

Pemetaan Kejadian Stunting Berdasarkan Status Ekonomi Keluarga Di Puskesmas Bangil Menggunakan Algoritma X-Means

Oong Alfa Salasa¹, Muhammad Faisal²

¹SMK 1 Pasuruan, ²Magister Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri
Maulana Malik Ibrahim Malang

¹2006015210014@student.uin-malang.ac.id, ²mfaisal@ti.uin-malang.ac.id

ABSTRAK

Stunting adalah masalah kekurangan asupan gizi dari makanan yang berlangsung cukup lama, upaya penurunan kejadian stunting merupakan salah satu prioritas program pemerintah dalam bidang kesehatan. Pemetaan wilayah berdasarkan kejadian stunting akan sangat membantu petugas kesehatan pada wilayah kerja untuk menurunkan prevalensi stunting. Dalam penelitian ini, penulis akan memetakan wilayah yang rentan terjadi stunting berdasarkan faktor status ekonomi keluarga dengan teknik *data mining Clustering* menggunakan algoritma X-Means. Untuk mengetahui *scoring cluster* yang optimal menggunakan *Davies Bouldin Index (DBI)*. Agar memudahkan pembacaan hasil analisa *clustering* diintegrasikan dalam sebuah peta berbasis Sistem Informasi Geografis (SIG). Hasil penelitian memberikan gambaran mengenai pengelompokan wilayah dengan nilai *scoring* evaluasi DBI sebesar 0,47. Disarankan untuk memanfaatkan metode yang berbeda dan membandingkannya agar kevaliditasnya menjadi lebih maksimal.

Kata kunci: Stunting, Clustering, X-Means, Sistem Informasi Geografis (SIG)

ABSTRACT

Stunting is a problem of lack of nutritional intake from food that lasts quite a long time. Efforts to reduce the incidence of stunting are one of the priorities of government programs in the health sector. Regional mapping based on the incidence of stunting will greatly assist health workers in their work areas to reduce the prevalence of stunting. In this study, the authors will map areas that are prone to stunting based on family economic status factors using the Clustering data mining technique using the X-Means algorithm. To determine the optimal scoring cluster using the Davies Bouldin Index (DBI). In order to make it easier to read the results of the clustering analysis, it is integrated into a map based on a Geographic Information System (GIS). The results of the study provide an overview of the grouping of regions with a DBI evaluation scoring value of 0.47. It is recommended to use different methods and compare them so that their validity is maximized.

Keywords: Stunting, Clustering, X-Means, Geographic Information System (GIS)

1. PENDAHULUAN

Melalui halaman resmi UNICEF *Global Nutrition Report* mengeluarkan informasi tingkat malnutrisi anak tahun 2021. pernyataan itu menguak keadaan jutaan kanak-kanak yang obesitas, kurang nutrisi, sampai stunting di bumi. Diperkirakan terdapat 149, 2 juta kanak-kanak yang alami *stunting*. Nilai itu sebanding 22 persen balita di semua negara pada 2020 [1]. Masih banyak ditemukan balita *stunting* di Kabupaten Pasuruan, khususnya di wilayah kerja Puskesmas Bangil selama periode tahun 2022 yaitu sebanyak 132 kasus.

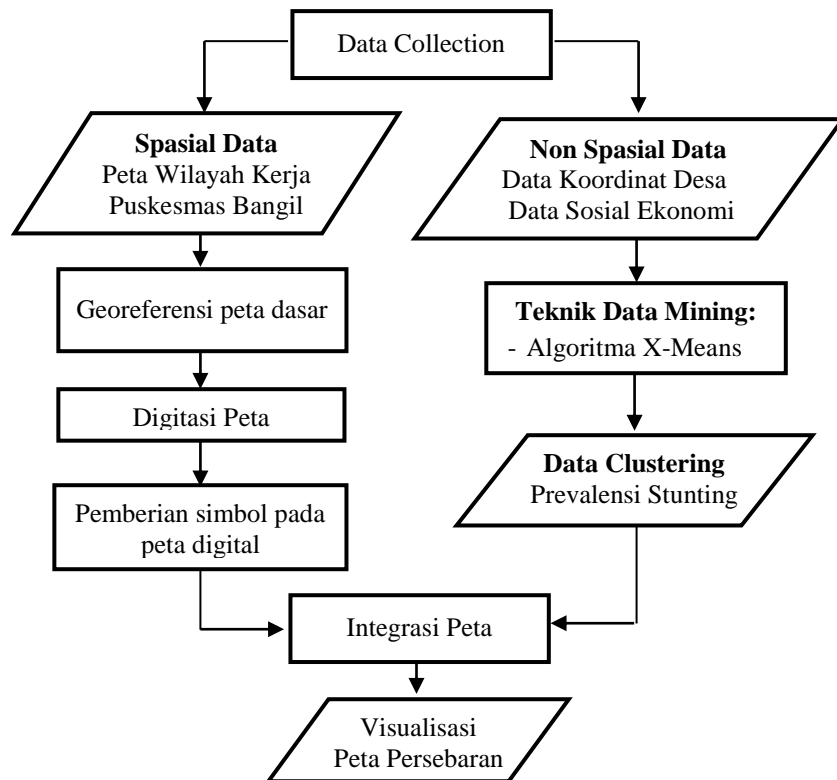
Dari data kejadian *stunting* tiap desa di seluruh wilayah Puskesmas Bangil Kabupaten Pasuruan, melalui metode *clustering* menggunakan algoritma X-Means akan mengelompokkan desa-desa berdasarkan kesamaan pola atau kesamaan kelompoknya.

Algoritma X-Means merupakan hasil perbaikan dari beberapa kekuangan dari algoritma K-Means, melalui perulangan pengelompokan beberapa kali sehingga menemukan titik terbaik dari nilai yang telah ditentukan. Menggunakan angka *Bayesian Information Criterion* (BIC) sebagai indikator untuk melakukan pemecahan kelompok data. [2]. Hasil pengelompokan dipetakan menggunakan program Sistem Informasi Geografis (SIG), dalam menghasilkan gambar denah suatu wilayah ataupun area tertutup yang bersumber pada informasi spasial mencakup posisi, objek serta keterikatan ruang di alam yang dikenal dengan peta. [3].

Dengan pengelompokan desa di wilayah Puskesmas Bangil berdasarkan faktor sosial ekonomi keluarga sebagai penyebab *stunting* dalam visualisasi peta persebaran, akan menjadi referensi untuk menetapkan wilayah yang menjadi prioritas dalam penanganan *stunting*.

2. METODOLOGI PENELITIAN

Tahapan dalam penelitian ini diawali mengumpulkan data *collection* yang terdiri dari data peta wilayah kerja Puskesmas Bangil (Spasial Data), kemudian diolah menggunakan SIG menjadi peta. Kemudian untuk data *collection* non spasial data sosial ekonomi dari balita *stunting* dikelompokkan menggunakan algoritma X-Means menjadi Data *clustering* Prevalensi *Stunting*. Dari hasil spasial data berupa peta dan non spasial data berupa data *clustering* diintegrasikan menghasilkan pemetaan persebaran *stunting* di wilayah Puskemas Bangil. Proses langkah-langkah dalam penelitian ditunjukkan oleh Gambar 1 berikut ini:



Gambar 1: Flowchart Desain Penelitian

2.1 Pengumpulan Data (*Data Collection*)

Pengumpulan informasi yang berkaitan dengan teori diambil dari artikel, literatur dan jurnal penelitian serta sumber-sumber lain berhubungan dengan permasalahan yang dibahas. Sebaliknya informasi yang hendak dipergunakan dalam riset ini adalah informasi inferior.

Untuk sumber data sekunder yang digunakan di bagi menjadi 2 data yaitu Data Spasial dan Data Non Spasial:

2.1.1 Data Spasial

Data spasial merupakan fakta yang dapat membuktikan posisi informasi di dataran alam. Informasi spasial mempunyai rujukan posisi pemetaan yang di formatkan dalam bentuk koordinat. Informasi spasial kerap pula di katakan sebagai informasi geospasial, informasi geografis, ataupun geodata [4].

1. Georeferensi Peta Dasar

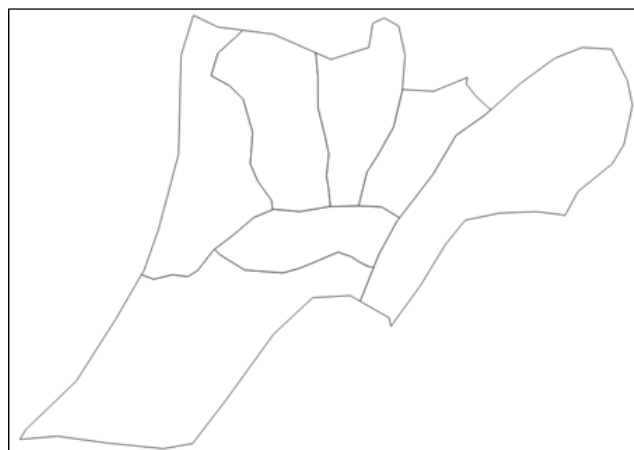
Data spasial mempunyai rujukan letak posisi pemetaan dan digambarkan dalam suatu titik koordinat UTM (*Universal Transerve Mercator*), dengan data koordinat sebagai berikut:

Tabel 1 Titik Koordinat Desa Wilayah Kerja Puskesmas Bangil

No.	Desa	Titik Koordinat	
		X	Y
1	Bendo Mungal	112.79202	-7.59892
2	Gempeng	112.80353	-7.59960
3	Kauman	112.78560	-7.60354
4	Kersikan	112.79507	-7.60366
5	Kidul Dalem	112.78954	-7.60918
6	Kolursari	112.77770	-7.61584
7	Pogar	112.78063	-7.60558

2. Digitasi Peta

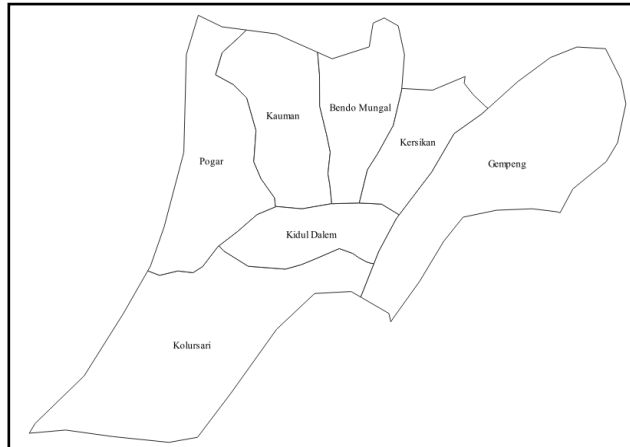
Pembuatan peta dilakukan dengan digitasi (penggambaran peta dilakukan secara on-screen pada layer monitor) yang akan menjadi sebuah peta digital.



Gambar 2. Peta Digital Batas Administrasi Desa

3. Pemberian model simbol pada peta digital

Melalui pemberian simbol atau *labelling* berupa nama desa dapat memberikan keterangan yang bermanfaat bagi pengguna.



Gambar 3. Peta Batas Administrasi Desa

2.1.2 Data Non Spasial

Data sekunder ini berupa data yang di peroleh dari Puskesmas Bangil Kesehatan tentang faktor sosial ekonomi keluarga dari anak yang mengalami *stunting* tahun 2022 sebanyak 246 anak. Adapun dataset faktor sosial ekonomi keluarga penyebab terbanyak *stunting* digunakan variabel dalam penelitian yang di tampilkan pada tabel berikut:

Tabel 2. Dataset Faktor Sosial Ekonomi Keluarga

No.	Desa	Faktor Sosial Ekonomi Keluarga				Jumlah
		Tingkat pendidikan ibu tinggi	Kategori pendapatan keluarga tinggi	Status Pekerjaan Orang Tua Rendah	Rumah yang memiliki sanitasi layak	
1.	Bendomunggal	12	15	15	6	48
2.	Gempeng	13	18	6	6	43
3.	Kauman	13	19	5	10	47
4.	Kersikan	2	3	2	2	9
5.	Kidul Dalem	17	24	19	9	69
6.	Kolursari	5	7	6	5	23
7.	Pogar	2	2	2	1	7
Jumlah Total						246

2.2 Teknik Data Mining

Dalam penelitian ini proses pengelompokan data, teknik *data mining* yang digunakan algoritma X-Means. Pengembangan perbaikan Algoritma yang dilakukan Dan Pelleg dan Andre Moore pada tahun 2000[6]. Langkah penyelesaian untuk menentukan kelompok diproses secara terus menerus memakai nilai maksimal serta minimal sajian oleh pemakai. Prosedur penyelesaian ini meliputi 2 tahap hingga berakhir.

1. Menaikkan Nilai Paramater, menjalankan algoritam k-means sampai tidak ada perubahan yang terjadi dalam pengelompokan. Menggunakan nilai k sebagai batasan bawah yang sudah ditentukan diawal oleh pemakai.
2. Memperbaiki susunan kelompok, dalam tahap ini memecah kelompok berdasarkan titik pusat menjadi dua bagian dengan seleksi random. Untuk menentukan nilai pusat *cluster* dengan menyamakan nilai BIC.
3. Bila $K = k_{\max}$ (batasan atas) menyudahi serta menginformasikan ke bentuk evaluasi terbaik yang ditemui sepanjang pencarian, bila tidak berangkat ke Tahap 1.[7]

Untuk mengetahui hasil evaluasi clustering seberapa baik dalam pengelompokan data, maka dipergunakan prosedur *Davies Bouldin Index* (DBI). Prosedur DBI mengukur penilaian *cluster* pada sesuatu prosedur pembentukan kelompok, keterikatan data dikonfirmasi selaku jumlah berdasarkan keefektan informasi pada nilai pusat kelompok dari *cluster* yang disertai, sebaliknya separasi ditentukan pada jarak dampingi nilai pusat kelompok dari *cluster*. Angka perbandingan yang didapat itu dikenakan guna mencari angka DBI dengan persesuaian berikut:

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \dots\dots\dots [8]$$

Berdasarkan rumus DBI, sejumlah kelompok diidentifikasi menggunakan nilai k. Perolehan hasil perhitungan mendekati nilai 1 (non-negatif ≥ 0), maka pengelompokan sangat ideal.

2.3 Integrasi Peta

Dalam melakukan penggabungan peta dengan data klaster dibutuhkan aplikasi Bahasa pemrograman Phyton yang menggunakan *library* geopandas. Peta digitasi yang sudah dibuat dengan file ekstensi .shp diintegrasikan dengan data klaster dimana setiap klaster data sudah ditentukan warna background.

2.4 Visualisasi Peta

Visualisasi peta digunakan untuk menganalisis dan menampilkan data geografis dan menyajikannya dalam bentuk tipe peta yang berbeda. Informasi yang lebih mungkin diungkapkan melalui visualisasi menjadi lebih jelas dan lebih intuitif. Dapat melihat distribusi atau proporsi data di setiap area, sehingga memudahkan orang untuk terlibat dengan informasi dan membuat keputusan yang lebih baik[9].

3. HASIL DAN PEMBAHASAN

Berdasarkan penjelasan pada bab sebelumnya, proses penelitian mulai dari pembacaan data sampai proses visualisasi peta persebaran menggunakan aplikasi jupiter notebook dengan menggunakan bahasa pemrograman Python.

Dalam penyelesaian masalah menggunakan bahasa pemrograman python dibutuhkan beberapa *library* pendukung[10]. *Library* bantuan yang dipakai antara lain:

- a. *PyClustering* merupakan *library* untuk *machine learning* digunakan untuk proses pengelompokan data melalui proses pembobotan sebelumnya
- b. *Scikit-learn (sklearn)* termasuk *library* python, dalam penelitian ini digunakan dalam perhitungan *score* dari proses pembentukan kelompok dengan metode *DBI*.
- c. *Pandas* digunakan dalam mengerjakan sekumpulan *dataset*.
- d. *Matplotlib* digunakan untuk melakukan visualisasi data seperti membuat plot grafik untuk satu sumbu atau lebih.
- e. *Geopandas* digunakan untuk membuat visualisasi data Geospasial.

3.1. Pembacaan data

Proses pertama dalam melakukan *machine learning* yaitu membaca dan menampilkan *dataset* faktor sosial ekonomi keluarga pada balita *stunting* yang sudah tersimpan dengan nama SosialEkonomi.csv

```
# Membaca dataset Sosial Ekonomi
dt_ekonomi = pd.read_csv('SosialEkonomi.csv')
print(dt_ekonomi)
```

	Desa	X1	X2	X3	X4
0	BENDO MUNGAL	12	15	15	6
1	GEMPENG	13	18	6	6
2	KAUMAN	13	19	5	10
3	KERSIKAN	2	3	2	2
4	KIDUL DALEM	17	24	19	9
5	KOLURSARI	5	7	6	5
6	POGAR	2	2	2	1

Gambar 4. Perintah membaca dan menampilkan data set

Hasil tampilan pembacaan data SosialEkonomi.csv, terlihat data jumlah balita *stunting* setiap desa yang dipengaruhi oleh faktor sosial ekonomi dipakai sebagai variabel penelitian. Untuk memudahkan dan efisiensi penulisan variabel dituliskan dalam bentuk kode. Keterangan kode variabel di jelaskan pada tabel berikut:

Tabel 2. Kode Variabel Faktor Sosial Ekonomi Keluarga

Kode Variabel	Keterangan Faktor Sosial Ekonomi Keluarga
X1	Tingkat pendidikan ibu tinggi
X2	Kategori pendapatan keluarga tinggi
X3	Status Pekerjaan Orang Tua Rendah
X4	Rumah tangga yang memiliki sanitasi layak.

3.2. Menentukan Inisialisasi Center dari kluster

Inisialisasi jumlah kluster dimulai dengan kluster terkecil ($C = 2$) dengan titik pusat kelompok 1 dan kelompok 2 diambil secara acak menggunakan metode *means_plusplus_initializer* yang sudah disediakan oleh *library python PyClustering*.

```
# Prepare initial centers - amount of initial centers defines amount of clusters from which X-Means will
# start analysis.
random_seed = 1000
amount_initial_centers = 2
sample = dt_ekonomi_x
initial_centers = kmeans_plusplus_initializer(sample, amount_initial_centers).initialize()
print(initial_centers)
```

Gambar 5. Perintah inisialisasi pusat kluster

Dihasilkan pusat kluster awal bisa diamati pada bagan 5 di bawah ini:

Tabel 4. Inisialisasi Pusat Kluster Awal

Desa	X1	X2	X3	X4
Kidul Dalem	17	24	19	9
Kolursari	5	7	6	5

3.3. Teknik Data Mining

Proses teknik data mining *clustering* menggunakan metode *X-Means* dengan kelebihan memudahkan dalam menentukan jumlah *cluster* terbaik pada data yang besar.

```
# Buat instance dari algoritma X-Means. Algoritme akan memulai analisis dari minimal 2 cluster
# jumlah cluster maksimal yang dapat dialokasikan adalah 20.
xmeans_instance = xmeans(sample, xm_c2, 20, splitting_type=splitting_type.BAYESIAN_INFORMATION_CRITERION,
                          random_state=random_seed)
xmeans_instance.process()
```

Gambar 5. Perintah *clustering* metode *X-Means*

Proses kerja dari metode *X-Means* adalah data dikelompokkan menggunakan metode *Bayesian Information Criterion (BIC)*, untuk mencari pengelompokan data secara optimal berdasarkan inisialisasi pusat kluster yang sudah ditentukan secara acak.

3.4. Data Clustering

Hasil proses *clustering* menggunakan *X-Means* menggunakan perintah python seperti berikut.

```
# Ekstrak hasil pengelompokan: cluster dan pusatnya
clusters = xmeans_instance.get_clusters()
centers = xmeans_instance.get_centers()
print(clusters)

[[1, 2], [5], [3, 6], [0, 4]]

print(centers)

[[13.0, 18.5, 5.5, 8.0], [5.0, 7.0, 6.0, 5.0], [2.0, 2.5, 2.0, 1.5], [14.5, 19.5, 17.0, 7.5]]
```

Gambar 6. Perintah menampilkan hasil *clustering* dan Nilai Pusat *cluster*

Hasil proses *clustering* menggunakan cara pengelompokan *Bayesian Information Criterion (BIC)* menghasilkan 4 kelompok data.

```
#Hasil Clustering
dt_ekonomi['Klaster'] = encoder.get_clusters()
print(dt_ekonomi)
```

	Desa	X1	X2	X3	X4	Klaster
0	BENDO MUNGAL	12	15	15	6	3
1	GEMPENG	13	18	6	6	0
2	KAUMAN	13	19	5	10	0
3	KERSIKAN	2	3	2	2	2
4	KIDUL DALEM	17	24	19	9	3
5	KOLURSARI	5	7	6	5	1
6	POGAR	2	2	2	1	2

Gambar 7. Perintah menampilkan hasil kluster

3.5. Hasil Evaluasi Clustering

Berdasarkan perhitungan proses clustering yang dihasilkan, maka untuk mengetahui evaluasi internal cluster, dimana baik atau tidaknya hasil pengelompokan *Davies-Bouldin Index (DBI)* merupakan metode menghitung validitas kelompok pada suatu metode pengelompokan[11].

```
# Ekstrak hasil pengelompokan BIC X-Means.
BIC_clusters = xmeans_instance.get_clusters()

# Secara default x-means mengembalikan representasi CLUSTER_INDEX_LIST_SEPARATION
type_repr = xmeans_instance.get_cluster_encoding()
encoder = cluster_encoder(type_repr, BIC_clusters, dt_ekonomi_x)

# Perhitungan Davies-Bouldin Index (DBI)
db = davies_bouldin_score(dt_ekonomi_x, encoder.get_clusters())
print(db)

0.4711128653718212
```

Gambar 8. Perintah Perhitungan Perhitungan Davies-Bouldin Index (DBI)

Bersumber pada perhitungan skoring DBI, semakin rendah nilai DBI (tidak negatif) ≥ 0 maka *cluster* yang didapat semakin berkualitas dengan menggunakan pengelompokan *x-means*.

3.6. Integrasi Peta

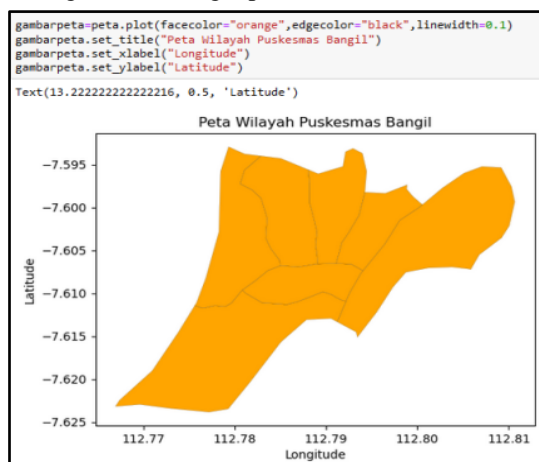
Data georeferensi peta dasar wilayah Puskesmas Bangil di tampilkan menggunakan *geopandas* dan ditampilkan dalam bentuk peta plot.

```
peta = gpd.read_file('C:/Users/smk/Documents/Python/SosialEkonomi - 12-1/Bangil/SHP_Kecamatan Bangil/Desa/Indo_Desa_region.shp')
```

	DESA	KODE	KECAMATAN	KABUPATEN	PROPINSI	LUAS_WILAY	geometry
0	Kolursari	None	Bangil	Pasuruan	Jawa Timur	1830	POLYGON Z ((112.76695 -7.62317 0.00000, 112.76...
1	Pogar	None	Bangil	Pasuruan	Jawa Timur	1874	POLYGON Z ((112.77564 -7.61143 0.00000, 112.77...
2	Kauman	None	Bangil	Pasuruan	Jawa Timur	540	POLYGON Z ((112.78499 -7.60681 0.00000, 112.78...
3	Kidul Dalem	None	Bangil	Pasuruan	Jawa Timur	494	POLYGON Z ((112.78085 -7.60963 0.00000, 112.78...
4	Bendo Mungal	None	Bangil	Pasuruan	Jawa Timur	500	POLYGON Z ((112.78911 -7.60658 0.00000, 112.78...
5	Kersikan	None	Bangil	Pasuruan	Jawa Timur	442	POLYGON Z ((112.79113 -7.60653 0.00000, 112.79...
6	Gempeng	None	Bangil	Pasuruan	Jawa Timur	1292	POLYGON Z ((112.79405 -7.60738 0.00000, 112.79...

Gambar 9. Perintah menampilkan data geometry wilayah kerja Puskesmas Bangil

Perintah menampilkan data geometry wilayah kerja Puskesmas dari data georeferensi peta dasar menggunakan *geopandas*, digunakan sebagai penentuan titik koordinat dalam pemetaan.



Gambar 10. Perintah menampilkan peta wilayah kerja Puskesmas Bangil

Peta dihasilkan dari koordinat wilayah yang di rendering dalam bentuk plot menggunakan library *matplotlib.pyplot*.

3.7. Visualisasi Peta

Peta plot hasil integrasi peta di kombinasikan dengan data kluster dimana setiap kluster sudah ditentukan warna background.

```
# Menentukan Keterangan Kluster
dt_ekonomii = pd.read_csv('SosialEkonomiii.csv')
Kondisi = [
    (dt_ekonomii['Kluster']==0),
    (dt_ekonomii['Kluster']==1),
    (dt_ekonomii['Kluster']==2),
    (dt_ekonomii['Kluster']==3)]
Pilihan = ['Rendah','Sedang','Tinggi','Sangat Tinggi']
dt_ekonomii['Keterangan'] = np.select(Kondisi, Pilihan)
print(dt_ekonomii)
```

	Desa	Kluster	Keterangan
0	Bendo Mungal	3	Sangat Tinggi
1	Gempeng	0	Rendah
2	Kauman	0	Rendah
3	Kersikan	2	Tinggi
4	Kidul Dalem	3	Sangat Tinggi
5	Kolursari	1	Sedang
6	Pogar	2	Tinggi

Gambar 11. Perintah menampilkan keterangan kluster

Berdasarkan ketentuan pemerintah Jawa Timur tentang status wilayah *stunting*, terdapat 4 kategori yaitu status wilayah *stunting* Rendah (*Background* Warna Biru), Sedang (*Background* Warna Hijau), Tinggi (*Background* Warna Kuning) dan Sangat Tinggi (*Background* Warna Merah)

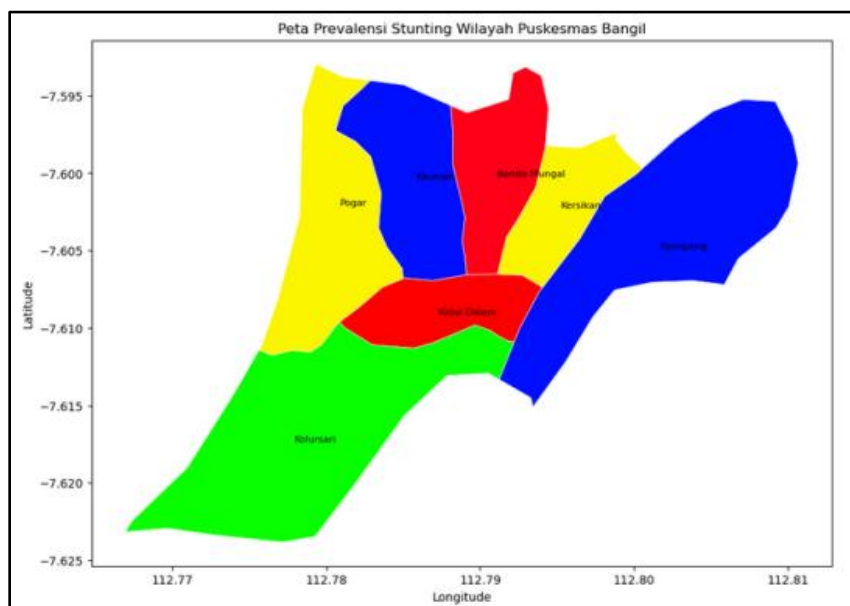
```
dt_ekonomii = pd.read_csv('SosialEkonomii.csv')
gabungan=pd.merge(
    left=dt_ekonomii,
    right=peta,
    how="inner",
    left_on=["Desa"],
    right_on=["DESA"]
)
gabunganGDF=gpd.GeoDataFrame(gabungan)
gambarpeta=gabunganGDF.plot(
    column="kluster",
    cmap="hsv",
    figsize=(30, 8),
    edgecolor="white",
    linewidth=0.5,
)
gambarpeta.set_xlabel("Longitude")
gambarpeta.set_ylabel("Latitude")
gambarpeta.set_title("Peta Prevalensi Stunting Wilayah Puskesmas Bangil")

gabunganGDF["lokasilabel"]=gabunganGDF.geometry.representative_point()

labelpeta=[]
for namakec,posisi in zip(gabunganGDF["DESA"],gabunganGDF["lokasilabel"]):
    labelpeta.append(gambarpeta.text(posisi.x,posisi.y,namakec,fontsize=8))
```

Gambar 12. Perintah menampilkan integrasi peta dengan kluster

Perintah integrasi melalui tiga proses ketika menampilkan dalam bentuk peta. Proses pertama menggabungkan data kluster ke dalam data geometri, langkah ke dua peta sesuai dengan data kluster dan proses ke tiga memberi label nama wilayah dalam tampilan peta.



Gambar 13. Peta Persebaran Prevalensi *Stunting*

Berdasarkan hasil *clustering* menggunakan algoritma X-Means terbentuk 4 kelompok, ketika di integrasikan kedalam peta menghasilkan peta persebaran prevalensi *stunting* di wilayah Puskesmas bangil seperti yang terlihat pada gambar 13.

Pada peta persebaran prevalensi stunting terlihat status *stunting* desa berdasarkan *background* warna, dengan keterangan sebagai berikut:

- *Background* warna merah
Legenda peta dengan warna *background* merah merupakan wilayah desa dengan status **stunting sangat tinggi**, yaitu desa Bendo Mungal dan desa Kidul Dalem.
- *Background* warna biru
Legenda peta dengan warna *background* biru merupakan wilayah desa dengan status **stunting tinggi**, yaitu desa Kauman dan desa Gempeng.
- *Background* warna hijau
Legenda peta dengan warna *background* kuning merupakan wilayah desa dengan status **stunting sedang**, yaitu desa Kolursari.
- *Background* warna kuning
Legenda peta dengan warna *background* kuning merupakan wilayah desa dengan status **stunting rendah**, yaitu desa Pogar dan kersikan.

4. KESIMPULAN

Hasil penelitian menghasilkan beberapa kesimpulan sebagai berikut:

1. Pengelompokan wilayah desa kejadian *stunting*, berdasarkan data status ekonomi keluarga di wilayah Puskesmas Bangil menggunakan algoritma X-Means menghasilkan 4 kelompok terbaik. Dengan hasil keakuratan cluster sebesar 0.47 menggunakan DBI, dimana nilai yang mempunyai tingkatan ketepatan yang bagus dengan angka DBI mendekati 0.

2. Melalui proses menggunakan *library* geopandas metode plot pada python tidak terjadi kesalahan pada proses integrasi peta dengan hasil pengelompokan data. Serta didapatkan hasil kluster data sesuai dengan *background* warna yang telah ditentukan.
3. visualisasi peta persebaran prevalensi *stunting* dapat digambarkan status desa dengan *background* warna merah untuk status kejadian *stunting* sangat tinggi, *background* warna biru untuk status kejadian *stunting* tinggi, *background* warna hijau untuk status kejadian *stunting* sedang dan *background* warna kuning untuk status kejadian *stunting* rendah.

DAFTAR PUSTAKA

- [1] Anggryni, M. et al. (2021) 'Faktor Pemberian Nutrisi Masa Golden Age dengan Kejadian Stunting pada Balita di Negara Berkembang', *Jurnal Obsesi : Jurnal Pendidikan Anak Usia Dini*, 5(2), pp. 1764–1776. doi: 10.31004/obsesi.v5i2.967.
- [2] R. Adhitama, A. Burhanuddin, and R. Ananda, "Penentuan Jumlah Cluster Ideal Smk Di Jawa Tengah Dengan Metode X-Means Clustering Dan K-Means Clusterin," *JIKO (Jurnal Inform. dan Komputer)*, vol. 3, no. 1, pp. 1–5, 2020, doi: 10.33387/jiko.v3i1.1635.
- [3] Purwanto, A. and Iswandi, I. (2019) 'Pemanfaatan Sistem Informasi Geografis Untuk Menentukan Lokasi Potensial Pengembangan Kawasan Industri Di Kabupaten Pati', *Jurnal Tanah dan Sumberdaya Lahan*, 6(2), pp. 1219–1228. doi: 10.21776/ub.jtsl.2019.006.2.2.
- [4] Dahlia, S. et al. (2021) 'Analisis Pola Spasial Pesebaran Kasus Covid-19 Menggunakan Sistem Informasi Geografis Di DKI Jakarta', 5(2), pp. 101–108.
- [5] Sanam, S. M., Manurung, I. and Purnawan, S. (2021) 'Pemetaan Kejadian Stunting Di Wilayah Kerja Puskesmas Bati Berdasarkan Ketersediaan Air Bersih', *Media Kesehatan Masyarakat*, 3(2), pp. 119–127.
- [6] D. Pelleg & A. Moore. 2014. X-means: Extending K-means with Efficient Estimation of The Number of Clusters. *Proceedings of the Seventeenth International Conference on Machine Learning*, Palo Alto, CA.
- [7] F. Noorbehbahani and S. Mansoori, "A new semi-supervised method for network traffic classification based on X-means clustering and label propagation," 2018 8th Int. Conf. Comput.Knowl. Eng. ICCKE 2018, no. Iccke, pp. 120–125, 2018, doi:10.1109/ICCKE.2018.8566608.
- [8] Wahidin, A. J. and Sensuse, D. I. (2021) 'Perbandingan Algoritma K-Means, X-Means Dan K-Medoids Untuk Klasterisasi Awak Kabin Lion Air', *Jurnal ICT: Information Communication & Technology*, 20(2), pp. 298–302.
- [9] Danila, Pawa ID, Choiruni A, Wijayanti A. Geospatial Analysis pada Prevalensi Stunting di Kabupaten Manggarai. *Ber Kedokt Masy* [Internet]. 2018;34(11). Available from: <https://journal.ugm.ac.id/bkm/article/view/40618>
- [10] Edwardo, T. O. (2018). *Penggunaan Python untuk Data Mining*. Binus University School of Computer Science.
- [11] PyShark. (2021). Davies-Bouldin Index for K-Means Clustering Evaluation in Python. *PythonBloggers.Com*.<https://pythonbloggers.com/2021/06/davies-bouldin-index-fork-means-clustering-evaluation-in-python/>