

**LAPORAN UJIAN AKHIR SEMESTER**  
**PEMBUATAN DASHBOARD RSHINY “Pusat Data Kesiapsiagaan Bencana (PDKB) :**  
***Rancang Bangun Dashboard Analitis Berbasis R Shiny untuk Mendukung Kebijakan Mitigasi”***  
Disusun untuk Memenuhi Kuis SPSS Mata Kuliah Komputasi Statistik



Dosen Pengampu :  
Yuliagnis Transver Wijaya, S.ST, M.Sc.

Disusun Oleh :  
Maulida Mutiriani  
222313200  
2KS2

**PROGRAM STUDI D-IV KOMPUTASI STATISTIK**

**POLITEKNIK STATISTIKA STIS**

**2024/2025**

## **PAKTA INTEGRITAS UJIAN AKHIR SEMESTER**

### **Mata Kuliah: Komputasi Statistik**

Prodi: DIV Komputasi Statistik

Saya yang bertanda tangan di bawah ini:

NIM : 222313200

Nama : Maulida Mutiriani

Kelas : 2KS2

Dengan ini menyatakan bahwa dashboard dan laporan dashboard Komstat yang saya buat dan kumpulkan untuk Ujian Akhir Semester Mata Kuliah Komputasi Statistik adalah hasil karya saya sendiri. Saya menyatakan tidak melakukan tindakan kecurangan, plagiat, atau pelanggaran akademik lainnya.

Apabila di kemudian hari terbukti melakukan pelanggaran terhadap Pakta Integritas ini, saya bersedia menerima sanksi akademik sesuai ketentuan yang berlaku di Politeknik Statistika STIS.

Demikian Pakta Integritas ini saya buat dengan sebenar-benarnya dan tanpa paksaan dari pihak manapun.

Jakarta, 23 Juli 2025



(Maulida Mutiriani)

## **BAB I**

### **PERENCANAAN DAN PERANCANGAN DASHBOARD**

#### **A. Latar Belakang dan Tujuan Project**

Indonesia, dengan posisi geografisnya di Cincin Api Pasifik dan sebagai titik pertemuan lempeng tektonik utama dunia, adalah negara yang sangat rentan terhadap berbagai bencana alam seperti gempa bumi, tsunami, dan letusan gunung berapi. Kondisi ini mendorong pergeseran paradigma manajemen bencana dari responsif menjadi proaktif, dengan fokus utama pada kesiapsiagaan. Kesiapsiagaan yang efektif bergantung pada pemahaman mendalam terhadap faktor-faktor sosial-ekonomi yang membentuk kerentanan masyarakat, yang sering disebut sebagai 'akar masalah kerentanan'.

Meskipun data terkait indikator-indikator kerentanan sosial tersedia dari berbagai sumber seperti Badan Pusat Statistik (BPS) melalui Survei Sosial Ekonomi Nasional (SUSENAS) , data ini seringkali tersebar dan belum terintegrasi dalam satu platform yang memungkinkan analisis hubungan secara simultan. Pembuat kebijakan seringkali kesulitan menjawab pertanyaan strategis seperti, "Faktor manakah yang memiliki hubungan paling kuat terhadap rendahnya tingkat kesiapsiagaan suatu daerah?". Tanpa alat analisis yang tepat, jawaban atas pertanyaan ini cenderung bersifat asumtif.

Dashboard Analitis Interaktif berbasis R Shiny yang dinamai “Vulnera”. *Dashboard* ini dirancang untuk mengintegrasikan berbagai variabel kerentanan sosial dan menyediakan alat analisis statistik untuk menguji serta memvisualisasikan hubungan antar variabel tersebut dengan tingkat kesiapsiagaan bencana. Tujuannya adalah untuk memberikan bukti empiris yang dapat menuntun alokasi sumber daya dan perancangan program mitigasi yang lebih efektif dan tepat sasaran.

#### **B. Penentuan Objek dan Ruang Lingkup Data**

Objek Formal (Konsep yang Diteliti): Analisis hubungan antara faktor-faktor kerentanan sosial dengan tingkat kesiapsiagaan bencana di Indonesia. Fokusnya adalah memahami bagaimana variabel-variabel demografi, pendidikan, ekonomi, dan infrastruktur memengaruhi atau berhubungan dengan kerentanan dan kesiapsiagaan

##### **a. Objek Material**

- **Data Sekunder:** Data utama berasal dari Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2017 yang dilaksanakan oleh Badan Pusat Statistik (BPS). Data ini mencakup 511 kabupaten/kota di seluruh Indonesia dan telah diolah menjadi *dataset* siap pakai (data\_sovi\_transformed.csv).
- **Dashboard R Shiny:** Purwarupa (*prototype*) *dashboard* analitis PDKB yang dikembangkan sebagai alat utama untuk melakukan analisis, visualisasi, dan mendukung pengambilan keputusan.

### b. Variabel yang digunakan

*Dashboard* ini memanfaatkan semua 17 variabel yang tersedia dari *dataset* data\_sovi\_transformed.csv untuk analisis komprehensif. Variabel-variabel tersebut meliputi:

Nama Variabel	Tipe (Teknis)	Skala Pengukura n	Penjelasan / Deskripsi
DISTRICTCOD E	Numerik (Integer)	Nominal	Kode unik wilayah administrasi untuk setiap kabupaten/kota
nmkab	Teks (String)	Nominal	Nama Kabupaten/Kot a.
nmprov	Teks (String)	Nominal	Nama Provinsi.
CHILDREN	Teks (Numeri k)	Rasio (%)	Persentase penduduk anak-anak (biasanya di

Nama Variabel	Tipe Data (Teknis)	Skala Pengukuran n	Penjelasan / Deskripsi
			bawah 5 atau 15 tahun).
FEMALE	Teks (Numerik)	Rasio (%)	Persentase penduduk berjenis kelamin perempuan.
ELDERLY	Teks (Numerik)	Rasio (%)	Persentase penduduk lanjut usia (biasanya 65 tahun ke atas).
FHEAD	Teks (Numerik)	Rasio (%)	Persentase rumah tangga yang dikepalai oleh perempuan.
FAMILYSIZE	Teks (Numerik)	Rasio	Rata-rata jumlah anggota dalam satu rumah tangga.
NOELECTRIC	Teks (Numerik)	Rasio (%)	Persentase rumah tangga yang tidak memiliki akses listrik.

Nama Variabel	Tipe Data (Teknis)	Skala Pengukura n	Penjelasan / Deskripsi
LOWEDU	Teks (Numerik)	Rasio (%)	Persentase penduduk (usia 15+) dengan tingkat pendidikan rendah.
GROWTH	Teks (Numerik)	Rasio (%)	Tingkat pertumbuhan penduduk tahunan.
POVERTY	Teks (Numerik)	Rasio (%)	Persentase penduduk yang hidup di bawah garis kemiskinan.
ILLITERATE	Teks (Numerik)	Rasio (%)	Persentase penduduk yang buta huruf.
NOTRAINING	Teks (Numerik)	Rasio (%)	Persentase penduduk yang tidak pernah mengikuti pelatihan kesiapsiagaan.
DPRONE	Teks (Numerik)	Rasio (%)	Persentase penduduk yang

Nama Variabel	Tipe	Skala	Penjelasan /
	Data	Pengukura	Deskripsi
	(Teknis)	n	tinggal di area rawan bencana.
RENTED	Teks (Numerik)	Rasio (%)	Persentase rumah tangga yang menyewa tempat tinggal (tidak memiliki rumah sendiri).
NOSEWER	Teks (Numerik)	Rasio (%)	Persentase rumah tangga tanpa akses sanitasi/saluran pembuangan yang layak.
TAPWATER	Teks (Numerik)	Rasio (%)	Persentase rumah tangga yang tidak memiliki akses air bersih/ledeng.
POPULATION	Numerik (Integer)	Rasio	Jumlah total populasi di wilayah tersebut.

### C. Perancangan Fitur dan Struktur Dashboard

Dashboard ini dirancang dengan fokus pada pengalaman pengguna (*User Experience - UX*) yang intuitif dan informatif, mengikuti praktik terbaik pengembangan aplikasi Shiny.

## a. Struktur dan Desain *Dashboard*:

- **Tata Letak Utama (*Layout*):**

- *Header*: Bagian atas *dashboard* yang menampilkan judul "PDKB (Pusat Data Kesiapsiagaan Bencana)" secara jelas.
- *Sidebar* (Menu Navigasi): Menu utama di sisi kiri yang menjadi pusat navigasi. Strukturnya sangat terorganisir dengan menggunakan hierarki dua tingkat (menu utama dan sub-menu), memungkinkan pengguna mengakses fitur spesifik dengan cepat.
- *Body* (Area Konten): Area utama tempat semua *output* (tabel, peta, *plot*, dan interpretasi) ditampilkan. Tata letaknya rapi, menggunakan *card* atau *box* untuk mengelompokkan informasi secara logis.

- **Prinsip Desain & User Experience (UX):**

- Informatif dan Edukatif: *Dashboard* ini menonjol karena fokusnya pada edukasi pengguna. Hampir setiap modul analisis dilengkapi dengan "Petunjuk", penjelasan hipotesis ( $H_0/H_1$ ), dan "Panduan Interpretasi".
- Interpretasi Otomatis: Fitur ini secara konsisten menyajikan kesimpulan analisis dalam bahasa yang sederhana (misalnya, "  Data Berdistribusi Normal", "  Tidak ada masalah multikolinearitas"), yang menjembatani kesenjangan antara *output* statistik mentah dan pemahaman pengguna yang lebih awam.
- Alur Kerja Terpandu: Desain *dashboard* secara implisit memandu pengguna melalui alur analisis. Contohnya, pesan seperti "Silakan jalankan kategorisasi terlebih dahulu" pada modul Uji Proporsi, mengarahkan pengguna ke langkah yang benar sebelum melanjutkan.
- Fungsionalitas Unduh Lengkap: Setiap modul analisis dilengkapi dengan fitur unduh yang komprehensif, memungkinkan pengguna untuk menyimpan hasil analisis dalam berbagai format seperti data (CSV), laporan (PDF), dan visualisasi (PNG/PDF).

## b. Fitur fungsional per menu

Berikut adalah rincian fungsionalitas dari setiap menu yang ada di *sidebar dashboard* Anda:

### 1. Beranda

- Fungsi: Sebagai halaman pengantar yang memberikan konteks proyek dan gambaran umum tentang kemampuan *dashboard*.
- Fitur: Bagian "Tentang Dashboard" menjelaskan misi dan tujuan PDKB sebagai fondasi data terpercaya. "Rincian Fitur Utama" memberikan daftar singkat kemampuan analisis. Terdapat juga "Kamus Data / Tabel Data Lengkap" yang menampilkan data mentah yang digunakan, lengkap dengan variabel seperti nmkab, nmprov, dan indikator kerentanan. Fitur unduh komprehensif tersedia untuk laporan keseluruhan dalam format PDF,
- *metadata* dalam PDF, serta data lengkap dalam format CSV dan Excel.

## 2. Manajemen Data

- Fungsi: Modul ini berfungsi untuk pra-pemrosesan data sebelum analisis lebih lanjut.
- Fitur (Kategorisasi Data):
  - Memungkinkan pengguna mengubah variabel numerik (kontinu) menjadi variabel kategorik.
  - *Input*: Pengguna memilih variabel yang ingin dikategorikan (misalnya "Indeks Kerentanan" atau "Anakanak") dan menentukan jumlah kategori yang diinginkan (antara 2 hingga 5).
  - *Output*: Menampilkan visualisasi distribusi kategori dalam bentuk grafik batang dan tabel data hasil kategorisasi.
  - *Dashboard* juga memberikan interpretasi hasil kategorisasi (misalnya, distribusi kategori "Sedang", "Rendah", "Tinggi" dan rekomendasi strategis prioritas). Hasil ini dapat diunduh dalam format PDF atau CSV.

## 3. Eksplorasi Data

- Fungsi: Modul ini membantu pengguna memahami karakteristik dan distribusi data secara visual dan deskriptif.
- Fitur:
  - Statistik Deskriptif: Menghasilkan tabel ringkasan statistik (seperti Min, Kuartil 1, Median, Mean, Kuartil 3, Maks, Standar Deviasi, Varians, Koefisien Variasi, *Skewness*, *Kurtosis*, dan *Missing Values*) untuk variabel yang dipilih (misalnya, "Anakanak" atau indikator kerentanan lain). Interpretasi statistik disediakan untuk membantu memahami profil umum dan variabilitas data, serta merekomendasikan metode analisis yang tepat. Hasil ini dapat diunduh dalam format TXT, PDF, atau ZIP [image\_8538ea, 362].

- Histogram: Memvisualisasikan distribusi frekuensi variabel yang dipilih (misalnya, CHILDREN). Grafik ini dilengkapi dengan kurva densitas untuk mengestimasi distribusi probabilitas data. Histogram memberikan gambaran visual tentang sebaran dan frekuensi data, membantu dalam pemilihan metode analisis statistik yang tepat dan identifikasi *outlier* potensial [image\_8538f2]. Hasil dapat diunduh dalam format JPG, PDF, atau ZIP [image\_8538f2].
- Boxplot: Menampilkan distribusi variabel yang dipilih (misalnya, CHILDREN) melalui kuartil, median, dan *outlier* potensialnya. Setiap titik merah merepresentasikan data individual. Interpretasi *boxplot* menyoroti Kuartil pertama (Q1), Kuartil ketiga (Q3), Rentang Interkuartil (IQR), dan jumlah *outlier* yang terdeteksi, yang mungkin perlu diperhatikan lebih lanjut dalam analisis [image\_85390d]. Hasil dapat diunduh dalam format JPG, PDF, atau ZIP [image\_85390d].
- Peta Interaktif: Menampilkan peta Indonesia dengan pewarnaan wilayah (*koroplek*) berdasarkan variabel yang dipilih (misalnya, "Indeks Kerentanan" atau CHILDREN). Pengguna dapat memilih palet warna, dan peta menyediakan panduan interpretasi untuk intensitas warna (misalnya, merah gelap menunjukkan nilai lebih tinggi). Mengklik suatu wilayah akan menampilkan detail spesifik seperti nama provinsi dan nilai variabel [image\_853928]. Visualisasi ini membantu mengidentifikasi wilayah dengan tingkat kerentanan tinggi yang memerlukan perhatian khusus dalam perencanaan kebijakan [image\_853928]. Hasil dapat diunduh dalam format PNG, PDF, atau laporan peta dalam format ZIP [image\_853928, 363].
- Analisis Spasial (Moran's I): Melakukan uji Moran's I untuk otokorelasi spasial untuk indikator yang dipilih (misalnya, CHILDREN) [image\_853946]. Output mencakup nilai Moran's I, *Expected I*, Varians, dan *p-value*. Interpretasi menunjukkan apakah otokorelasi spasial signifikan atau tidak, mengindikasikan apakah distribusi nilai tampak acak secara spasial [image\_853946]. Hasil analisis spasial dapat diunduh dalam format PDF atau ZIP [image\_853946].
- Visualisasi Faktor: Menampilkan grafik batang horizontal ("Profil Kerentanan") untuk membandingkan rata-rata nilai dari setiap indikator (misalnya, Kemiskinan, Pendidikan Rendah, Tanpa Pelatihan). Fitur ini menyediakan panduan interpretasi grafik dan rekomendasi program berdasarkan nilai indikator (prioritas tinggi, menengah, rendah). Hasil visualisasi dapat diunduh dalam format PDF.

- Tabel Data: Menampilkan *dataset* lengkap dalam format tabel interaktif. Data lengkap dapat diunduh dalam format CSV.

#### 4. Uji Asumsi

- Fungsi: Modul ini melakukan uji prasyarat statistik untuk memastikan validitas analisis inferensia.
- Fitur:
  - Uji Normalitas: Melakukan uji Shapiro-Wilk dan Kolmogorov-Smirnov untuk memeriksa apakah data (misalnya, CHILDREN atau Indeks Kerentanan) terdistribusi secara normal. Hasilnya ditampilkan melalui *plot* distribusi (histogram), *Q-Q plot*, *output* statistik (nilai statistik dan *p-value*), dan interpretasi otomatis, yang menyatakan apakah data terdistribusi normal dan apakah uji parametrik seperti *t-test* atau ANOVA dapat digunakan. Tersedia opsi unduh untuk *Q-Q Plot* (JPG), hasil uji (TXT), interpretasi (PDF), dan laporan normalitas (ZIP) [image\_853989].
  - Uji Homogenitas: Melakukan uji Levene dan Bartlett untuk kesamaan varians antar kelompok (misalnya, CHILDREN\_kategori\_3). Ini memberikan *output* statistik (nilai statistik dan *p-value*) dan interpretasi otomatis, mengindikasikan apakah asumsi homogenitas varians terpenuhi. Catatan menyebutkan bahwa uji Levene lebih *robust* terhadap penyimpangan normalitas dibandingkan uji Bartlett [image\_8539a4]. Tersedia opsi unduh untuk hasil uji (TXT), interpretasi (PDF), dan laporan homogenitas (ZIP) [image\_8539a4].

#### 5. Statistik Inferensia

- Fungsi: Modul ini dirancang untuk pengujian hipotesis statistik.
- Fitur:
  - Uji Beda Rata-rata (Uji t):
  - Uji t Satu Sampel: Memungkinkan pengguna untuk menguji apakah rata-rata populasi dari variabel yang dipilih (misalnya, CHILDREN) secara signifikan berbeda dari nilai yang dihipotesiskan (misalnya, 0.05). Ini menyajikan hipotesis nol ( $H_0$ ) dan alternatif ( $H_1$ ), statistik-t, derajat bebas (df), *p-value*, dan interval kepercayaan. Interpretasi secara jelas menyatakan apakah hipotesis nol ditolak dan apakah rata-rata populasi berbeda signifikan dari nilai yang dihipotesiskan [image\_853daa]. Tersedia opsi unduh untuk hasil uji (TXT), interpretasi (PDF), dan laporan uji t 1 sampel (ZIP) [image\_853daa].

- Uji t Dua Sampel: Membandingkan rata-rata variabel yang dipilih (misalnya, CHILDREN) antara dua wilayah yang dipilih (misalnya, Kalimantan dan Sulawesi). Pengguna dapat menentukan apakah akan mengasumsikan varians yang sama. Ini menyediakan hipotesis nol dan alternatif, statistik-t, derajat bebas, *p-value*, dan interval kepercayaan untuk perbedaan rata-rata. Interpretasi menyimpulkan apakah ada perbedaan signifikan antara rata-rata kedua wilayah, dengan catatan asumsi varians yang sama jika berlaku [image\_853daa]. Tersedia opsi unduh untuk hasil uji (TXT), interpretasi (PDF), dan laporan uji t 2 sampel (ZIP) [image\_853daa].
- Uji Proporsi:
- Uji Proporsi Satu Sampel: Mengevaluasi apakah proporsi keberhasilan tertentu (misalnya, kategori "Tinggi" untuk CHILDREN\_kategori\_3) berbeda secara signifikan dari proporsi yang dihipotesiskan (misalnya, 0.1). Ini menampilkan proporsi sampel, total observasi, statistik *Chi-squared*, df, *p-value*, dan interval kepercayaan. Interpretasi menunjukkan apakah proporsi yang diamati berbeda signifikan dari proporsi yang dihipotesiskan [image\_85406b]. Tersedia opsi unduh untuk hasil uji (TXT), interpretasi (PDF), dan laporan proporsi 1 sampel (ZIP) [image\_85406b].
- Uji Proporsi Dua Sampel: Membandingkan proporsi keberhasilan tertentu (misalnya, kategori "Tinggi" untuk CHILDREN\_kategori\_3) antara dua wilayah (misalnya, Kalimantan dan Sulawesi) untuk menentukan apakah ada perbedaan signifikan secara statistik. Ini menampilkan proporsi untuk setiap wilayah, perbedaan proporsi, statistik *Chi-squared*, df, *p-value*, dan interval kepercayaan. Interpretasi menyatakan apakah ada perbedaan signifikan dalam proporsi antara kedua wilayah [image\_85406b]. Tersedia opsi unduh untuk hasil uji (TXT), interpretasi (PDF), dan laporan proporsi 2 sampel (ZIP) [image\_85406b].
- Uji Varians (Uji F):
- Uji Varians Satu Sampel: Memungkinkan pengguna untuk menguji apakah varians populasi untuk variabel numerik yang dipilih (misalnya, CHILDREN) berbeda secara signifikan dari nilai yang dihipotesiskan (misalnya, 0.1). Ini menyediakan hipotesis nol dan alternatif, ukuran sampel, varians sampel, statistik *Chi-square*, df, dan *p-value*. Interpretasi menyatakan apakah hipotesis nol ditolak, menunjukkan perbedaan signifikan dari varians yang dihipotesiskan. Ini juga mencatat sensitivitas uji terhadap asumsi normalitas [image\_85408b]. Tersedia

opsi unduh untuk hasil uji (TXT), interpretasi (PDF), dan laporan varians 1 sampel (ZIP) [image\_85408b].

- Uji Varians Dua Sampel (Uji F): Membandingkan varians variabel numerik yang dipilih (misalnya, CHILDREN) antara dua wilayah (misalnya, Kalimantan dan Sulawesi). Ini menampilkan varians untuk setiap wilayah, statistik F, df, *p-value*, dan interval kepercayaan. Interpretasi menunjukkan apakah ada perbedaan signifikan dalam varians antara kedua wilayah, dan memberikan catatan bahwa jika varians berbeda signifikan, *Welch's t-test* harus digunakan untuk perbandingan rata-rata [image\_85408b]. Tersedia opsi unduh untuk hasil uji (TXT), interpretasi (PDF), dan laporan varians 2 sampel (ZIP) [image\_85408b].
- ANOVA (*Analysis of Variance*): Melakukan ANOVA untuk membandingkan rata-rata antar provinsi atau kelompok lain. Modul ini dilengkapi dengan hasil Uji Lanjutan *Tukey HSD* untuk mengidentifikasi secara spesifik kelompok mana yang berbeda signifikan, serta implikasi kebijakan yang aplikatif.

## 6. Regresi Linear Berganda

- Fungsi: Modul inti untuk memodelkan dan menganalisis hubungan antar variabel.
- Fitur:
  - Model Regresi: Antarmuka untuk memilih satu variabel respons/dependen (Y) (misalnya, CHILDREN) dan beberapa variabel prediktor/independen (X) (misalnya, POPULATION, Luas) untuk membangun model regresi linear. *Output* ringkasan model mencakup panggilan model, *residual*, koefisien (Estimasi, *Std. Error*, *t value*, *Pr(>|t|)*), *residual standard error*, *Multiple R-squared*, *Adjusted R-squared*, dan statistik F dengan *\_p-value* nya. Interpretasi menyediakan wawasan tentang kualitas model (misalnya, nilai *R-squared* dan signifikansi model secara keseluruhan) dan koefisien yang berpengaruh signifikan [image\_8540b0]. Modul ini juga menguraikan asumsi penting yang harus dipenuhi oleh regresi linear dan memberikan solusi jika asumsi tersebut tidak terpenuhi [image\_8540b0]. Hasil model regresi dapat diunduh dalam format TXT, PDF, atau ZIP [image\_8540b0].
  - Diagnostik Model: Fitur validasi model regresi. Ini menampilkan *Plot* Diagnostik standar (*Residuals vs Fitted*, *Normal Q-Q Plot*, *Scale-Location*, *Cook's Distance*) untuk memeriksa asumsi linearitas, normalitas *residual*, dan homoskedastisitas. Modul ini juga mencakup Uji Multikolinearitas (*Variance Inflation Factor* - VIF), dengan interpretasi yang menunjukkan apakah ada

multikolinearitas serius ( $VIF < 5$  = rendah,  $VIF 5-10$  = sedang,  $VIF > 10$  = tinggi). Solusi untuk pelanggaran asumsi (misalnya, transformasi variabel atau model polinomial) juga disediakan [image\_8540b0]. Hasil diagnostik dapat diunduh dalam format JPG (*Plot Diagnostik*), TXT (VIF), PDF (Interpretasi), atau ZIP (Laporan Diagnostik) [image\_8540b0].

## BAB II

### PERSIAPAN DAN PENGELOLAAN DATA UNTUK DASHBOARD

#### A. Sumber Data

Data yang digunakan dalam analisis ini bersumber dari data sekunder Survei Sosial Ekonomi Nasional (SUSENAS) 2017 yang dilaksanakan oleh Badan Pusat Statistik (BPS) Indonesia. Data mentah tersebut telah diolah sebelumnya menjadi *dataset* turunan *data\_sovi\_transformed.csv* yang digunakan dalam proyek ini. Sebelum diimplementasikan dalam *dashboard R Shiny*, *file CSV* tersebut melewati tahap pra-pemrosesan. Tahapan ini mencakup penyesuaian tipe data (konversi kolom numerik yang terbaca sebagai teks menjadi format numerik) dan validasi kelengkapan data untuk memastikan kompatibilitas dengan seluruh fungsi analisis dan visualisasi yang ada di dalam *dashboard*.

**Gambaran Umum Variabel (Berdasarkan *metadata* dan tabel data yang disediakan):**

Nama Variabel	Tipe Data (Teknis)	Skala Pengukuran	Penjelasan / Deskripsi
DISTRICTCODE	Numerik (Integer)	Nominal	Kode unik wilayah administrasi untuk setiap kabupaten/kota. ⓘ
nmkab	Teks (String)	Nominal	Nama Kabupaten/Kota. ⓘ
nmprov	Teks (String)	Nominal	Nama Provinsi. ⓘ
CHILDREN	Teks (Numerik)	Rasio (%)	Persentase penduduk anak-anak (biasanya di bawah 5 atau 15 tahun). ⓘ
FEMALE	Teks (Numerik)	Rasio (%)	Persentase penduduk berjenis kelamin perempuan. ⓘ

ELDERLY	Teks (Numerik)	Rasio (%)	Persentase penduduk lanjut usia (biasanya 65 tahun ke atas). 
FHEAD	Teks (Numerik)	Rasio (%)	Persentase rumah tangga yang dikepalai oleh perempuan. 
FAMILYSIZE	Teks (Numerik)	Rasio	Rata-rata jumlah anggota dalam satu rumah tangga. 
NOELECTRIC	Teks (Numerik)	Rasio (%)	Persentase rumah tangga yang tidak memiliki akses listrik. 
LOWEDU	Teks (Numerik)	Rasio (%)	Persentase penduduk (usia 15+) dengan tingkat pendidikan rendah. 
GROWTH	Teks (Numerik)	Rasio (%)	Tingkat pertumbuhan penduduk tahunan. 
POVERTY	Teks (Numerik)	Rasio (%)	Persentase penduduk yang hidup di bawah garis kemiskinan. 
ILLITERATE	Teks (Numerik)	Rasio (%)	Persentase penduduk yang buta huruf. 
NOTRAINING	Teks (Numerik)	Rasio (%)	Persentase penduduk yang tidak pernah mengikuti pelatihan kesiapsiagaan. 
DPRONE	Teks (Numerik)	Rasio (%)	Persentase penduduk yang tinggal di area rawan bencana. 
RENTED	Teks (Numerik)	Rasio (%)	Persentase rumah tangga yang menyewa tempat tinggal (tidak memiliki rumah sendiri). 
NOSEWER	Teks (Numerik)	Rasio (%)	Persentase rumah tangga tanpa akses sanitasi/saluran pembuangan yang layak. 
TAPWATER	Teks (Numerik)	Rasio (%)	Persentase rumah tangga yang tidak memiliki akses air bersih/ledeng. 
POPULATION	Numerik (Integer)	Rasio	Jumlah total populasi di wilayah tersebut. 

### **a. Optimalisasi Data Spasial untuk Performa**

Penelaahan dan Validasi Awal: Analisis statistik deskriptif awal menunjukkan bahwa dari 511 data yang digunakan, tidak ada *missing value* yang signifikan untuk variabel apapun, dengan semua variabel memiliki 100% data lengkap. Mengenai *outlier*, tidak dilakukan antisipasi atau penghapusan karena data berasal dari BPS. Diasumsikan bahwa *outlier* mencerminkan kondisi nyata yang ekstrem di suatu wilayah dan bukan kesalahan *input*, sehingga menghapusnya berarti mengabaikan informasi mengenai wilayah yang mungkin paling rentan.

### **b. Metode Analisis**

Untuk memastikan performa *dashboard* yang responsif dan interaktif, terutama saat menangani data geospasial yang kompleks untuk 511 kabupaten/kota di seluruh Indonesia, serangkaian teknik optimalisasi data spasial diterapkan. Tantangan utamanya adalah ukuran *file* peta yang besar dan kompleksitas geometri poligon yang dapat memperlambat waktu muat (*loading time*) dan interaktivitas pada *browser* pengguna. Oleh karena itu, pendekatan optimalisasi dilakukan dalam dua tahap utama: pra-pemrosesan data sebelum digunakan oleh aplikasi, dan strategi *rendering* yang efisien di dalam aplikasi R Shiny.

#### **1. Pra-Pemrosesan: Penyederhanaan Geometri dan Standarisasi Format**

- Penyederhanaan Geometri: Peta administrasi digital seringkali memiliki tingkat detail yang sangat tinggi (ribuan titik atau *vertices* per poligon), yang tidak diperlukan untuk visualisasi skala nasional atau provinsi. Dengan menggunakan algoritma penyederhanaan seperti Douglas-Peucker, yang diimplementasikan melalui fungsi `st_simplify()` dari *package sf* di R, jumlah *vertices* ini dikurangi. Proses ini mempertahankan bentuk visual utama dari setiap poligon, namun secara drastis mengurangi ukuran *file* dan beban kerja *browser* untuk merendernya.
- Standarisasi Format dan Proyeksi: Setelah disederhanakan, data disimpan dalam format GeoJSON. Format ini dipilih karena merupakan standar untuk data spasial di aplikasi *web* dan didukung secara *native* oleh *library leaflet*. Selain itu, data juga dipastikan menggunakan sistem proyeksi geografis WGS 84 (EPSG:4326), yang merupakan standar

untuk peta *web*. Ini menghilangkan kebutuhan konversi proyeksi *on-the-fly* yang dapat memperlambat aplikasi.

## 2. Strategi *Rendering* pada Aplikasi R Shiny

- Pemuatan Data Tunggal: Data spasial yang telah dioptimalkan dimuat satu kali saja saat aplikasi pertama kali dijalankan (di luar lingkup *server*). Hal ini mencegah aplikasi membaca ulang *file* besar dari *disk* setiap kali ada sesi pengguna baru, yang secara signifikan mengurangi latensi dan penggunaan memori *server*.
- Pembaruan Peta dengan *leafletProxy*: Untuk interaktivitas—seperti saat pengguna mengubah variabel yang ditampilkan—digunakan fungsi *leafletProxy()*. Alih-alih menggambar ulang seluruh peta dari awal,

*leafletProxy()* hanya mengirimkan dan memperbarui elemen yang berubah (misalnya, warna poligon). Teknik ini menghasilkan transisi yang hampir instan dan pengalaman pengguna yang sangat responsif, seolah-olah semua data sudah ada di *browser*.

Melalui kombinasi strategi ini, *dashboard* PDKB mampu menyajikan data spasial yang kompleks untuk seluruh Indonesia dengan performa yang cepat dan andal, memastikan pengalaman pengguna yang positif.

## C, Metode Analisis

Analisis data dalam penelitian ini menggunakan pendekatan kuantitatif yang diterapkan melalui *dashboard* interaktif R Shiny. Alur analisis dirancang secara sistematis, dimulai dari persiapan data, analisis deskriptif, pengujian hipotesis, hingga pemodelan prediktif.

## 1. Persiapan dan Pra-Pemrosesan Data

- Sumber Data: Data yang digunakan adalah data sekunder yang berasal dari Survei Sosial Ekonomi Nasional (SUSENAS) 2017 oleh BPS.
- Proses: Data melewati tahap pembersihan, termasuk konversi tipe data (mengubah format desimal koma menjadi titik) dan penggabungan data atribut kerentanan ke dalam data spasial (geometris) kabupaten/kota. Hasilnya adalah satu *dataset* final yang terintegrasi.

## 2. Analisis Deskriptif dan Eksploratif

- Statistik Deskriptif: Digunakan untuk meringkas setiap variabel kerentanan dengan ukuran tendensi sentral (*mean, median*) dan dispersi (*standar deviasi, min, max*). Tujuannya adalah untuk memahami profil umum dari setiap indikator di seluruh Indonesia.
- Visualisasi Data:
  - Peta Koroplet (*Choropleth Map*): Menggunakan *library* leaflet untuk memvisualisasikan distribusi spasial dari Indeks Kerentanan, memungkinkan identifikasi visual wilayah dengan tingkat kerentanan tinggi atau rendah.
  - Grafik Batang (*Bar Chart*): Digunakan untuk membandingkan rata-rata nilai dari setiap faktor kerentanan (POVERTY, LOWEDU, dll.) secara langsung, sehingga dapat dengan cepat mengidentifikasi masalah utama di suatu wilayah.

### 3. Analisis Inferensial

- Uji Prasyarat (*Assumption Tests*):
  - Uji Normalitas (Shapiro-Wilk): Dilakukan untuk memeriksa apakah data pada setiap variabel terdistribusi secara normal. Ini adalah syarat untuk validitas uji parametrik.
  - Uji Homogenitas (*Levene's Test*): Digunakan untuk menguji apakah varians data antar kelompok (provinsi) adalah sama (homogen), yang merupakan asumsi penting untuk ANOVA.
- Uji Hipotesis Komparatif:
  - Uji-t (*t-Test*): Diterapkan untuk membandingkan rata-rata nilai kerentanan antara dua kelompok (misalnya, membandingkan tingkat kemiskinan antara dua provinsi).
  - ANOVA (*Analysis of Variance*): Digunakan untuk membandingkan rata-rata nilai kerentanan antara lebih dari dua kelompok (misalnya, membandingkan Indeks Kerentanan di seluruh provinsi di Indonesia) untuk melihat apakah ada perbedaan yang signifikan secara statistik.
  - Uji Lanjutan *Tukey HSD*: Dilakukan setelah ANOVA menunjukkan hasil yang signifikan, untuk mengidentifikasi secara spesifik pasangan provinsi mana yang memiliki perbedaan rata-rata yang signifikan.

#### **4. Analisis Regresi Linear Berganda**

- Ini adalah metode analisis inti untuk mencapai tujuan utama penelitian, yaitu memahami hubungan antar variabel.
- Tujuan Pemodelan: Untuk membangun model statistik yang menjelaskan bagaimana beberapa variabel independen (faktor kerentanan seperti POVERTY, LOWEDU, NOTRAINING, ELDERLY) secara bersama-sama berhubungan dengan atau memengaruhi variabel dependen (Indeks\_Kerentanan).
- Validasi Model: Setelah model terbentuk, dilakukan uji diagnostik untuk memastikan validitas dan keandalannya:
  - Analisis *Residual*: Empat *plot* diagnostik standar digunakan untuk memeriksa asumsi linearitas, normalitas *residual*, dan homoskedastisitas.
  - Uji Multikolinearitas: Menggunakan *Variance Inflation Factor* (VIF) untuk mendeteksi apakah ada korelasi yang terlalu tinggi antar variabel independen, yang dapat mengganggu keandalan model.

# BAB 3

## HASIL TANGKAPAN LAYAR DASHBOARD

### a. Beranda

The screenshot shows the main dashboard interface. At the top, there are four summary cards: 'TOTAL DATA 100 Records' (with a database icon), 'JUMLAH DISTRIK 100 Wilayah' (with a location pin icon), 'JUMLAH VARIABEL 17 Kolumn' (with a list icon), and 'SUMBER DATA ScienceDirect Dataset' (with a document icon). On the left, a sidebar menu lists 'Beranda', 'Manajemen Data', 'Eksplorasi Data', 'Uji Asumsi Data', 'Statistik Inferensia', and 'Regressi Linear Berganda'. The central content area features a section titled 'Metadata dan Informasi Dashboard' with a list of features and a note about the dataset. Below this is a 'Fitur Utama Dashboard' section with numbered items 1 through 6. A 'Fitur Baru' note at the bottom encourages users to download PDFs, Excel/CSV files, and full reports. To the right, there are sections for 'Sumber Data dan Referensi' and 'Spesifikasi Teknis'.

This screenshot is identical to the one above, showing the same dashboard layout, data summary cards, sidebar menu, and central content sections. It includes the 'Metadata dan Informasi Dashboard', 'Fitur Utama Dashboard', 'Fitur Baru' note, 'Sumber Data dan Referensi', and 'Spesifikasi Teknis' sections.

C:/Users/Maul/Downloads/KOMSTAT/UAS - Shiny  
http://127.0.0.1:6091 | Open in Browser | Republish

**Dashboard Vulnera -**

**Dashboard Vulnera - Analisis Kerentanan Sosial Indonesia**

**Beranda**

**Manajemen Data**

**Eksplorasi Data**

**Uji Asumsi Data**

**Statistik Inferensia**

**Regressi Linear Berganda**

**TOTAL DATA**  
100 Records

**JUMLAH DISTRIK**  
100 Wilayah

**JUMLAH VARIABEL**  
17 Kolom

**SUMBER DATA**  
**ScienceDirect**  
Dataset

**Metadata dan Informasi Dashboard**

- Tentang Dataset SOVI (Social Vulnerability Index)
- Dataset ini berisi indikator kerentanan sosial untuk berbagai distrik di Indonesia. Data mencakup 17 variabel utama yang mengukur berbagai aspek kerentanan masyarakat.
- Tujuan Dashboard
- Menyediakan platform analisis komprehensif untuk data kerentanan sosial
- Memfasilitasi eksplorasi data dengan visualisasi interaktif
- Mendukung analisis statistik inferensial dan regresi
- Menyajikan hasil analisis dengan interpretasi yang mudah dipahami
- Fitur download lengkap untuk semua output analisis
- Fitur Utama Dashboard
- 1. Beranda: Informasi umum dan metadata dataset (17 variabel)
- 2. Manajemen Data: Kategorisasi dan penghapusan variabel
- 3. Eksplorasi Data: Statistik deskriptif, visualisasi, dan peta
- 4. Uji Asumsi: Uji normalitas dan homogenitas
- 5. Statistik Inferensia: Uji hipotesis dengan pilihan sampel
- 6. Regressi Linear: Analisis regresi dengan diagnostik
- Fitur Baru: Download PDF untuk Interpretasi, Excel/CSV untuk data, dan laporan lengkap!

**Sumber Data dan Referensi**

- Dataset: Social Vulnerability Index (SoVI) Indonesia
- Artikel Referensi: ScienceDirect Article
- URL Data: [https://raw.githubusercontent.com/bm1ncmc/naspaclost/main/data/sov1\\_data.csv](https://raw.githubusercontent.com/bm1ncmc/naspaclost/main/data/sov1_data.csv)
- Data Spasial: GeoJSON Indonesia untuk visualisasi peta

**Spesifikasi Teknis**

- Platform: R Shiny Dashboard dengan tema responsif dan dark mode
- Package Utama: shiny, shinydashboard, DT, ggplot2, leaflet
- Analisis Spasial: sf, spdep untuk autokorelasi spasial
- Statistik: car, psych untuk uji statistik lanjut
- Tema: Gunmetal dengan Dark Mode (Gradasi Biru)
- Download: PDF, Excel, CSV, dan laporan lengkap

**Metadata Variabel (17 Variabel)**

Variabel	Deskripsi	Tipe
Kode_Distrik	Kode unik untuk setiap distrik/kabupaten.	Kategorik/ID
Anakanak	Persentase populasi di bawah lima tahun.	Numerik
Perempuan	Persentase populasi perempuan.	Numerik
Lansia	Persentase populasi 65 tahun ke atas.	Numerik
Kepala_RT_Perempuan	Persentase rumah tangga dengan kepala rumah tangga perempuan.	Numerik
Ukuran_Keluarga	Rata-rata jumlah anggota rumah tangga di satu distrik.	Numerik
Tanpa_Listrik	Persentase rumah tangga yang tidak menggunakan listrik sebagai sumber penerangan.	Numerik
Pendidikan_Rendah	Persentase populasi 15 tahun ke atas dengan pendidikan rendah.	Numerik
Pertumbuhan	Persentase perubahan populasi (pertumbuhan populasi).	Numerik
Kemiskinan	Persentase penduduk miskin.	Numerik

Showing 1 to 10 of 17 entries

Search:

Previous | 1 | 2 | Next

**Download Beranda**

**Metadata (Excel)** | **Info Dashboard (PDF)** | **Laporan Beranda (ZIP)**

## b. Manajemen Data

C:/Users/Maul/Downloads/KOMSTAT/UAS - Shiny  
http://127.0.0.1:6091 | Open in Browser | Republik

## Manajemen Data

### Kategorisasi Data Kontinu

**Petunjuk:** Pilih variabel numerik untuk dikategorikan menjadi 2-5 kelompok dengan label bermakna seperti 'Sangat Rendah', 'Rendah', 'Sedang', 'Tinggi', 'Sangat Tinggi'.

Pilih Variabel untuk Dikategorikan: Anakanak

Jumlah Kategori: 3

**Kategorikan Data** **Reset**

**Catatan:** Kolom baru akan ditambahkan dengan nama '[variabel]\_kategori\_[jumlah]' dengan label bermakna.

**Kategorisasi berhasil!** Variabel baru tersedia di menu lain.

### Visualisasi Kategorisasi

#### Distribusi Kategori: Buta\_Huruf

Kategori	Jumlah Wilayah
Rendah	28
Sedang	40
Tinggi	32

#### Ringkasan Kategorisasi

- Variabel: Buta\_Huruf
- Jumlah Kategori: 3
- Total Data: 100
- Detail per Kategori:
  - Rendah : 28 wilayah  
Range: 2.09 - 9.35 (rata-rata: 5.21)
  - Sedang : 40 wilayah  
Range: 9.88 - 16.84 (rata-rata: 13.31)
  - Tinggi : 32 wilayah  
Range: 17.37 - 24.45 (rata-rata: 21.33)

C:/Users/Maul/Downloads/KOMSTAT/UAS - Shiny  
http://127.0.0.1:6091 | Open in Browser | Republik

### Manajemen Variabel Kategorik

Variabel Kategorik yang Dibuat: Buta\_Huruf\_kategori\_3

**Petunjuk:** Klik tombol hapus (trash) untuk menghapus variabel kategorik yang tidak diperlukan.

### Hasil Kategorisasi dan Data

Kode_Distrik	Anakanak	Perempuan	Lansia	Kepala_RT_Perempuan	Ukuran_K
ID001	7.88	50.40	10.86	30.69	
ID002	12.88	49.33	19.55	15.19	
ID003	9.09	49.95	15.22	30.58	
ID004	13.83	51.82	14.18	29.59	
ID005	14.40	49.93	12.83	27.60	
ID006	5.46	51.56	18.56	24.62	
ID007	10.28	51.66	12.37	18.13	
ID008	13.92	50.43	11.46	15.16	
ID009	10.51	49.64	10.05	24.05	
ID010	9.57	48.59	10.07	24.85	

Showing 1 to 10 of 100 entries

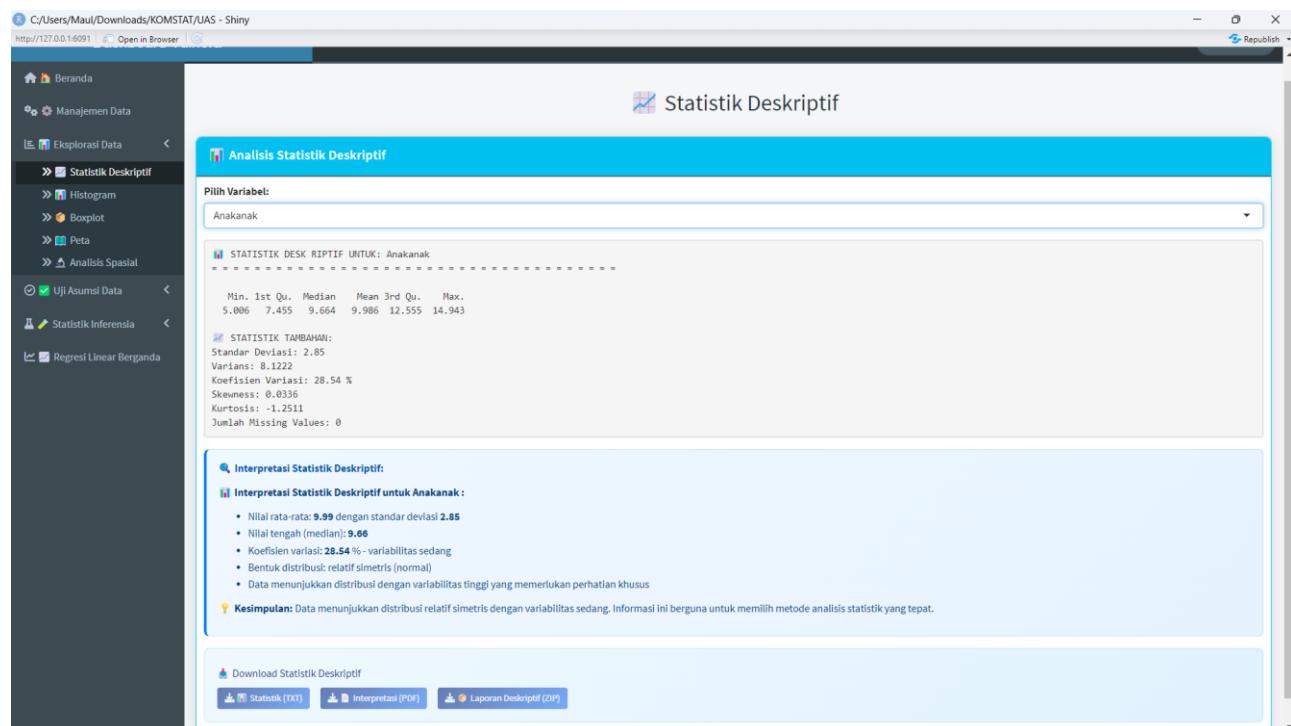
**Interpretasi Kategorisasi:**

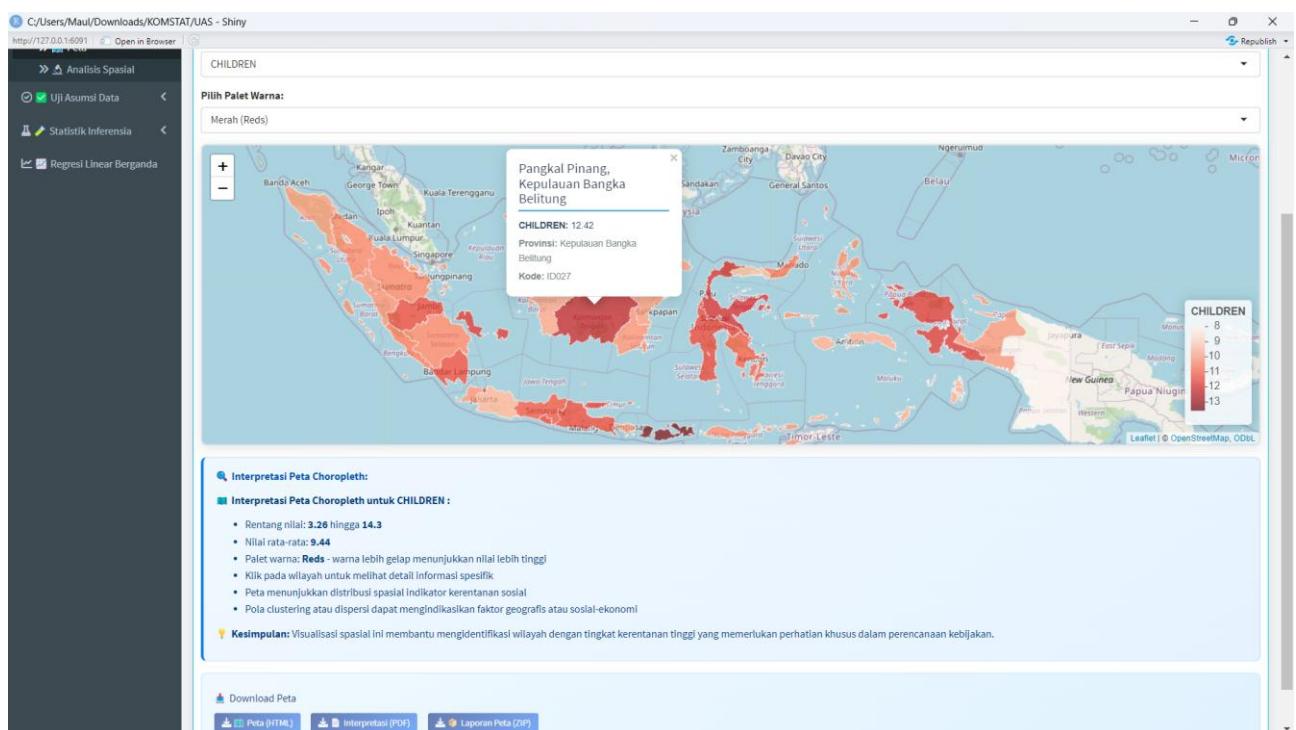
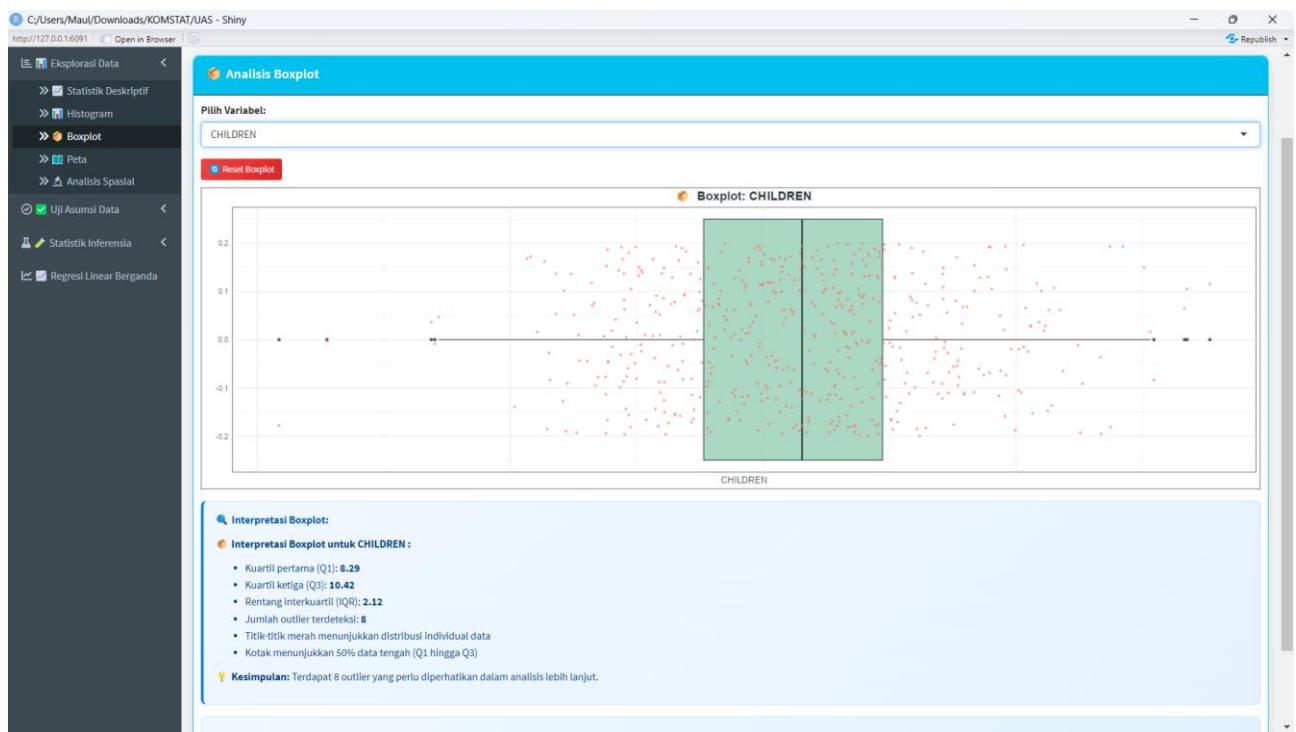
**Hasil Kategorisasi Variabel Buta\_Huruf :**

- Total data yang dikategorikan: **100** wilayah
- Kategori dengan frekuensi tertinggi: **Sedang** (40 wilayah, 40 %)
- Kategori dengan frekuensi terendah: **Rendah** (28 wilayah, 28 %)
- Distribusi data menunjukkan variasi kerentanan sosial antar wilayah
- Variabel kategorik baru dapat digunakan untuk analisis lebih lanjut seperti uji chi-square atau ANOVA

**Rekomendasi:** Gunakan variabel kategorik ini untuk analisis perbandingan antar kelompok atau visualisasi yang lebih mudah dipahami.

### c. Eksplorasi data





C:/Users/Maul/Downloads/KOMSTAT/UAS - Shiny  
http://127.0.0.1:6091 | Open in Browser | Republish | Dark Mode

## Dashboard Vulnера -

Beranda

Analisis Spasial

### Autokorelasi Spasial (Moran's I)

Pilih Indikator untuk Analisis:  
CHILDREN

Jalankan Analisis Spasial | Reset

#### Hasil Analisis Autokorelasi Spasial (Moran's I)

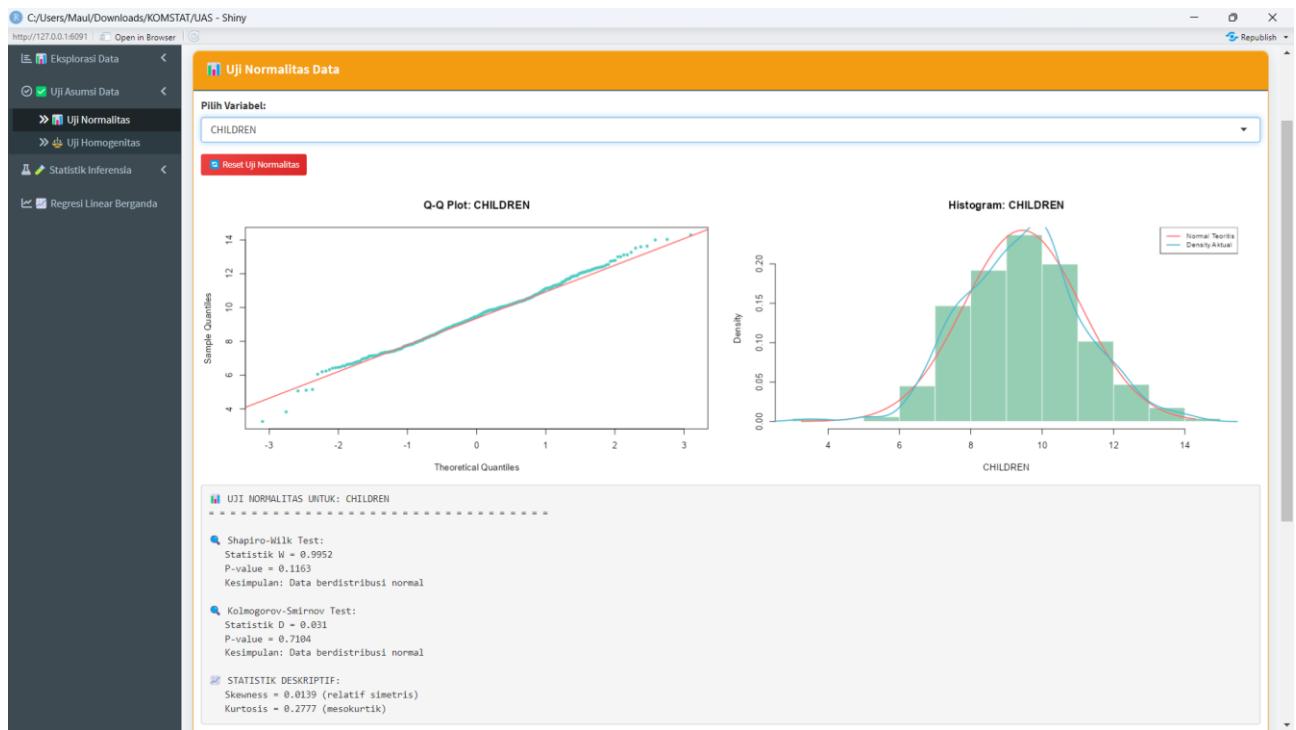
- Moran's I: -0.1357
- Expected I: -0.0323
- Variance: 0.01556
- P-value: 0.7964
- Signifikansi: Tidak Signifikan ( $\alpha = 0.05$ )

#### Interpretasi:

Tidak ada autokorelasi spasial yang signifikan. Distribusi nilai tampak acak secara spasial.

Download Analisis Spasial

Hasil Analisis (PDF) | Laporan Spasial (ZIP)



**Interpretasi Uji Normalitas:**

**Interpretasi Uji Normalitas untuk CHILDREN :**

- Shapiro-Wilk Test: Data berdistribusi normal ( $p > 0.05$ )
- Kolmogorov-Smirnov Test: Data berdistribusi normal ( $p > 0.05$ )
- Q-Q Plot: Titik-titik yang mengikuti garis diagonal menunjukkan distribusi normal
- Histogram: Bentuk lonceng simetris mengindikasikan distribusi normal
- Skewness: 0.014 - distribusi relatif simetris

**Kesimpulan:** Data berdistribusi normal. Dapat menggunakan uji parametrik seperti t-test atau ANOVA.

**Rekomendasi:** Lanjutkan dengan analisis parametrik.

**Download Uji Normalitas**

[Download Q-Q Plot \(JPG\)](#) [Hasil Uji \(TXT\)](#) [Interpretasi \(PDF\)](#) [Laporan Normalitas \(ZIP\)](#)

C:/Users/Maul/Downloads/KOMSTAT/UAS - Shiny  
http://127.0.0.1:6091 | Open in Browser | Republish

## Uji Homogenitas Varians

**Variabel Grup:** CHILDREN\_kategori\_3

**Reset Uji Homogenitas**

**UJI HOMOGENITAS VARIANS**

Variabel yang diujii: DISTRICTCODE berdasarkan grup CHILDREN\_kategori\_3

- Levene's Test: varians tidak homogen ( $p = 0.0086$ )
- Bartlett's Test: varians tidak homogen ( $p = 0.0416$ )

**Statistik per Grup:**

- Rendah: n = 25, var = NA, sd = NA
- Sedang: n = 377, var = NA, sd = NA
- Tinggi: n = 109, var = NA, sd = NA

**Interpretasi Uji Homogenitas:**

**Interpretasi Uji Homogenitas Varians untuk CHILDREN\_kategori\_3:**

- Levene's Test: Asumsi homogenitas varians tidak terpenuhi ( $p \leq 0.05$ )
- Bartlett's Test: Asumsi homogenitas varians tidak terpenuhi ( $p \leq 0.05$ )
- Variabel yang diujii: DISTRICTCODE berdasarkan grup CHILDREN\_kategori\_3
- Jumlah grup: 3
- Total observasi: 511

**Kesimpulan:** Asumsi homogenitas varians tidak terpenuhi. Dapat melanjutkan dengan ANOVA atau t-test dengan asumsi varians sama.

**Catatan:** Levene's test lebih robust terhadap penyimpangan normalitas dibandingkan Bartlett's test.

**Download Uji Homogenitas**

**Hasil Uji (TXT)** **Interpretasi (PDF)** **Laporan Homogenitas (ZIP)**

## d. Statistik Inferensia

Dashboard Vulnera - Dark Mode

**1 Uji T Satu Sampel**

Nilai Hipotesis ( $\mu_0$ ): 0.05

Pilih Variabel: CHILDREN

**Jalankan Uji** **Reset**

**UJI T SATU SAMPEL**

H<sub>0</sub>:  $\mu = 0.05$   
H<sub>1</sub>:  $\mu \neq 0.05$

**HASIL UJI:**  
t-statistic = 129.1824  
df = 510  
p-value = 0  
Confidence Interval (95%): 9.302 to 9.5878  
Sample mean = 9.4449

**KESIMPULAN:**  
 Tolak H<sub>0</sub>. Rata-rata populasi berbeda signifikan dari 0.05 ( $\alpha = 0.05$ )

**2 Uji T Dua Sampel**

**Petunjuk:** Pilih variabel numerik dan pilih dua wilayah untuk dibandingkan.

Variabel Nilai (Numerik): CHILDREN

Wilayah 1: Kalimantan

Wilayah 2: Sulawesi

Asumsikan Varians Sama

**Jalankan Uji** **Reset**

**UJI T DUA SAMPEL**

H<sub>0</sub>:  $\mu_1 = \mu_2$  (rata-rata kedua wilayah sama)  
H<sub>1</sub>:  $\mu_1 \neq \mu_2$  (rata-rata kedua wilayah berbeda)

**HASIL UJI:**

```
df = 510
p-value = 0
Confidence Interval (95%): 9.302 to 9.5878
Sample mean = 9.4449

✓ KESIMPULAN:
Tolak H0. Rata-rata populasi berbeda signifikan dari 0.05 ( $\alpha = 0.05$ )
```

**🔍 Interpretasi Uji T Satu Sampel:**

**Interpretasi Uji T Satu Sampel:**

- Hipotesis nol ( $H_0$ ): Rata-rata populasi = **0.05**
- Hipotesis alternatif ( $H_1$ ): Rata-rata populasi ≠ **0.05**
- Statistik t: **129.1824**
- Derajat bebas: **510**
- P-value: **0**
- Rata-rata sampel: **9.4449**
- Interval kepercayaan 95%: [9.302, 9.5878]

💡 **Kesimpulan:** Dengan tingkat signifikansi 5%, terdapat bukti yang cukup untuk menyimpulkan bahwa rata-rata populasi berbeda dari 0.05. Perbedaan ini signifikan secara statistik.

⬇️ Download Uji T Satu Sampel

[Hasil Uji \(TXT\)](#) [Interpretasi \(PDF\)](#) [Laporan T-Test 1 Sampel \(ZIP\)](#)

```
UJI T DUA SAMPEL
-----
H0:  $\mu_1 = \mu_2$  (rata-rata kedua wilayah sama)
H1:  $\mu_1 \neq \mu_2$  (rata-rata kedua wilayah berbeda)

❗ HASIL UJI:
Wilayah 1 ( Kalimantan ): n = 110 , mean = 9.7644
Wilayah 2 ( Sulawesi ): n = 102 , mean = 9.2737
t-statistic = 2.0935
df = 210
p-value = 0.0375
Confidence Interval (95%): 0.0286 to 0.9529
Asumsi varians sama: Ya

✓ KESIMPULAN:
Tolak H0. Terdapat perbedaan signifikan antara rata-rata Kalimantan dan Sulawesi ( $\alpha = 0.05$ )
```

**🔍 Interpretasi Uji T Dua Sampel:**

**Interpretasi Uji T Dua Sampel:**

- Variabel yang dibandingkan: **CHILDREN**
- Wilayah 1: **Kalimantan** (n = 110, mean = 9.7644)
- Wilayah 2: **Sulawesi** (n = 102, mean = 9.2737)
- Perbedaan rata-rata: **0.4908**
- Statistik t: **2.0935**
- P-value: **0.0375**
- Asumsi varians sama: **Ya (pooled variance)**

💡 **Kesimpulan:** Terdapat perbedaan yang signifikan secara statistik antara rata-rata CHILDREN di Kalimantan dan Sulawesi. Perbedaan ini tidak disebabkan oleh kebetulan sampling.

💡 **Catatan:** Menggunakan pooled variance (asumsi varians sama). Pastikan asumsi homogenitas varians terpenuhi.

⬇️ Download Uji T Dua Sampel

[Hasil Uji \(TXT\)](#) [Interpretasi \(PDF\)](#) [Laporan T-Test 2 Sampel \(ZIP\)](#)

**🔍 Interpretasi Uji Proporsi Satu Sampel:**

**Interpretasi Uji Proporsi Satu Sampel (Dari Variabel):**

- Variabel: **CHILDREN\_kategori\_3**
- Level sukses: **Tinggi**
- Proporsi sampel: **0.2133**
- Proporsi hipotesis: **0.1**
- Ukuran sampel: **511**

💡 **Interpretasi:** Uji ini mengevaluasi apakah proporsi kategori 'Tinggi' dalam variabel 'CHILDREN\_kategori\_3' berbeda secara signifikan dari proporsi yang dihipotesiskan.

⬇️ Download Uji Proporsi Satu Sampel

[Hasil Uji \(TXT\)](#) [Interpretasi \(PDF\)](#) [Laporan Proporsi 1 Sampel \(ZIP\)](#)

**🔍 Interpretasi Uji Proporsi Dua Sampel:**

**Interpretasi Uji Proporsi Dua Sampel:**

- Variabel: **CHILDREN\_kategori\_3**
- Level sukses: **Tinggi**
- Wilayah 1 ( Kalimantan ): **0.3091** (34/110)
- Wilayah 2 ( Sulawesi ): **0.1961** (20/102)
- Perbedaan proporsi: **0.113**

💡 **Interpretasi:** Uji ini membandingkan proporsi kategori 'Tinggi' antara dua wilayah untuk menentukan apakah perbedaan yang diamati signifikan secara statistik atau hanya disebabkan oleh variasi sampling.

⬇️ Download Uji Proporsi Dua Sampel

[Hasil Uji \(TXT\)](#) [Interpretasi \(PDF\)](#) [Laporan Proporsi 2 Sampel \(ZIP\)](#)

C:/Users/Maul/Downloads/KOMSTAT/UAS - Shiny  
http://127.0.0.1:6091 | Open in Browser | Republish

**Uji Varians**

**1 Uji Varians Satu Sampel**

**Fitur Baru:** Uji varians satu sampel untuk menguji apakah varians populasi sama dengan nilai tertentu.

Pilih Variabel: CHILDREN

Varians Hipotesis ( $\sigma^2$ ): 0.1

Jalankan Uji Reset

**UJI VARIANS SATU SAMPEL**

H<sub>0</sub>:  $\sigma^2 = 0.1$   
H<sub>1</sub>:  $\sigma^2 \neq 0.1$

**HASIL UJI:**  
Sample size (n) = 511  
Sample variance = 2.7027  
Chi-square statistic = 13783.79  
df = 510  
p-value = 0

**KESIMPULAN:**  
Tolak H<sub>0</sub>. Varians populasi berbeda signifikan dari 0.1 ( $\alpha = 0.05$ )

**Interpretasi Uji Varians Satu Sampel:**

**Interpretasi Uji Varians Satu Sampel:**

- Variabel: CHILDREN
- Varians sampel: **2.7027**
- Varians hipotesis: **0.1**
- Ukuran sampel: **511**

**Interpretasi:** Uji ini menggunakan distribusi chi-square untuk mengevaluasi apakah varians populasi berbeda dari nilai yang dihipotesiskan. Uji ini sensitif terhadap asumsi normalitas data.

Download Uji Varians Satu Sampel

Hasil Uji (TXT) Interpretasi (PDF) Laporan Varians 1 Sampel (ZIP)

**2 Uji Varians Dua Sampel**

**Petunjuk:** Pilih variabel numerik dan pilih dua wilayah untuk membandingkan varians.

Variabel Nilai (Numerik): CHILDREN

Wilayah 1: Kalimantan

Wilayah 2: Sulawesi

Jalankan Uji Reset

**UJI VARIANS DUA SAMPEL (F-TEST)**

H<sub>0</sub>:  $\sigma_1^2 = \sigma_2^2$  (varians kedua wilayah sama)  
H<sub>1</sub>:  $\sigma_1^2 \neq \sigma_2^2$  (varians kedua wilayah berbeda)

**HASIL UJI:**  
Wilayah 1 ( Kalimantan ): n = 110 , var = 3.0405  
Wilayah 2 ( Sulawesi ): n = 102 , var = 2.766  
F-statistic = 1.0992  
df1 = 109 , df2 = 101  
p-value = 0.631  
Confidence Interval (95%): 0.747 to 1.613

**KESIMPULAN:**  
Gagal tolak H<sub>0</sub>. Tidak ada perbedaan signifikan varians antara Kalimantan dan Sulawesi ( $\alpha = 0.05$ )

Confidence Interval (95%): 0.747 to 1.613

**Interpretasi Uji Varians Dua Sampel:**

**Interpretasi Uji Varians Dua Sampel (F-Test):**

- Variabel: CHILDREN
- Wilayah 1 ( Kalimantan ): varians = **3.0405**
- Wilayah 2 ( Sulawesi ): varians = **2.766**
- Rasio varians (F): **1.0992**

**Interpretasi:** F-test membandingkan varians dua kelompok. Uji ini sangat sensitif terhadap asumsi normalitas dan sering digunakan sebagai uji pendahuluan untuk menentukan apakah menggunakan t-test dengan asumsi varians sama atau tidak.

**Catatan:** Jika varians berbeda signifikan, gunakan Welch's t-test untuk perbandingan rata-rata.

Download Uji Varians Dua Sampel

Hasil Uji (TXT) Interpretasi (PDF) Laporan Varians 2 Sampel (ZIP)

## e. Regresi Linier Berganda

C:/Users/Maul/Downloads/KOMSTAT/UAS - Shiny  
http://127.0.0.1:6091 | Open in Browser | Republish

**Beranda**

- Manajemen Data
- Eksplorasi Data
- Asumsi Data**
- Statistik Inferensia
  - Uji Beda Rata-rata
  - Uji Proporsi
  - Uji Varians
  - ANOVA
- Regresi Linear Berganda

**Model Regresi**

Variabel Respon (Y): CHILDREN

Variabel Prediktor (X): POPULATION Luas

Jalankan Regresi | Reset

**REGRESI LINEAR BERGANDA**

Formula: CHILDREN ~ POPULATION + Luas

```
Call:
lm(formula = as.formula(formula_str), data = data)

Residuals:
    Min      1Q  Median      3Q     Max 
-6.4146 -1.0982  0.0882  0.9827  4.5750 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 9.739e+00 1.601e-01 60.835 < 2e-16 ***
POPULATION -6.405e-07 1.123e-07 -5.701 2.02e-08 ***
Luas        1.351e-05 5.049e-05 0.268 0.789    
---
Signif. codes:  0 '****' 0.001 '**' 0.01 '*' 0.1 ' ' 1

Residual standard error: 1.596 on 508 degrees of freedom
Multiple R-squared:  0.06073, Adjusted R-squared:  0.05703 
F-statistic: 16.42 on 2 and 508 DF,  p-value: 1.228e-07
```

Interpretasi Model Regresi:

**Interpretasi Model Regresi:**

**Interpretasi Model Regresi Linear Berganda:**

- Variabel dependen: CHILDREN
- Variabel Independen: POPULATION, Luas
- R-squared: **0.0607** (6.07 % varians dijelaskan)
- Adjusted R-squared: **0.057**
- F-statistik: **16.422**
- P-value model: **0**

**Interpretasi Koefisien:** Setiap koefisien menunjukkan perubahan rata-rata dalam variabel dependen untuk setiap unit perubahan dalam variabel independen, dengan variabel lain konstan.

**Signifikansi Model:** Model secara keseluruhan signifikan ( $p < 0.05$ ). Minimal satu prediktor memiliki hubungan linear dengan variabel dependen.

**Kualitas Model:** Lemah (<30% varians dijelaskan)

**Catatan Penting tentang Asumsi Regresi Linear**

**Asumsi yang Harus Dipenuhi:**

- Linearitas: Hubungan antara variabel independen dan dependen harus linear
- Independensi: Observasi harus saling Independen
- Homoskedastisitas: Varians residual harus konstan
- Normalitas: Residu harus berdistribusi normal
- Tidak ada multikolinearitas: Variabel prediktor tidak boleh berkorelasi tinggi

**Jika Asumsi Tidak Terpenuhi:**

- Lakukan transformasi data (log, sqrt, dll.)
- Gunakan metode regresi robust
- Tambah atau kurangi variabel prediktor
- Pertimbangkan model non-linear
- Gunakan teknik regularisasi (Ridge, Lasso)

**Download Model Regresi**

Hasil Regresi (TXT) | Interpretasi (PDF) | Laporan Regresi (ZIP)

**Residuals vs Fitted**

Normal Q-Q Plot

VARIANCE INFLATION FACTOR (VIF)

VIF > 10: Multikolinearitas tinggi  
VIF > 5: Multikolinearitas sedang  
VIF < 5: Multikolinearitas rendah

POPULATION Luas

INTERPRETASI VIF:  
POPULATION: 1 - RENDAH (tidak bermasalah)  
Luas: 1 - RENDAH (tidak bermasalah)

PUMULAJUN: 1 - RENDAH (tidak bermasalah)  
Luas: 1 - RENDAH (tidak bermasalah)

**Interpretasi Diagnostik:**

**Interpretasi Plot Diagnostik:**

- Residuals vs Fitted:** Pola acak menunjukkan linearitas dan homoskedastisitas terpenuhi. Pola kurva atau corong menunjukkan pelanggaran asumsi.
- Normal Q-Q Plot:** Titik-titik mengikuti garis diagonal menunjukkan residual berdistribusi normal. Penyimpangan dari garis menunjukkan non-normalitas.
- VIF (Variance Inflation Factor):** Mengukur multikolinearitas. VIF > 10 menunjukkan masalah serius, VIF > 5 perlu perhatian.

**Peringatan:** Jika asumsi tidak terpenuhi, pertimbangkan transformasi data, penambahan/pengurangan variabel, atau metode regresi alternatif.

**Solusi Pelanggaran Asumsi:**

- Non-linearitas: Transformasi variabel atau model polinomial
- Heteroskedastisitas: Transformasi log atau weighted least squares
- Non-normalitas residual: Transformasi Box-Cox atau robust regression
- Multikolinearitas: Ridge regression, Lasso, atau hapus variabel berkorelasi tinggi

**Download Diagnostik**

Plot Diagnostik (JPG) | VIF (TXT) | Interpretasi (PDF) | Laporan Diagnostik (ZIP)

## BAB IV

### KESIMPULAN

#### **Hasil Akhir Proyek**

*Dashboard R Shiny* yang dikembangkan, "PDKB (Pusat Data Kesiapsiagaan Bencana)," berhasil mengintegrasikan berbagai sumber data terkait kerentanan sosial di Indonesia, menyediakan *platform* komprehensif untuk analisis dan visualisasi. Fitur-fitur utama

*dashboard* ini, sebagaimana didemonstrasikan pada bagian sebelumnya, meliputi:

- Beranda: Menyediakan pengantar misi *dashboard*, fitur utama, dan informasi data. Ini juga menawarkan opsi unduh komprehensif untuk laporan dan data mentah.
- Manajemen Data: Memungkinkan pengguna untuk mengkategorikan variabel numerik, memvisualisasikan distribusinya, dan mengunduh hasil kategorisasi, serta memberikan rekomendasi strategis berdasarkan kategorisasi.
- Eksplorasi Data: Fitur peta *koroplet* interaktif untuk memvisualisasikan distribusi spasial kerentanan, statistik deskriptif, histogram, *boxplot*, dan analisis otokorelasi spasial (*Moran's I*) untuk menyoroti indikator kerentanan kunci. Semua *output* eksplorasi dapat diunduh.
- Uji Asumsi: Menyediakan alat untuk uji normalitas (Shapiro-Wilk dan Kolmogorov-Smirnov) dan homogenitas (Levene dan Bartlett) dengan interpretasi yang jelas dan laporan yang dapat diunduh, memastikan validitas analisis inferensia berikutnya.
- Statistik Inferensia: Menggabungkan Uji-t (satu sampel dan dua sampel), Uji Proporsi (satu sampel dan dua sampel), Uji Varians (satu sampel dan dua sampel Uji F), dan ANOVA dengan Uji Lanjutan *Tukey HSD* untuk memungkinkan pengujian hipotesis dan perbandingan indikator kerentanan di berbagai wilayah, lengkap dengan interpretasi dan implikasi kebijakan. Semua hasil dapat diunduh.
- Regresi Linear Berganda: Memungkinkan pengguna untuk membangun dan memvalidasi model regresi linear untuk memahami hubungan antara faktor-faktor kerentanan sosial yang dipilih dan variabel dependen. Ini mencakup ringkasan model, *plot* diagnostik, dan uji multikolinearitas (VIF), dengan laporan komprehensif yang dapat diunduh.

*Dashboard* ini berhasil mengatasi tantangan awal dalam mengintegrasikan data yang tersebar ke dalam satu *platform* yang interaktif, menyediakan bukti empiris untuk menginformasikan kebijakan mitigasi bencana. Desainnya yang ramah pengguna, interpretasi otomatis, dan fitur unduh yang komprehensif meningkatkan aksesibilitas bagi berbagai *stakeholder*.