘apply’. Data kemudian dinormalisasi dengan menggunakan fungsi ‘scale’, di mana setiap nilai dalam dataset diubah sedemikian rupa sehingga memiliki rata-rata nol dan deviasi standar satu

```
# Normalisasi
z <- Data[, -c(1, 2)]
m <- apply(z, 2, mean)
s <- apply(z, 2, sd)
z <- scale(z, m, s)
```

Gambar 5 Langkah-langkah Agglomerative Hierarchical Clustering

Langkah selanjutnya adalah menghitung matriks Euclidean antar observasi dalam dataset yang telah dinormalisasi ‘z’

```
# Menghitung Euclidean Distance
distance <- dist(z)
print(distance, digits = 3)
```

Gambar 6 Langkah-langkah Agglomerative Hierarchical Clustering

Pertama, fungsi dist digunakan untuk menghitung jarak Euclidean antar baris data, menghasilkan matriks jarak. Hasil perhitungan tersebut disimpan dalam variabel distance. Dengan menciptakan matriks jarak, kita memperoleh informasi tentang seberapa dekat atau jauh setiap pasangan observasi dalam ruang Euclidean. Kemudian, fungsi print digunakan untuk menampilkan matriks jarak dengan batasan angka desimal hingga tiga digit untuk memperjelas tampilan output. Matriks jarak ini akan menjadi dasar untuk proses clustering selanjutnya, membantu dalam pembentukan dendrogram dan identifikasi pola kelompok dalam dataset.

Setelah itu data diproses menggunakan metode single linkage pada dataset yang telah dinormalisasikan.

```
# Clustering Dendrogram (Single Linkage)
hc.1 <- hclust(distance)
dend <- as.dendrogram(hc.1)
labels(dend) <- Data$Nama
dend <- hang.dendrogram(dend, hang_height = -0.1)
plot(dend, main = "Dendrogram Hierarchical Clustering (Single Linkage)", xlab = "SMA Negeri & Kantor Kecamatan", cex = 2)
rect.hclust(hc.1, k = 4, border = c("blue", "green", "orange", "red"))
```

Gambar 7 Langkah-langkah Agglomerative Hierarchical Clustering

Pertama, matriks jarak Euclidean yang telah dihitung sebelumnya ‘distance’ digunakan untuk membentuk hirarki clustering dengan fungsi hclust. Hasilnya disimpan dalam variabel hc.1. Selanjutnya, dendrogram dibentuk dari hasil clustering dengan menggunakan fungsi as.dendrogram. Label pada dendrogram diatur sesuai dengan nilai "Nama" dari dataset menggunakan perintah labels(dend)← Data\$Nama. Dendrogram tersebut kemudian dimodifikasi dengan menggeser tinggi cabang menggunakan hang.dendrogram untuk meningkatkan kejelasan visualisasi. Plot dari dendrogram ditampilkan dengan fungsi plot, dan judul, label sumbu x, dan ukuran font diatur untuk memberikan informasi kontekstual yang jelas. Terakhir, fungsi rect.hclust digunakan untuk menandai pembagian data ke dalam empat klaster dengan warna batas yang berbeda (biru, hijau, oranye, dan merah). Kemudian data dihitung menggunakan agglomerative hierarchical clustering menggunakan metode average linkage pada dataset yang telah dinormalisasi.

```
# Clustering Dendrogram (Average Linkage)
hc.a <- hclust(distance, method = "average")
dend <- as.dendrogram(hc.a)
labels(dend) <- Data$Nama # Mengganti Label dengan Data$Nama
dend <- hang.dendrogram(dend, hang_height = 0.1)
plot(dend, main = "Dendrogram Hierarchical Clustering (Average Linkage)", xlab = "SMA Negeri & Kantor Kecamatan", cex = 2)
rect.hclust(hc.a, k = 4, border = c("blue", "green", "orange", "red"))
```

Gambar 8 Langkah-langkah Agglomerative Hierarchical Clustering

Matriks jarak Euclidean ‘distance’ digunakan dalam fungsi hclust dengan opsi method = "average", menghasilkan dendrogram clustering yang disimpan dalam variabel hc.a. Dendrogram tersebut kemudian diubah menjadi format dendrogram menggunakan as.dendrogram, dan label pada cabang dendrogram diatur sesuai dengan nilai kolom "Nama" dari dataset menggunakan labels(dend)←Data\$Nama. Selanjutnya, tinggi cabang dendrogram dimodifikasi dengan menggunakan fungsi hang.dendrogram untuk meningkatkan kejelasan visualisasi. Plot dari dendrogram ditampilkan dengan judul dan label sumbu x yang sesuai dengan konteks analisis. Terakhir, fungsi rect.hclust digunakan untuk menandai pembagian data ke dalam empat klaster dengan warna batas yang berbeda (biru, hijau, oranye, dan merah).

Lalu data dihitung menggunakan agglomerative hierarchical clustering dengan metode complete linkage pada dataset yang telah dinormalisasi.

```
# Clustering Dendrogram (Complete Linkage)
hc.c <- hclust(distance, method = "complete")
dend <- as.dendrogram(hc.c)
labels(dend) <- Data$Nama # Mengganti label dengan Data$Nama
dend <- hang.dendrogram(dend, hang_height = 0.1)
plot(dend, main = "Dendrogram Hierarchical Clustering (Complete Linkage)", xlab = "SMA Negeri & Kantor Kecamatan", cex = 2)
rect.hclust(hc.c, k = 4, border = c("blue", "green", "orange", "red"))
```

Gambar 9 Langkah-langkah Agglomerative Hierarchical Clustering

Matriks jarak Euclidean ‘distance’ digunakan dalam fungsi hclust dengan opsi method = "complete", menghasilkan dendrogram clustering yang disimpan dalam variabel hc.c. Dendrogram tersebut diubah menjadi format dendrogram menggunakan as.dendrogram, dan label pada cabang dendrogram diatur sesuai dengan nilai kolom "Nama" dari dataset dengan perintah labels(dend)←Data\$Nama. Selanjutnya, tinggi cabang dendrogram dimodifikasi dengan menggunakan fungsi hang.dendrogram untuk meningkatkan kejelasan visualisasi. Plot dari dendrogram ditampilkan dengan judul dan label sumbu x yang sesuai dengan konteks analisis. Terakhir, fungsi rect.hclust digunakan untuk menandai pembagian data ke dalam empat klaster dengan warna batas yang berbeda (biru, hijau, oranye, dan merah). Setelah itu hasil clustering ditampilkan yang dihasilkan dari tiga metode agglomerative hierarchical clustering yang telah dilakukan sebelumnya (single linkage, average linkage, dan complete linkage).

```
# Menampilkan hasil clustering
member.l <- cutree(hc.l, 4)
clustersl <- data.frame(School = Data$Nama, Cluster = member.l)
print(clustersl)

member.a <- cutree(hc.a, 4)
clustersa <- data.frame(School = Data$Nama, Cluster = member.a)
print(clustersa)

member.c <- cutree(hc.c, 4)
clustersc <- data.frame(School = Data$Nama, Cluster = member.c)
print(clustersc)
```

Gambar 10 Langkah-langkah Agglomerative Hierarchical Clustering

Penggunaan fungsi cutree digunakan untuk memotong dendrogram yang telah dibentuk sebelumnya menjadi empat klaster, sesuai dengan jumlah klaster yang diinginkan (k=4). Hasil klasterisasi kemudian disimpan dalam variabel member.l, member.a, dan member.c untuk masing-masing metode clustering. Selanjutnya, data hasil klasterisasi tersebut dikombinasikan dengan kolom "Nama" dari dataset menggunakan fungsi data.frame, dan hasilnya disimpan dalam variabel clustersl, clustersa, dan clustersc. Setiap variabel tersebut mengandung dua kolom, yaitu "School" yang berisi informasi nama sekolah atau entitas, dan "Cluster" yang

menunjukkan klaster tempat entitas tersebut tergabung. Fungsi print kemudian digunakan untuk menampilkan hasil klasterisasi dalam format data frame.

```
# Menampilkan rata-rata nilai dalam setiap cluster
aggregate(z, by = list(cluster = member.l), mean)
aggregate(z, by = list(cluster = member.a), mean)
aggregate(z, by = list(cluster = member.c), mean)
```

Gambar 11 Langkah-langkah Agglomerative Hierarchical Clustering

Kode di atas bertujuan untuk menampilkan rata-rata nilai dalam setiap klaster yang dihasilkan dari tiga metode agglomerative hierarchical clustering sebelumnya (single linkage, average linkage, dan complete linkage). Dengan menggunakan fungsi aggregate, dataset yang telah dinormalisasi (z) diagregasi berdasarkan klaster yang telah ditentukan sebelumnya menggunakan variabel member.l, member.a, dan member.c. Dalam setiap baris hasil, kolom "Cluster" menyajikan nomor klaster, sementara kolom-kolom lainnya menunjukkan rata-rata nilai setiap variabel numerik dalam klaster tersebut.

BAB III

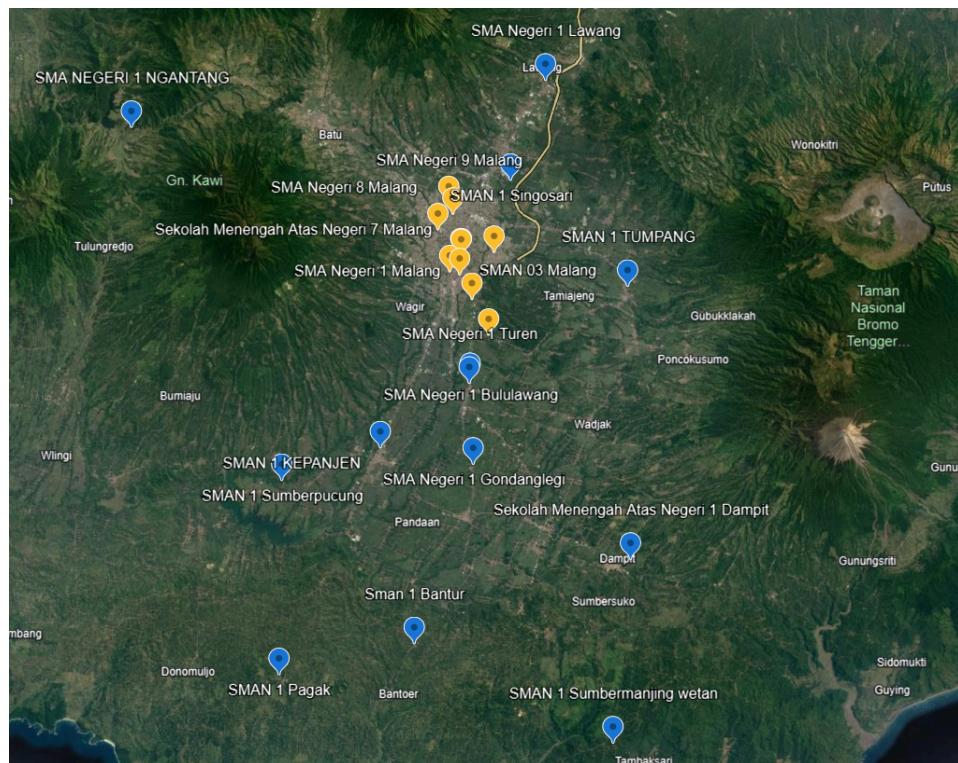
PEMBAHASAN

3.1 Pengolahan Data

Data yang kami gunakan dalam laporan ini diantaranya yaitu data seluruh SMA Negeri yang ada di Kota dan Kabupaten Malang (Sumber: *dapo.kemdikbud.go.id*), data wilayah, penduduk, dan jumlah SMA Negeri per kecamatan di Kota Malang tahun 2020-2022 dan Kabupaten Malang tahun 2015 - 2020 (Sumber: *malangkota.bps.go.id* dan *malangkab.bps.go.id*), data kantor kecamatan di Kota dan Kabupaten Malang (Sumber: *pemerintahan.malangkota.go.id* dan *malangkab.go.id*), serta data jarak antara SMA Negeri dan 4 kantor kecamatan dengan kepadatan penduduk tertinggi (Masing-masing diambil 2 kantor kecamatan dari Kota Malang dan Kabupaten Malang.

3.1.1 Data SMA Negeri di Kota dan Kabupaten Malang

Jumlah total SMA Negeri di Kota dan Kabupaten Malang adalah sebanyak 24 SMA, dengan rincian 11 SMA Negeri di Kota Malang (Ikon warna kuning) dan 13 SMA Negeri di Kabupaten Malang (Ikon warna biru). Berikut merupakan visualisasi lokasi dari ke-24 SMA Negeri tersebut dalam Google Earth:



Gambar 3.1 Visualisasi Lokasi SMA Negeri Kota dan Kabupaten Malang

3.1.2 Data Kecamatan dengan Kepadatan Penduduk Tertinggi di Kota dan Kabupaten Malang

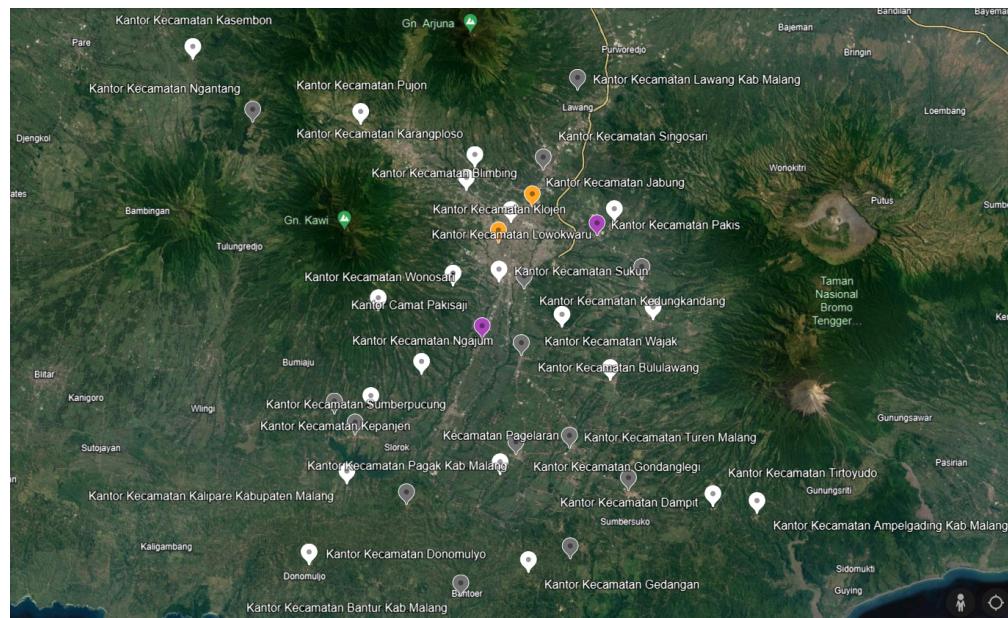
Jumlah total kecamatan di Kota dan Kabupaten Malang adalah sebanyak 38 kecamatan dengan rincian 5 kecamatan di Kota Malang dan 33 kecamatan di Kabupaten Malang. Di Kota Malang terdapat 3 dari 5 kecamatan yang memiliki SMA negeri, sedangkan untuk Kabupaten Malang hanya 13 dari 33 kecamatan yang memiliki SMA negeri. Kami memilih 4 kecamatan (masing-masing 2 kecamatan dari Kota Malang dan Kabupaten Malang) dengan kepadatan penduduk tertinggi (jumlah penduduk : luas wilayah). Dengan memilih kecamatan dengan penduduk tertinggi dapat dianggap sebagai representasi wilayah yang memiliki dampak signifikan terhadap distribusi dan zonasi SMA Negeri dimana kecamatan dengan penduduk tinggi cenderung mencerminkan pusat kegiatan dan kepadatan penduduk yang tinggi, dan dapat dianalisis bagaimana jumlah penduduk memengaruhi penempatan sekolah.

Tabel 7. Data Kantor Kecamatan dengan Kepadatan Penduduk Tertinggi
di Kota dan Kabupaten Malang

| No. | Kecamatan | Alamat | X (Latitude) | Y (Longitude) |
|-----|-----------|---|--------------|---------------|
| 1 | Blimbing | Jl. Raden Intan No.14, Polowijen, Kec. Blimbing, Kota Malang, Jawa Timur 65125 | -7.9303969 | 112.651949 |
| 2 | Klojen | Jl. Surabaya No.3, 65112, Sumbersari, Kec. Lowokwaru, Kota Malang, Jawa Timur 65115 | -7.9663394 | 112.6179851 |
| 3 | Pakis | Jl. Raya Pakis No.69, Krajan, Pakisjajar, Kec. Pakis, Kabupaten Malang, Jawa Timur 65154 | -7.9592009 | 112.7168389 |
| 4 | Pakisaji | Jl. Raya Pakisaji No.9, Jatirejo, Pakisaji, Kec. Pakisaji, Kabupaten Malang, Jawa Timur 65162 | -8.0610009 | 112.601334 |

3.1.3 Data Kantor Kecamatan di Kota dan Kabupaten Malang

Dari data sebelumnya, terdapat total 16 dari 38 kecamatan di Kota dan Kabupaten Malang yang tidak memiliki SMA negeri (Ikon warna putih), 22 dari 38 kecamatan yang memiliki setidaknya 1 SMA (Ikon warna abu-abu), 2 kantor kecamatan di Kota Malang terpadat (Ikon warna oranye), dan 2 kantor kecamatan di Kabupaten Malang terpadat (Ikon warna ungu) seperti digambarkan oleh visualisasi Google Earth berikut:

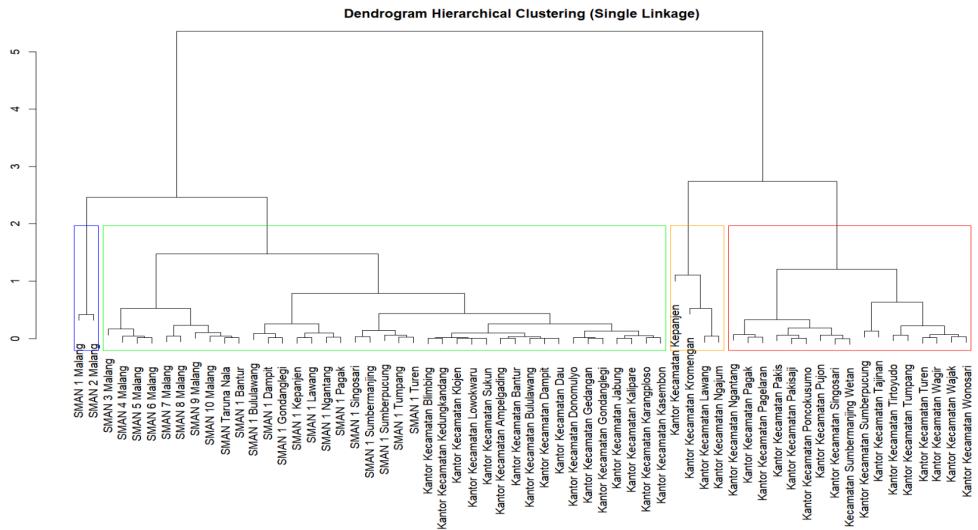


Gambar 3.2 Visualisasi Lokasi Kantor Kecamatan Kota dan Kabupaten Malang

3.2. Analisis Hasil Agglomerative Hierarchical Clustering

Setelah kami melakukan pengolahan data SMA negeri dan kantor kecamatan dengan menggunakan koordinat geografis pada software RStudio, kami menentukan 4 sebagai jumlah cluster yang kami amati. Terdapat 3 metode *linkage* yang kami gunakan sebagai penentu hasil clustering kami. Hasil dari pengolahan data yang telah dilakukan sebagai berikut:

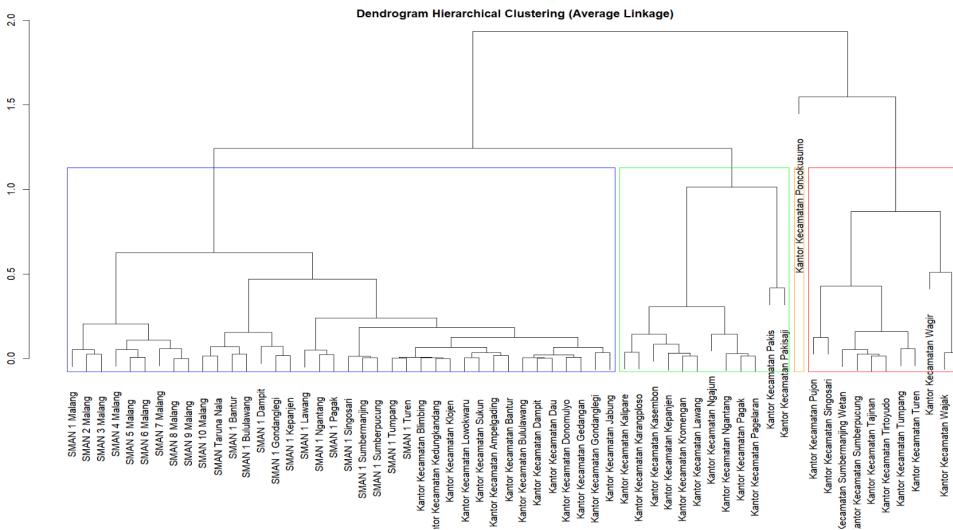
3.2.1 AHC Metode Single Linkage:



Gambar 3.3 Dendrogram Hierarchical Clustering (Single Linkage)

Single Linkage merupakan metode AHC yang mengukur jarak antara dua kelompok dengan mempertimbangkan jarak minimum antara satu elemen dari kelompok pertama ke satu elemen dari kelompok kedua. Ketika kami menggunakan metode ini terdapat 2 cluster besar dan 2 cluster yang merupakan outliers yaitu cluster biru dengan anggota SMAN 1 dan 2 Malang, serta cluster oranye dengan anggota Kantor Kecamatan Kepanjen, Kromengan, Lawang, dan Ngajum.

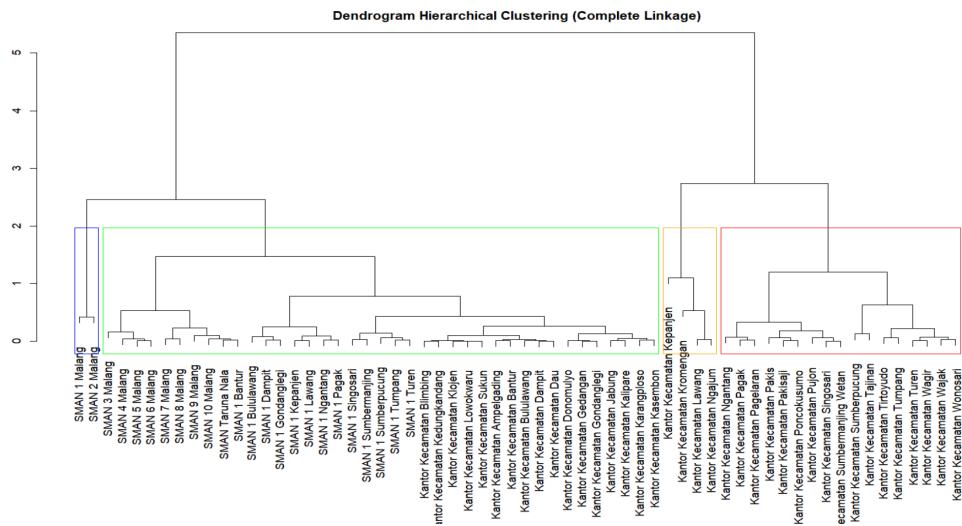
3.2.2 AHC Metode Average Linkage:



Gambar 3.4 Dendrogram Hierarchical Clustering (Average Linkage)

Average Linkage merupakan metode AHC yang mengukur jarak antara dua kelompok dengan menggunakan rata-rata jarak antara semua pasangan elemen yang satu berasal dari kelompok pertama dan yang lainnya dari kelompok kedua. Dengan menggunakan metode ini kita mendapatkan 3 cluster besar dan hanya 1 cluster yang merupakan outliers yaitu cluster oranye dengan anggota Kantor Kecamatan Kepanjen.

3.2.3 AHC Metode Complete Linkage:



Gambar 3.3 Dendrogram Hierarchical Clustering (Complete Linkage)

Complete Linkage merupakan metode AHC yang mengukur jarak antara dua kelompok dengan mempertimbangkan jarak maksimum antara satu elemen dari kelompok pertama ke satu elemen dari kelompok kedua. Cluster yang terbentuk dengan menggunakan metode ini sama dengan ketika kami menggunakan metode Single Linkage.

Setelah dilakukan pengamatan, terdapat 2 jenis clustering yang terbentuk yaitu hasil dengan tendensi untuk membentuk kelompok yang kompak dan padat (Metode Single dan Complete Linkage) dan hasil yang lebih cenderung menghasilkan kelompok yang lebih seimbang dan seragam dalam ukuran (Metode Average Linkage). Dari hasil pengelompokan SMA Negeri dan Kantor Kecamatan di daerah Malang, diketahui bahwa persebaran SMA di daerah Kota dan Kabupaten Malang kurang merata. Ditandai dengan adanya outlier yaitu SMAN 1 dan SMAN 2 Malang yang berarti kebanyakan SMA Negeri di Malang kebanyakan terdapat di daerah pusat Kota Malang dengan kepadatan penduduk yang tinggi.

BAB IV

PENUTUP

4.1. Simpulan

Dari hasil analisis Agglomerative Hierarchical Clustering kami menyimpulkan bahwa distribusi SMA Negeri di Kota dan Kabupaten Malang kurang merata, wilayah yang memiliki kepadatan penduduk tinggi, khususnya di Kota Malang, memang benar sudah terdapat SMA Negeri pada daerah itu. Hal tersebut dapat dilihat dari hasil analisis clustering yang menunjukkan terbentuknya cluster-cluster yang didominasi SMA negeri di wilayah perkotaan dengan kepadatan penduduk yang lebih tinggi. Namun, pada daerah Kabupaten Malang tiap kecamatannya rata-rata hanya memiliki 1 SMA negeri atau bahkan tidak ada SMA negeri sama sekali. Sehingga dapat dikatakan persebaran SMA negeri-nya hanya fokus pada daerah perkotaan saja.

4.2. Saran

Berdasarkan penelitian yang telah dilakukan, dapat disarankan untuk melakukan pengaturan ulang sistem zonasi SMA Negeri di Kota dan Kabupaten Malang dengan mempertimbangkan faktor distribusi SMA negeri dan kepadatan penduduk. Dengan demikian, diharapkan dapat meningkatkan pemerataan akses pendidikan dan keadilan bagi siswa di seluruh wilayah Kota dan Kabupaten Malang. Penelitian lanjutan diperlukan untuk menggali faktor-faktor lain yang mempengaruhi zonasi, seperti ketersediaan infrastruktur pendidikan, tingkat kelulusan sekolah, dan preferensi siswa agar menghasilkan rekomendasi zonasi yang lebih optimal dan sesuai dengan kebutuhan masyarakat.

LAMPIRAN

Visualisasi Google Earth:

<https://earth.google.com/earth/d/1ELULkwOZEQ4vCrAGb4QQPa02qnsUQ2ds?usp=sharing>

Data Final Project ADD Kelas D (Kota dan Kabupaten Malang):

https://docs.google.com/spreadsheets/d/1tUYehwB0JgYoKy1STx_4PHy-zDBIupGmJfVwxGyiySY/edit?usp=sharing

Pembagian Tugas:

- Farhan Bilawa Gita Maulana (5026211158):
 - Mengerjakan Sub Bab 2.1
 - Mengerjakan Sub Bab 2.4
 - Mengerjakan Sub Bab 3.2
- Alif Faturrohman (5026221040):
 - Mengerjakan Sub Bab 1.1
 - Mengerjakan Sub Bab 3.1
 - Mengolah Data
- Dewi Maharani (5026221046):
 - Mengerjakan Sub Bab 2.2
 - Mengerjakan Sub Bab 2.3
 - Mengerjakan Sub Bab 2.5
- Muhammad Iqbal Baiduri Yamani (5026221103):
 - Mengerjakan Sub Bab 2.2
 - Mengerjakan Sub Bab 3.1
 - Mengerjakan Sub Bab 3.2
 - Mengolah Data
- Keysha Amelia (5026221122):
 - Mengerjakan Sub Bab 1.2
 - Mengerjakan Sub Bab 1.3
 - Mengerjakan Sub Bab 2.1
 - Mengerjakan Sub Bab 2.4
 - Mengerjakan Bab 4

DAFTAR PUSTAKA

- [1] H. Hikmah, F. Fardinah, L. Qadrini, and E. Tande, “Analisis Klaster Pengelompokan Kecamatan di Sulawesi Barat Berdasarkan Indikator Pendidikan,” SAINTIFIK, vol. 8, no. 2, pp. 188–196, Jul. 2022, doi: 10.31605/SAINTIFIK.V8I2.383
- [2] K. P. Simanjuntak and U. Khaira, “Pengelompokan Titik Api di Provinsi Jambi dengan Algoritma Agglomerative Hierarchical Clustering,” MALCOM Indones. J. Mach. Learn. Comput. Sci., vol. 1, no. 1, pp. 7–16, 2021, doi: 10.57152/malcom.v1i1.6
- [3] Iglewicz, Boris & Banerjee, Sharmila. (2001). A simple univariate outlier identification procedure.