

PENGEMBANGAN *DASHBOARD SELF-SERVICE ANALYTICS* UNTUK ANALISIS *SOCIAL VULNERABILITY INDEX (SOVI)* DI INDONESIA

Dashboard ini dikembangkan sebagai bentuk pemenuhan tugas
Ujian Akhir Semester pada mata kuliah Komputasi Statistik



Disusun Oleh:

Amanda Tri Hapsari
222312966
2KS3

Dosen Pengampu:

Yuliagnis Transver Wijaya, S.ST., M.Sc.

**PROGRAM STUDI D-IV KOMPUTASI STATISTIK
POLITEKNIK STATISTIKA STIS
TAHUN AKADEMIK 2024/2025**

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi telah membawa dampak besar dalam berbagai aspek kehidupan, termasuk dalam bidang analisis data dan statistika. Di era digital saat ini, data tidak hanya semakin melimpah tetapi juga semakin beragam, mencakup data numerik, tekstual, hingga data spasial. Oleh karena itu, diperlukan metode dan alat yang mampu mengolah serta menyajikan data secara efektif agar dapat memberikan informasi yang bermakna bagi para pengambil keputusan maupun masyarakat umum. Salah satu pendekatan yang banyak digunakan dalam beberapa tahun terakhir adalah dashboard interaktif. Dashboard merupakan platform visualisasi data yang menyajikan informasi dalam bentuk grafik, tabel, maupun peta interaktif, yang memungkinkan pengguna untuk mengeksplorasi dan menganalisis data secara langsung. Keunggulan utama dashboard interaktif adalah kemampuannya dalam menyajikan data secara ringkas, dinamis, dan mudah dipahami bahkan oleh pengguna non-teknis.

Dalam konteks statistika sosial, pemanfaatan dashboard interaktif menjadi sangat relevan karena kompleksitas data yang digunakan seringkali tinggi, baik dari segi jumlah variabel, keterkaitan antar indikator, hingga distribusi geografisnya. Salah satu indikator penting dalam statistika sosial adalah Social Vulnerability Index (SOVI) atau Indeks Kerentanan Sosial. Indeks ini digunakan untuk mengukur tingkat kerentanan sosial suatu wilayah berdasarkan kombinasi berbagai indikator sosial, ekonomi, dan demografis, seperti tingkat kemiskinan, kepadatan penduduk, tingkat pendidikan, usia penduduk, dan sebagainya. SOVI bertujuan untuk mengidentifikasi kelompok atau wilayah yang memiliki potensi paling tinggi untuk mengalami dampak buruk akibat bencana alam, krisis ekonomi, atau tekanan sosial lainnya.

Penggunaan SOVI menjadi lebih kuat ketika dikombinasikan dengan data spasial dalam bentuk peta geospasial. Data geospasial memungkinkan visualisasi informasi SOVI dalam konteks wilayah, sehingga pola kerentanan sosial dapat dipetakan secara lebih intuitif dan aplikatif. Misalnya, daerah-daerah dengan nilai SOVI tinggi dapat diidentifikasi dengan cepat pada peta, memungkinkan intervensi kebijakan yang lebih tepat sasaran. Integrasi data kuantitatif dan spasial ini kemudian menjadi dasar dalam pengembangan dashboard interaktif yang tidak hanya informatif, tetapi juga responsif terhadap kebutuhan pengguna dalam mengakses dan menganalisis data kerentanan sosial.

Dalam proyek ini, dikembangkan sebuah dashboard interaktif berbasis R Shiny yang mengintegrasikan data SOVI dengan data geospasial. Pengembangan dashboard ini dilaksanakan dalam rangka memenuhi tugas ujian akhir mata kuliah Komputasi Statistik, dengan tujuan untuk menerapkan prinsip-prinsip komputasi data dalam penyajian informasi statistik. Melalui platform ini, data SOVI yang semula bersifat tabular dapat ditampilkan dalam bentuk visual yang lebih mudah dianalisis, seperti peta interaktif, grafik, dan tabel dinamis. Penggunaan teknologi *R Shiny* memungkinkan integrasi antara pemrosesan data, analisis statistik, dan penyajian visual dalam satu kesatuan sistem yang terstruktur dan responsif.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, didapatkan rumusan masalah sebagai berikut:

- a. Bagaimana cara menyajikan informasi indeks kerentanan sosial secara interaktif, visual, dan mudah dipahami?
- b. Bagaimana mengintegrasikan data statistik dan spasial untuk menghasilkan visualisasi yang informatif?

- c. Bagaimana memberikan fleksibilitas kepada pengguna untuk melakukan analisis mandiri melalui dashboard?

1.3 Tujuan

Tujuan dari proyek ini adalah:

- a. Mengembangkan dashboard interaktif menggunakan R Shiny untuk menyajikan analisis terhadap data SOVI.
- b. Mengintegrasikan data statistik dan spasial ke dalam satu platform visualisasi.
- c. Menyediakan fitur eksplorasi dan analisis data statistik dasar hingga lanjutan (inferensia dan regresi), yang dapat diakses dan digunakan secara mandiri oleh pengguna.

1.4 Manfaat

Manfaat dari pengembangan dashboard ini antara lain:

- a. Menyediakan alat bantu analisis berbasis data yang mudah digunakan oleh akademisi, peneliti, maupun pengambil kebijakan
- b. Meningkatkan pemahaman terhadap pola kerentanan sosial secara visual dan terstruktur.
- c. Mendukung proses pembelajaran dalam penerapan komputasi statistik dan pengembangan sistem informasi berbasis data.

BAB II

TAHAPAN ANALISIS DATA

2.1 Identifikasi Kebutuhan Data

Tahap awal dalam proses analisis adalah mengidentifikasi dan memahami kebutuhan informasi yang hendak disampaikan kepada pengguna. Dalam konteks proyek ini, isu yang menjadi fokus utama adalah kerentanan sosial di berbagai wilayah, yang menjadi salah satu dimensi penting dalam perencanaan pembangunan berkelanjutan, mitigasi bencana, serta intervensi sosial-ekonomi. Oleh karena itu, diperlukan data yang tidak hanya bersifat numerik, tetapi juga memiliki keterkaitan spasial agar dapat dianalisis dan divisualisasikan secara komprehensif.

Jenis data yang digunakan dalam proyek ini melibatkan dua komponen utama, yaitu:

a. Data Atribut SOVI (Social Vulnerability Index):

Merupakan data cross-section yang mencakup beragam indikator sosial, ekonomi, dan demografi. Beberapa variabel yang termasuk di dalamnya antara lain tingkat kemiskinan, kepadatan penduduk, rasio ketergantungan, tingkat pendidikan, dan akses terhadap fasilitas kesehatan. Skala pengukuran variabel bervariasi, dengan sebagian besar berskala rasio (misalnya persentase penduduk miskin) dan sebagian lainnya nominal (misalnya klasifikasi wilayah).

b. Data Spasial Wilayah Administratif:

Data ini merepresentasikan struktur wilayah dalam bentuk spasial (geometris), seperti batas desa, kecamatan, atau kabupaten. Format data disiapkan dalam bentuk GeoJSON, yang mendukung penyimpanan informasi geometris serta atribut tambahan yang relevan. Data spasial ini memungkinkan visualisasi data SOVI secara geografis, mendukung pemetaan interaktif, serta membantu dalam analisis spasial berbasis wilayah.

2.2 Pengambilan Data

Data yang digunakan dalam proyek ini diperoleh secara daring dari sumber-sumber resmi dan kredibel, guna menjamin validitas serta keterandalan analisis yang dilakukan. Terdapat dua jenis data utama yang dikumpulkan, yaitu data statistik dan data spasial wilayah administratif, yang kemudian akan diintegrasikan untuk membentuk satu kesatuan analisis berbasis spasial.

a. Data Statistik SOVI

Data statistik diperoleh dari repositori publik yang tersedia di GitHub melalui tautan:

https://raw.githubusercontent.com/bmlmcmc/naspaclust/main/data/sovi_data.csv

Dataset ini berisi informasi terkait indikator kerentanan sosial (*Social Vulnerability Index* - SOVI) pada tingkat kabupaten/kota di Indonesia. Terdapat total 17 variabel, meliputi dimensi sosial, ekonomi, demografi, serta akses terhadap infrastruktur dasar. Variabel dalam data ini antara lain:

Nama Kolom	Deskripsi
DISTRICTCODE	Kode wilayah administratif
CHILDREN	Persentase penduduk usia di bawah lima tahun
FEMALE	Persentase penduduk perempuan
ELDERLY	Persentase penduduk usia ≥ 65 tahun dan populasi lanjut usia
FHEAD2	Persentase rumah tangga dengan kepala rumah tangga perempuan
FAMILYsiz	Rata-rata jumlah anggota rumah tangga dalam satu wilayah
NOELECTRIC	Persentase rumah tangga yang tidak menggunakan listrik sebagai sumber penerangan
LOWEDU	Persentase penduduk ≥ 15 tahun yang tidak tamat sekolah (berpendidikan rendah)
GROWTH	Laju pertumbuhan penduduk (persentase perubahan jumlah penduduk)
POVERTY	Persentase penduduk miskin
ILLITERATE	Persentase penduduk yang tidak dapat membaca dan menulis
NOTRAINING	Persentase rumah tangga yang tidak mengikuti pelatihan bencana
DPRONE	Persentase rumah tangga yang tinggal di daerah rawan bencana
RENTED	Persentase rumah tangga yang tinggal di rumah sewa

NOSWERE	Persentase rumah tangga yang tidak memiliki sistem drainase
TAPWATER	Persentase rumah tangga yang menggunakan air leding sebagai sumber air utama
POPULATION	Jumlah total penduduk di wilayah tersebut

b. Data Spasial Wilayah Administratif

Data spasial diperoleh dalam bentuk file GeoJSON yang berisi representasi geometris wilayah administratif tingkat kabupaten/kota di Indonesia. Data ini peneliti dapatkan dari dosen pengampu mata kuliah Sistem Informasi Geografis pada semester 3. Selain itu, data ini juga dapat diunduh dari referensi wilayah administratif nasional yang lazim digunakan dalam pemetaan dan analisis geospasial seperti melalui tautan berikut: <https://drive.google.com/file/d/1xG8Ma7ZHDeqnxWes6IZjBhawl6ihS9Kp/view>.

Setiap entitas wilayah dalam file ini dilengkapi atribut identifikasi seperti:

Nama Kolom	Deskripsi
FID	Nomor urut fitur spasial (internal ID)
gid	Global ID unik untuk setiap entitas wilayah
kdkab	Kode kabupaten/kota (biasanya 2 digit, sesuai standar BPS)
kdprov	Kode provinsi tempat kabupaten tersebut berada
nmkab	Nama kabupaten/kota
nmprov	Nama provinsi

2.3 Integritas Data

Setelah data statistik dan data spasial dikumpulkan, langkah penting berikutnya adalah melakukan integrasi data untuk memastikan keselarasan antara informasi atribut (data SOVI) dengan representasi geografisnya (data spasial). Proses ini bertujuan untuk menyatukan dua sumber data berbeda ke dalam satu kesatuan struktur data yang dapat digunakan untuk analisis dan visualisasi spasial secara interaktif.

1. Prinsip Penggabungan

Penggabungan dilakukan berdasarkan kode wilayah administratif, yang berfungsi sebagai kunci relasional antara kedua tabel data. Dalam hal ini, digunakan kombinasi kode provinsi (kdprov) dan kode kabupaten/kota (kdkab) dari data spasial untuk dicocokkan dengan kode wilayah pada data SOVI (DISTRICTCODE).

2. Teknik Penggabungan

Proses integrasi dilakukan menggunakan fungsi join yang tersedia dalam software GIS. Langkah-langkah umum yang dilakukan meliputi:

- a. Membuat variabel “kodeprkab” pada data spasial sebagai gabungan dari “kdprov” dan “kdkab” (misalnya "53" + "05" menjadi "5305").
- b. Melakukan left join atau inner join antara data spasial dan data SOVI berdasarkan variabel “kodeprkab” pada data spasial dan “DISTRICTCODE” pada data SOVI.
- c. Memastikan bahwa seluruh wilayah dalam data spasial memiliki kecocokan dengan data SOVI (atau sebaliknya).

3. Verifikasi Hasil Integrasi

Setelah proses penggabungan dilakukan, dilakukan verifikasi untuk menjamin integritas hasil integrasi, antara lain dengan:

- a. Mengecek jumlah baris hasil gabungan (harus sesuai dengan jumlah baris pada data SOVI).
- b. Memastikan bahwa tidak ada nilai NA yang muncul akibat ketidaksesuaian kode wilayah.
- c. Memvisualisasikan peta hasil gabungan sebagai uji awal keberhasilan integrasi.

Dengan integrasi ini, setiap entitas spasial (misalnya: Kabupaten Kupang, Provinsi Nusa Tenggara Timur) tidak hanya memiliki geometri wilayahnya, tetapi juga memuat seluruh indikator kerentanan sosial yang relevan dari data SOVI. Struktur data

hasil gabungan ini selanjutnya menjadi fondasi utama untuk analisis spasial, visualisasi peta interaktif, serta penyajian informasi dalam dashboard.

Data gabungan dari data spasial dan data SOVI dapat diakses melalui tautan berikut:

https://raw.githubusercontent.com/amandatarih/projek-dashboard/main/sovi_data.geojson

4. Pembangunan Business Intelligence

Dashboard dikembangkan sebagai implementasi business intelligence berbasis data statistik dan spasial untuk menyajikan informasi secara ringkas, informatif, dan mudah dipahami. Platform yang digunakan dalam pengembangan adalah RShiny, yaitu sebuah framework dari bahasa R yang memungkinkan pembuatan aplikasi web interaktif berbasis data secara dinamis. Dengan menggunakan R Shiny, pengguna dapat berinteraksi langsung dengan data tanpa perlu melakukan pengolahan manual di luar aplikasi.

Analisis utama pada dashboard mencakup:

- a. Statistik deskriptif
- b. Uji asumsi dan uji inferensia
- c. Regresi linear berganda
- d. Visualisasi spasial (peta tematik)
- e. Fitur interaktif seperti filter, download, dan transformasi data

Yang menjadi keunggulan dari dashboard ini adalah adanya interpretasi langsung yang disajikan pada setiap bagian tampilan. Setiap output analisis dilengkapi dengan narasi singkat yang menjelaskan arti hasil tersebut. Hal ini dirancang untuk meningkatkan keterpahaman pengguna, bahkan bagi mereka yang tidak memiliki latar belakang statistik, serta mengurangi ketergantungan terhadap referensi eksternal atau dokumen pendukung.

BAB III

FITUR DASHBOARD

3.1 Menu dashboard

Dashboard ini dikembangkan menggunakan bahasa pemrograman R, dengan memanfaatkan *framework* R Shiny sebagai platform utama untuk membangun aplikasi web yang dinamis dan interaktif. Pengembangan ini juga didukung oleh berbagai library R yang secara fungsional saling melengkapi, antara lain:

- a. shinydashboard – untuk membangun struktur tampilan dashboard yang responsif dengan sistem sidebar dan tab utama;
- b. ggplot2 – untuk menghasilkan grafik statistik yang informatif dan estetik;
- c. plotly – untuk menyajikan visualisasi interaktif berbasis grafik;
- d. leaflet – untuk menampilkan peta interaktif berbasis data spasial;
- e. DT – untuk membuat tabel dinamis yang dapat difilter, diurutkan, dan diekspor oleh pengguna;
- f. dplyr – untuk manipulasi data secara efisien;
- g. sf – untuk menangani data spasial dalam format simple features;
- h. dan lain-lain.

Struktur Navigasi dan Fungsionalitas

Desain antarmuka dashboard mengadopsi struktur sidebar menu, yang membagi fungsionalitas aplikasi ke dalam beberapa bagian utama. Setiap bagian dirancang untuk melayani fungsi spesifik secara modular, sehingga pengguna dapat mengakses dan menggunakan setiap fitur secara mandiri tanpa ketergantungan terhadap urutan penggunaan tertentu. Adapun menu utama dalam dashboard ini meliputi:

- a. Beranda
- b. Manajemen Data
- c. Eksplorasi Data
- d. Uji Asumsi
- e. Statistik Inferensia
- f. Regresi Linear

3.2 Deskripsi Menu Dashboard

3.2.1 Beranda

Menu ini menyajikan:

- a. Deskripsi umum dashboard
- b. Tujuan pengembangan
- c. Tautan ke metadata dan sumber data
- d. Penjelasan struktur data SOVI

Tujuan utama dari menu ini adalah memberikan konteks awal kepada pengguna sebelum melakukan eksplorasi data lebih lanjut.

3.2.2 Manajemen Data

Fitur utama:

- a. Konversi variabel numerik menjadi kategorik berdasarkan kuartil atau batas nilai kustom
- b. Tabel hasil transformasi yang dapat diunduh
- c. Penjelasan langsung terhadap arti hasil transformasi

Fungsi ini memudahkan pengguna dalam membandingkan wilayah dengan kategori tingkat kerentanan, misalnya: "tinggi", "sedang", atau "rendah".

3.2.3 Eksplorasi Data

Menu eksplorasi data menampilkan:

- a. Statistik deskriptif (mean, median, sd, min, max) dalam bentuk tabel
- b. Korelasi antarvariabel dalam bentuk korelogram (corrplot)
- c. Peta tematik interaktif yang memvisualisasikan nilai variabel tertentu untuk masing-masing wilayah
- d. Visualisasi lainnya seperti histogram, boxplot, dan barplot
- e. Tabel interaktif dengan opsi pencarian dan pemfilteran

Peta dibuat menggunakan leaflet, dengan warna yang disesuaikan berdasarkan skala nilai variabel.

3.2.4 Uji Asumsi

Fitur ini menyediakan:

- a. Uji normalitas (Shapiro-Wilk) dan visualisasi QQ Plot
- b. Uji homogenitas (Levene test)
- c. Interpretasi hasil uji secara otomatis

Hasil uji ditampilkan dalam format tabel dan grafik untuk memudahkan interpretasi.

3.2.5 Statistik Inferensia

Menu ini terdiri dari tiga bagian utama:

1. Uji Beda Rata-rata:
 - a. Uji satu kelompok
 - b. Uji dua kelompok
2. Uji Proporsi dan Uji Ragam:
 - a. `prop.test()` untuk satu atau dua proporsi
 - b. `var.test()` untuk membandingkan ragam dua kelompok

3. Analisis Varians (ANOVA):

- a. ANOVA satu arah
- b. ANOVA dua arah

Setiap hasil uji disertai:

- Nilai statistik uji dan p-value
- Grafik boxplot atau barplot sesuai konteks
- Interpretasi langsung terhadap hasil pengujian

3.2.6 Regresi Linear Berganda

Fitur ini mencakup:

- a. Pemodelan regresi linear berganda dengan pilihan variabel bebas dan terikat
- b. Tabel koefisien model dan goodness-of-fit (R^2 , Adjusted R^2)
- c. Visualisasi: Scatterplot
- d. Uji asumsi regresi:
 - Normalitas
 - Multikolinearitas
 - Homoskedastisitas

Pengguna dapat mengganti variabel input secara dinamis dan memperoleh interpretasi otomatis dari hasil model.

BAB IV

PENUTUP

5.1 Kesimpulan

Dashboard ini menyajikan informasi indeks kerentanan sosial secara visual, interaktif, dan mudah dipahami melalui berbagai komponen seperti grafik deskriptif, tabel dinamis, serta peta tematik interaktif. Dengan memanfaatkan kombinasi elemen visual dan interpretasi langsung pada tampilan, pengguna dapat memahami data tanpa memerlukan pengetahuan teknis yang mendalam. Antarmuka yang ramah pengguna dan responsif semakin memperkuat tujuan penyampaian informasi yang komunikatif dan efektif.

Integrasi data statistik dan spasial dilakukan melalui penggabungan atribut SOVI dengan geometri wilayah administratif ke dalam satu berkas geojson, yang memudahkan pemetaan dan analisis spasial. Proses ini memungkinkan setiap entitas spasial (kabupaten/kota) ditautkan langsung dengan indikator-indikator kerentanan sosialnya. Hasil integrasi ini menjadi fondasi utama dalam membangun visualisasi peta interaktif, yang mampu mengungkap pola spasial kerentanan secara lebih nyata dan eksploratif.

Dashboard ini memberikan fleksibilitas tinggi kepada pengguna untuk melakukan analisis mandiri, baik secara deskriptif maupun inferensial. Pengguna dapat memilih variabel, melakukan uji asumsi, membangun model regresi linear berganda, serta mengeksport hasil dalam berbagai format. Fitur-fitur seperti dropdown pemilihan wilayah, filter data, dan opsi transformasi menjadikan dashboard sebagai alat analisis yang tidak hanya statis, tetapi juga adaptif terhadap kebutuhan pengguna.

5.2 Saran

Agar pengembangan dashboard serupa di masa mendatang dapat lebih optimal dan aplikatif, terdapat beberapa saran yang dapat dipertimbangkan berdasarkan evaluasi fungsionalitas dan potensi pengembangannya:

1. Implementasi Implementasi Pembobot Jarak dalam Klasterisasi

Meskipun dashboard saat ini telah menerapkan algoritma K-Means clustering, namun proses pembentukan kluster masih belum mempertimbangkan dimensi spasial secara eksplisit. Oleh karena itu, disarankan untuk mengimplementasikan pembobot jarak geografis (spatial weight matrix) dalam proses klasterisasi, misalnya dengan pendekatan spatially-constrained clustering atau integrasi fungsi jarak Euclidean antar wilayah. Dengan demikian, kluster yang terbentuk tidak hanya serupa dari sisi atribut sosial, tetapi juga berdekatan secara geografis, yang lebih relevan dalam konteks pemetaan kerentanan sosial.

2. Integrasi Data Time Series

Jika tersedia, integrasi data dalam format deret waktu (time series) akan sangat berguna untuk memantau dinamika dan tren perubahan indeks kerentanan sosial antar periode. Pendekatan ini memungkinkan pengguna melakukan analisis longitudinal untuk mendeteksi wilayah yang mengalami peningkatan atau penurunan kerentanan secara signifikan.

3. Penyempurnaan Antarmuka Pengguna (UI/UX)

Meskipun dashboard telah interaktif, pengalaman pengguna dapat ditingkatkan lebih lanjut melalui penyediaan fitur tutorial bawaan atau panduan interaktif (interactive walkthrough), khususnya bagi pengguna awam. Fitur ini akan membantu pengguna memahami alur kerja aplikasi dan fungsi tiap menu tanpa perlu membuka dokumentasi terpisah.

4. Penguatan Keamanan dan Otentikasi Pengguna

Untuk penggunaan dashboard dalam konteks kebijakan internal, penelitian sensitif, atau publikasi terbatas, pengembangan sistem otentikasi pengguna (user authentication) dapat diterapkan. Langkah ini bertujuan untuk menjaga validitas penggunaan data serta mengatur akses berdasarkan otorisasi tertentu, seperti peran pengguna atau institusi.

5. Fitur Input Data Mandiri oleh Pengguna

Untuk meningkatkan fleksibilitas dan daya guna dashboard, disarankan penambahan fitur unggah data mandiri (custom data input) oleh pengguna. Dengan demikian, dashboard dapat berfungsi tidak hanya sebagai alat eksplorasi data yang sudah tersedia, tetapi juga sebagai kalkulator analisis statistik dinamis. Fitur ini memungkinkan pengguna mengunggah dataset mereka sendiri dan langsung memperoleh visualisasi, uji statistik, serta pemodelan regresi dengan struktur dan logika yang sama.

Implementasi saran-saran tersebut diharapkan dapat memperluas cakupan penggunaan dashboard, meningkatkan interaktivitas, serta memperkuat peran dashboard sebagai alat bantu komputasi statistik yang adaptif dan berkelanjutan.

LAMPIRAN

Akses Dashboard Online : https://amandahapsari.shinyapps.io/Dashboard_Sovi/







