

**ANALISIS KERENTANAN SOSIAL WILAYAH DI INDONESIA MENGGUNAKAN  
DASHBOARD BUSINESS INTELLIGENCE BERBASIS R SHINY**



Dosen Pengampu:

Yuliagnis T. Wijaya, SST, M.SC

Disusun Oleh:

Nama : Alifia Deanita

NIM : 222312960

Kelas : 2KS3

**PROGRAM STUDI D-IV KOMPUTASI STATISTIK**

**POLITEKNIK STATISTIKA STIS**

**2024/2025**

## KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan laporan proyek ini yang berjudul *"Analisis Kerentanan Sosial Wilayah di Indonesia Menggunakan Dashboard Business Intelligence Berbasis R Shiny"*. Laporan ini disusun sebagai bagian dari tugas akhir mata kuliah Komputasi Statistik.

Pembuatan dashboard interaktif ini bertujuan untuk membantu proses eksplorasi dan analisis data kerentanan sosial secara visual dan statistik. Dengan menggunakan tools seperti R dan R Shiny, diharapkan pengguna dapat memperoleh insight yang informatif dari data, serta menjadikannya sebagai dasar pengambilan keputusan berbasis data (data-driven).

Penulis menyadari bahwa laporan ini tidak lepas dari kekurangan. Oleh karena itu, penulis sangat terbuka terhadap kritik dan saran yang membangun demi perbaikan ke depannya. Ucapan terima kasih disampaikan kepada dosen pengampu mata kuliah, teman-teman seperjuangan, serta pihak-pihak yang telah membantu dalam proses penyusunan laporan ini.

Semoga laporan ini dapat memberikan manfaat, baik untuk penulis sendiri maupun untuk pihak lain yang membutuhkannya.

Jakarta, 23 Juli 2025



Alifia Deanita

NIM. 222312960

## DAFTAR ISI

<b>KATA PENGANTAR.....</b>	<b>ii</b>
<b>DAFTAR ISI.....</b>	<b>iii</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1    Latar Belakang .....	1
1.2    Rumusan Masalah.....	1
1.3    Tujuan .....	2
1.4    Manfaat .....	2
<b>BAB II TINJAUAN PUSTAKA.....</b>	<b>3</b>
2.1    Kerentanan Sosial Wilayah .....	3
2.2    Analisis Spasial dan Statistik Inferensia .....	3
2.3    Klasterisasi Wilayah Berbasis Matriks Penimbang .....	4
2.4    Dashboard Business Intelligence Berbasis R Shiny .....	4
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>6</b>
3.1    Desain Penelitian .....	6
3.2    Data dan Sumber Data .....	6
3.3    Variabel Penelitian .....	7
3.4    Teknik Analisis.....	7
3.4.1    Analisis Deskriptif dan Visualisasi Spasial.....	7
3.4.2    Uji Asumsi Statistik .....	8
3.4.3    Analisis Statistik Inferensia .....	8
3.4.4    Regresi Linier Berganda .....	8
3.4.5    Klasterisasi Spasial .....	9
3.5    Perangkat dan Tools .....	9

<b>BAB IV HASIL DAN PEMBAHASAN.....</b>	<b>10</b>
4.1    Tab Beranda .....	10
4.2    Tab Manajemen Data .....	11
4.3    Tab Eksplorasi Data .....	11
4.4    Tab Uji Asumsi Data .....	13
4.5    Tab Statistik Inferensia .....	15
4.6    Tab Regresi Linier Berganda .....	17
<b>BAB V PENUTUP.....</b>	<b>20</b>
5.1    Kesimpulan .....	20
5.2    Saran .....	21
<b>DAFTAR PUSTAKA .....</b>	<b>22</b>
<b>LAMPIRAN.....</b>	<b>23</b>

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Permasalahan sosial seperti kemiskinan, ketimpangan pendidikan, dan keterbatasan akses fasilitas dasar menjadi isu penting dalam pembangunan di Indonesia. Dalam konteks kebijakan pembangunan yang inklusif, diperlukan alat bantu yang mampu menggambarkan kondisi sosial suatu wilayah secara menyeluruh dan berbasis data. Salah satu pendekatan yang dapat digunakan adalah pengukuran kerentanan sosial wilayah melalui indeks atau skor komposit.

Dalam era digital, pemanfaatan teknologi analisis data dan visualisasi interaktif sangat diperlukan untuk menunjang pemahaman yang lebih baik terhadap data sosial. Business Intelligence (BI) menjadi solusi yang memungkinkan pengambilan keputusan berbasis data yang lebih cepat dan efektif. Melalui dashboard interaktif yang dibangun dengan R Shiny, pengguna dapat melakukan eksplorasi, analisis statistik, hingga pemodelan secara dinamis terhadap data kerentanan sosial.

Dashboard ini tidak hanya menyajikan visualisasi peta dan grafik, tetapi juga menyertakan fitur-fitur statistik inferensial, pengujian asumsi, hingga analisis regresi dan klusterisasi. Diharapkan, aplikasi ini mampu menjadi sarana analisis sosial spasial yang informatif dan dapat digunakan oleh pemangku kebijakan, peneliti, maupun masyarakat umum.

### **1.2 Rumusan Masalah**

1. Bagaimana memvisualisasikan sebaran kerentanan sosial wilayah di Indonesia secara interaktif?
2. Bagaimana melakukan eksplorasi dan uji asumsi statistik terhadap data kerentanan sosial?
3. Bagaimana pengaruh beberapa faktor sosial terhadap tingkat kerentanan sosial menggunakan analisis regresi?
4. Bagaimana mengelompokkan wilayah berdasarkan karakteristik sosialnya menggunakan analisis kluster?

5. Bagaimana dashboard ini dapat berfungsi sebagai alat Business Intelligence dalam pengambilan keputusan?

### **1.3 Tujuan**

1. Mengembangkan dashboard interaktif berbasis R Shiny untuk analisis kerentanan sosial.
2. Menyajikan visualisasi data dalam bentuk peta dan grafik yang informatif.
3. Melakukan analisis statistik dan pengujian asumsi untuk memahami karakteristik data.
4. Membangun model regresi linier berganda untuk mengetahui faktor-faktor yang berpengaruh terhadap kerentanan sosial.
5. Mengelompokkan wilayah berdasarkan kesamaan karakteristik sosial menggunakan teknik klasterisasi.
6. Menyediakan alat bantu Business Intelligence yang dapat mendukung pengambilan keputusan berbasis data.

### **1.4 Manfaat**

- Bagi akademisi: Memberikan contoh nyata penerapan analisis statistik dan visualisasi data spasial dalam satu kesatuan dashboard.
- Bagi pemerintah dan pemangku kebijakan: Memberikan insight terhadap wilayah yang rentan secara sosial dan mempermudah perencanaan intervensi.
- Bagi masyarakat umum: Menyediakan informasi visual yang mudah dipahami tentang kondisi sosial wilayah di Indonesia.
- Bagi pengembang: Melatih kemampuan dalam membangun aplikasi analitik menggunakan bahasa R dan R Shiny.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Kerentanan Sosial Wilayah**

Kerentanan sosial wilayah merupakan konsep multidimensional yang mengacu pada sejauh mana suatu komunitas atau wilayah memiliki kapasitas untuk menghadapi dan pulih dari tekanan atau bencana sosial, ekonomi, maupun lingkungan. Kerentanan ini tidak hanya dipengaruhi oleh satu faktor, melainkan oleh kombinasi karakteristik sosial seperti kemiskinan, pendidikan, jenis kelamin kepala keluarga, kepadatan penduduk, tingkat pengangguran, hingga akses terhadap layanan dasar.

Menurut Sudirman, Ambarwati, dan Lubis (2022), kerentanan sosial merupakan suatu kondisi yang dapat menyebabkan kelompok masyarakat tertentu sulit untuk bertahan atau pulih dari tekanan sosial, ekonomi, dan bencana karena terbatasnya sumber daya, kapasitas, dan posisi sosial dalam struktur masyarakat.

SoVI digunakan dalam berbagai konteks kebijakan publik untuk mengidentifikasi wilayah prioritas intervensi, terutama dalam perencanaan pembangunan berkelanjutan dan mitigasi bencana. Dengan pendekatan ini, data mikro (kabupaten/kota) dapat dianalisis secara komparatif dan visual melalui peta spasial.

#### **2.2 Analisis Spasial dan Statistik Inferensia**

Analisis spasial merupakan pendekatan yang digunakan untuk memahami fenomena sosial dengan mempertimbangkan dimensi lokasi atau geografi. Dalam konteks kerentanan sosial, analisis spasial memegang peranan penting untuk mengetahui persebaran dan pola ketimpangan sosial antardaerah.

Menurut Yulia (2022) dalam bukunya Navigasi Penelitian Sosial, pendekatan spasial tidak hanya berfokus pada penggambaran data dalam bentuk peta, tetapi juga memungkinkan peneliti untuk mengidentifikasi hubungan antarwilayah berdasarkan kedekatan geografis dan karakteristik sosial yang serupa. Penelitian sosial yang memanfaatkan spasial, lanjutnya, membuka kemungkinan lahirnya kebijakan yang

lebih tepat sasaran karena memperhatikan konteks lokal dan pola keterhubungan antarwilayah.

Di sisi lain, statistik sosial berfungsi untuk mengkuantifikasi fenomena sosial seperti kemiskinan, ketimpangan, pendidikan, kesehatan, dan lainnya. Dengan memadukan statistik sosial dan spasial, maka dapat dilakukan pemetaan indikator sosial yang lebih representatif dan akurat.

Penggunaan pendekatan ini sangat relevan dalam mengkaji kerentanan sosial wilayah karena indikator-indikator sosial yang digunakan tidak bersifat tunggal, tetapi saling berinteraksi dan memiliki persebaran yang tidak merata secara geografis. Oleh karena itu, integrasi metode statistik sosial dan analisis spasial dalam penelitian ini menjadi dasar dalam penyusunan indeks kerentanan sosial wilayah.

### **2.3 Klasterisasi Wilayah Berbasis Matriks Penimbang**

Klasterisasi merupakan teknik pengelompokan wilayah berdasarkan kemiripan karakteristik, sehingga setiap klaster berisi wilayah-wilayah dengan profil sosial yang mirip. Dalam konteks spasial, digunakan matriks penimbang yang mencerminkan jarak antar wilayah sebagai dasar klasterisasi. Salah satu metode yang umum digunakan adalah Hierarchical Clustering dengan pendekatan Ward's Method yang meminimalkan varians intra-klaster.

Dengan menggabungkan variabel spasial dan sosial, klasterisasi dalam studi ini bertujuan untuk menyederhanakan kompleksitas wilayah menjadi beberapa kelompok prioritas yang bisa ditindaklanjuti secara lebih fokus oleh pembuat kebijakan.

### **2.4 Dashboard Business Intelligence Berbasis R Shiny**

Dashboard Business Intelligence (BI) adalah alat visualisasi interaktif yang memungkinkan pengguna untuk menelusuri, menganalisis, dan mengambil keputusan berdasarkan data. Dalam konteks sosial spasial, dashboard memudahkan eksplorasi kerentanan sosial melalui peta, grafik, uji statistik, dan interpretasi otomatis.



Penggunaan R Shiny sebagai platform dashboard memberikan fleksibilitas tinggi dalam pengembangan antarmuka pengguna yang interaktif dan kaya fitur, serta integrasi langsung dengan paket analisis statistik dan spasial dalam R seperti sf, leaflet, dplyr, dan ggplot2.

Dashboard dalam penelitian ini mencakup fitur:

- Pemilihan variabel untuk eksplorasi peta,
- Uji asumsi distribusi data,
- Statistik inferensia untuk analisis perbandingan,
- Regresi linear untuk identifikasi faktor dominan,
- Klasterisasi spasial untuk prioritas intervensi.

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Desain Penelitian**

Penelitian ini menggunakan pendekatan kuantitatif dengan jenis penelitian deskriptif eksploratif, yang bertujuan untuk menggambarkan, menganalisis, dan menginterpretasikan tingkat kerentanan sosial wilayah kabupaten/kota di Indonesia. Kerentanan sosial ini dianalisis menggunakan indikator-indikator sosial yang relevan, dan hasilnya disajikan dalam bentuk dashboard interaktif berbasis R Shiny, guna menunjang pengambilan keputusan berbasis data.

Penelitian ini juga mengintegrasikan teknik analisis spasial dan statistik inferensial dalam satu sistem visualisasi, sehingga pengguna tidak hanya dapat melihat kondisi suatu wilayah secara visual, tetapi juga memperoleh wawasan mendalam dari analisis statistik yang mendasarinya. Desain ini memungkinkan proses analisis yang menyeluruh mulai dari eksplorasi data, uji asumsi, pengujian hipotesis, hingga pemodelan regresi dan segmentasi spasial wilayah menggunakan metode kluster.

#### **3.2 Data dan Sumber Data**

Penelitian ini menggunakan tiga jenis data utama, yaitu:

##### **1. Data Sosial Wilayah (SoVI – Social Vulnerability Index)**

Merupakan data utama yang berisi indikator-indikator sosial untuk masing-masing kabupaten/kota di Indonesia. Data ini dikompilasi dalam file `sovi_data.csv`, dan mencakup berbagai variabel seperti:

- Persentase anak-anak (CHILDREN)
- Perempuan (FEMALE)
- Lansia (ELDERLY)
- Kepala keluarga perempuan (FHEAD)
- Ukuran keluarga (FAMILYSIZE)
- Tidak memiliki listrik, pendidikan rendah, tidak dilatih, tinggal di rumah sewa. Tidak memiliki akses air bersih, dan lainnya.

## 2. Data Spasial Wilayah Administratif

Digunakan untuk visualisasi peta, data ini berasal dari file indonesia.geojson yang berisi batas administratif kabupaten/kota di Indonesia. Data ini memungkinkan analisis spasial dan pemetaan nilai indikator maupun hasil klasterisasi.

## 3. Matriks Jarak Antar Wilayah (Distance Matrix)

Berasal dari file distance.csv, data ini merepresentasikan kedekatan spasial antar wilayah dalam bentuk matriks jarak simetris. Matriks ini sangat penting untuk proses clustering spasial, karena mempertimbangkan kemiripan antar wilayah berdasarkan jarak geografis.

### 3.3 Variabel Penelitian

Variabel-variabel dalam penelitian ini terdiri atas:

- Variabel numerik: digunakan untuk perhitungan indeks dan analisis regresi, misalnya: POVERTY, FEMALE, LOWEDU, ILLITERATE, dan sebagainya.
- Variabel kategorik biner: dibentuk berdasarkan median dari variabel numerik, digunakan untuk keperluan uji beda rata-rata dua kelompok, uji proporsi, dan ANOVA.
- Variabel spasial: berupa batas wilayah administratif dalam data sf untuk pemetaan dan klasterisasi.

### 3.4 Teknik Analisis

Penelitian ini menggunakan beberapa teknik analisis sebagai berikut:

#### 3.4.1 Analisis Deskriptif dan Visualisasi Spasial

Visualisasi deskriptif dilakukan melalui peta tematik interaktif (choropleth map) menggunakan Leaflet. Pengguna dapat memilih satu variabel numerik untuk ditampilkan di peta, dan sistem akan memberikan interpretasi otomatis mengenai wilayah dengan nilai tertinggi dan terendah, serta rentang nilai variabel tersebut. Ini membantu mengidentifikasi sebaran kerentanan secara cepat.

### **3.4.2 Uji Asumsi Statistik**

Sebelum dilakukan analisis inferensial maupun regresi, dilakukan uji asumsi sebagai berikut:

- Normalitas (Shapiro-Wilk): menguji apakah distribusi residual maupun variabel target mengikuti distribusi normal.
- Homogenitas Varians (Levene Test): menguji apakah varians antar kelompok sama.
- Multikolinieritas (VIF): mengidentifikasi adanya hubungan antar prediktor dalam model regresi.
- Autokorelasi (Durbin-Watson): menguji apakah terdapat pola korelasi dalam residual model regresi.

### **3.4.3 Analisis Statistik Inferensia**

Analisis inferensial digunakan untuk menguji perbedaan antara kelompok-kelompok berdasarkan indikator kerentanan sosial:

- Uji Beda Rata-rata Dua Kelompok (t-test): digunakan jika variabel kategorik memiliki dua grup (misal: tinggi/rendah).
- Uji Proporsi Dua Kelompok: untuk menguji apakah proporsi suatu kejadian berbeda signifikan antara dua grup.
- Uji Varians Dua Kelompok (F-test): untuk menguji kesamaan varians antar dua grup.
- Uji ANOVA Satu Arah dan Dua Arah: digunakan untuk mengetahui pengaruh satu atau dua faktor kategorik terhadap variabel numerik, jika jumlah kelompok lebih dari dua.

### **3.4.4 Regresi Linier Berganda**

Model ini digunakan untuk melihat pengaruh gabungan dari beberapa variabel sosial terhadap satu variabel target (misalnya tingkat kemiskinan). Model dilengkapi dengan pengujian asumsi regresi, dan interpretasi otomatis terhadap koefisien serta kebaikan model.

### 3.4.5 Klasterisasi Spasial

Menggunakan metode Hierarchical Clustering dengan algoritma Ward's D2, wilayah dikelompokkan ke dalam klaster berdasarkan kemiripan karakteristik sosial dan jarak spasial. Hasilnya divisualisasikan dalam peta interaktif yang menyajikan distribusi klaster dan jumlah anggota per klaster. Interpretasi juga diberikan secara otomatis untuk menjelaskan klaster terbanyak dan distribusinya.

### 3.5 Perangkat dan Tools

Penelitian ini dilakukan dengan perangkat lunak R, memanfaatkan beberapa Pustaka (package) sebagai berikut:

<b>Tujuan</b>	<b>Paket R yang Digunakan</b>
Visualisasi dan dashboard	Shiny, shinydashboard, leaflet, plotly, ggplot2
Analisis statistik	Car, nortest, multcomp
Uji asumsi dan regresi	Lmtest, car, vif
Analisis spasial dan peta	Sf
Manipulasi data	Dplyr, readr, tidyr

## **BAB IV**

### **HASIL DAN PEMBAHASAN**

#### **4.1 Tab Beranda**

1. Tujuan dan Fungsi:

Tab Beranda merupakan halaman pembuka dari dashboard Business Intelligence yang dibangun. Tab ini berperan sebagai orientasi awal bagi pengguna untuk memahami konteks, tujuan analisis, serta gambaran umum data yang digunakan. Dengan tampilan yang sederhana dan informatif, tab ini dirancang agar pengguna dari berbagai latar belakang dapat segera memahami fokus analisis, yaitu kerentanan sosial wilayah Indonesia berdasarkan data SoVI (Social Vulnerability Index).

2. Langkah Penggunaan:

Pada tab ini, pengguna tidak perlu melakukan interaksi teknis seperti memilih input atau mengunggah data. Halaman beranda hanya perlu dibaca dan dipahami, karena berisi:

- Judul dashboard
- Deskripsi singkat mengenai tujuan analisis
- Sumber data yang digunakan
- Penjelasan umum mengenai variabel-variabel yang membentuk indeks kerentanan sosial.

3. Penjelasan dan Interpretasi

Isi deskriptif pada Beranda mengarahkan pengguna bahwa dashboard ini bertujuan untuk mengidentifikasi wilayah yang tergolong rentan secara sosial berdasarkan berbagai aspek, seperti pendidikan, kemiskinan, kepadatan penduduk, kelompok usia rentan (anak-anak dan lansia), dan akses terhadap fasilitas dasar.

Misalnya, dari penjelasan awal, pengguna diarahkan untuk memahami bahwa data yang digunakan mencakup 511 kabupaten/kota di Indonesia, dengan sejumlah variabel numerik yang dijadikan indikator.

Melalui gambaran umum pada tab ini, pengguna mulai memahami bahwa analisis tidak dilakukan berdasarkan satu indikator saja, melainkan kombinasi dari berbagai dimensi sosial. Tab ini menjadi titik awal bagi pengguna sebelum menjelajahi tab lainnya yang lebih interaktif.

## 4.2 Tab Manajemen Data

### 1. Tujuan dan Fungsi:

Tab Manajemen Data pada dashboard dirancang untuk memfasilitasi pengguna dalam mengelola dan menyiapkan data sebelum dilakukan analisis lebih lanjut. Salah satu fitur utama dalam tab ini adalah mengubah variabel numerik (kontinu) menjadi kategori dua kelompok secara otomatis, berdasarkan nilai median. Fitur ini bertujuan untuk memudahkan pengguna dalam:

- Melakukan analisis statistik yang memerlukan data kategorik seperti uji beda dua kelompok atau ANOVA.
- Membandingkan kelompok dengan nilai tinggi dan rendah terhadap variabel-variabel kerentanan sosial.
- Memberikan interpretasi awal terhadap distribusi variabel dan kategorisasinya.

### 2. Langkah penggunaan:

- Setiap variabel numerik secara otomatis dibuatkan versi kategoriknya. Pengguna dapat memilih nilai dibagi dalam berapa kategori.
- Untuk setiap variabel kategorik, disediakan tabel terbaru yang menampilkan hasil kategorisasi dan disertai interpretasi otomatis.

Dengan mengkategorikan data secara sistematis dan menyertakan interpretasi yang informatif, tab ini memberikan fondasi penting bagi pengguna untuk melakukan uji beda dan analisis perbandingan antarkelompok. Tab ini juga membantu menjembatani antara data mentah dan analisis statistik inferensia di tab berikutnya.

## 4.3 Tab Eksplorasi Data

Tab Eksplorasi Data pada dashboard dirancang untuk memberikan pemahaman awal terhadap kondisi data melalui statistik deskriptif, grafik interaktif, dan peta tematik. Pengguna dapat melakukan eksplorasi variabel-variabel penting dalam indeks kerentanan sosial serta melihat distribusinya baik dalam bentuk angka, grafik batang, boxplot, hingga peta sebaran wilayah.

### 1. Statistik Deskriptif

Fitur ini menyajikan nilai minimum, maksimum, rata-rata, dan standar deviasi dari setiap variabel numerik. Langkah penggunaannya:

- Pilih variabel yang ingin dianalisis dan pilih jenis grafik

- Sistem akan menampilkan ringkasan statistik secara otomatis.

Interpretasi:

Statistik deskriptif ini membantu memahami karakteristik umum dari masing-masing variabel. Sebagai contoh, jika variabel *POVERTY* memiliki rata-rata tinggi dan deviasi standar besar, maka sebaran kemiskinan antar wilayah sangat bervariasi, menandakan adanya ketimpangan.

## 2. Grafik visualisasi

Pengguna dapat memilih jenis grafik yang ingin ditampilkan, seperti:

- Histogram: menunjukkan distribusi frekuensi suatu variabel.
- Boxplot: memperlihatkan persebaran data dan deteksi outlier.
- Grafik batang rata-rata per provinsi: membandingkan nilai variabel antar provinsi.

Langkah penggunaan:

- Pilih variabel dan jenis grafik yang diinginkan.
- Grafik akan tergambar otomatis dan dapat berubah sesuai pilihan.

Interpretasi:

Visualisasi ini memudahkan pengguna dalam mengidentifikasi pola, outlier, dan perbandingan antar wilayah. Misalnya, boxplot dapat menunjukkan apakah suatu variabel cenderung simetris atau memiliki pencilan ekstrem.

## 3. Peta Interaktif

### a. Peta sebaran nilai variabel

- Menampilkan warna wilayah berdasarkan besar kecilnya suatu variabel (contoh: *POVERTY*).
- Dilengkapi legenda nilai dan tooltip nama kabupaten/kota.

Interpretasi:

Peta ini membantu mengenali wilayah dengan nilai ekstrem (tinggi/rendah) untuk setiap indikator. Misalnya, pengguna dapat mengetahui bahwa wilayah A dan B merupakan daerah dengan persentase penduduk lansia tertinggi.

### b. Peta klustering



- Menampilkan hasil pengelompokan wilayah ke dalam beberapa klaster berdasarkan kemiripan nilai semua variabel.
- Proses klasterisasi dilakukan dengan hierarchical clustering berbasis matriks penimbang yang dimuat dari file distance.csv.

Hasil klasterisasi ini memudahkan dalam segmentasi wilayah berdasarkan kerentanan sosial. Misalnya:

- Klaster 1 berisi wilayah dengan kerentanan sangat tinggi.
- Klaster 2 berisi wilayah dengan tingkat kerentanan sedang.
- Klaster 3 berisi wilayah dengan kerentanan relatif rendah.

Hal ini dapat membantu dalam pengambilan kebijakan yang lebih terarah, seperti menetapkan wilayah prioritas intervensi sosial. Melalui visualisasi dan peta, pengguna dapat memahami persebaran dan variasi nilai antar wilayah serta mengidentifikasi pola spasial yang muncul. Klastering spasial juga memungkinkan pengelompokan wilayah yang homogen, sehingga dapat digunakan sebagai dasar pengambilan keputusan kebijakan berbasis wilayah.

#### **4.4 Tab Uji Asumsi Data**

Tab Uji Asumsi Data berfungsi untuk memeriksa dua asumsi dasar sebelum melakukan analisis statistik inferensial atau regresi linear, yaitu uji normalitas dan uji homogenitas varians. Asumsi ini penting agar hasil analisis yang dilakukan (misalnya uji beda rata-rata atau regresi) valid dan dapat diinterpretasikan secara benar.

##### **1. Uji normalitas (Shapiro-Wilk Test)**

Uji normalitas dilakukan untuk mengetahui apakah data suatu variabel numerik terdistribusi normal. Dalam dashboard ini, digunakan Shapiro-Wilk test, yang cocok untuk ukuran sampel kecil maupun sedang.

Langkah penggunaan:

- Pilih variabel numerik yang ingin diuji dari menu pilihan.
- Hasil uji akan langsung ditampilkan, termasuk nilai p-value dan interpretasinya.

Interpretasi:

- Jika  $p\text{-value} > 0.05$ , maka data dapat dikatakan berdistribusi normal.
- Jika  $p\text{-value} \leq 0.05$ , maka data tidak berdistribusi normal.

Contoh interpretasi dari output:

“P-value = 0.224 → Data terdistribusi normal. Uji parametrik dapat digunakan.”

Asumsi ini sangat penting dalam menentukan jenis uji statistik yang tepat. Jika data tidak normal, maka disarankan menggunakan uji non-parametrik seperti Mann-Whitney atau Kruskal-Wallis.

## 2. Uji Homogenitas Varians (Levene Test)

Uji homogenitas dilakukan untuk memeriksa apakah dua atau lebih kelompok memiliki varians yang sama (homogen). Dalam dashboard ini digunakan Levene's Test, yang relatif kuat terhadap penyimpangan dari normalitas.

Langkah penggunaan:

- Pilih variabel numerik yang akan diuji.
- Pilih variabel kategori (biasanya hasil kategorisasi otomatis yang sudah disediakan di tab Manajemen Data).
- Hasil uji ditampilkan bersama dengan interpretasi otomatis.

Interpretasi:

- Jika  $p\text{-value} > 0.05$ , maka varians antar kelompok homogen (asumsi terpenuhi).
- Jika  $p\text{-value} \leq 0.05$ , maka varians antar kelompok tidak homogen.

Contoh interpretasi:

“P-value = 0.013 → Varians antar kelompok tidak homogen. Sebaiknya gunakan uji non-parametrik untuk uji beda rata-rata.”

Sebelum melakukan uji beda rata-rata atau regresi, perlu dipastikan bahwa data memenuhi asumsi dasar. Jika asumsi tidak terpenuhi, maka jenis uji yang digunakan harus disesuaikan agar hasilnya valid. Dengan adanya uji asumsi ini, pengguna dapat menentukan metode analisis lanjutan yang paling sesuai berdasarkan karakteristik data.

## 4.5 Tab Statistik Inferensia

Tab Statistik Inferensia merupakan pusat analisis statistik pada dashboard ini. Tab ini berisi berbagai metode statistik inferensial untuk menguji hipotesis, yang meliputi:

1. Uji Beda Rata-rata (1 dan 2 Kelompok)
2. Uji Proporsi (1 dan 2 Kelompok)
3. Uji Varians (1 dan 2 Kelompok)
4. Uji Beda Rata-rata Lebih dari 2 Kelompok (ANOVA Satu Arah dan Dua Arah)

Setiap metode telah dilengkapi dengan antarmuka pemilihan variabel dan interpretasi otomatis berdasarkan hasil uji.

### 1. Uji Beda Rata-rata

#### a. Satu Kelompok (One Sample T-Test)

Digunakan untuk membandingkan rata-rata satu variabel terhadap nilai tertentu.

Langkah:

- Pilih variabel numerik.
- Masukkan nilai pembanding ( $\mu_0$ ).

Interpretasi:

Jika  $p\text{-value} < 0.05$ , maka rata-rata berbeda signifikan dari nilai pembanding.

#### b. Dua Kelompok (Independent Sample T-Test)

Digunakan untuk menguji apakah dua kelompok berbeda secara signifikan dalam satu variabel numerik.

Langkah:

- Pilih variabel numerik.
- Pilih variabel kategorik 2 kelompok (otomatis dari hasil kategorisasi).

Interpretasi:

Jika  $p\text{-value} < 0.05$ , maka terdapat perbedaan rata-rata yang signifikan antar kelompok.

### 2. Uji Proporsi

#### a. Satu Kelompok

Mengukur apakah proporsi keberhasilan (contoh: nilai 1 pada variabel biner) berbeda dari proporsi yang diharapkan.

Langkah:

- Pilih variabel biner (bernilai 0/1).
- Masukkan proporsi pembandingan (misal 0.5).

Interpretasi:

Jika  $p\text{-value} < 0.05$ , maka proporsi berbeda signifikan dari pembandingan.

b. Dua Kelompok

Membandingkan proporsi antara dua kelompok dalam suatu variabel biner.

Langkah:

- Pilih variabel biner (0/1).
- Pilih variabel kategorik 2 kelompok.

Interpretasi:

Jika  $p\text{-value} < 0.05$ , maka proporsi antar kelompok berbeda signifikan.

3. Uji Varians

a. Satu Kelompok

Mengukur apakah varians suatu variabel berbeda dari nilai tertentu (jarang digunakan, tapi tetap tersedia sebagai referensi).

b. Dua Kelompok

Digunakan untuk membandingkan varians dua kelompok. Dapat membantu memastikan asumsi homoskedastisitas.

Langkah:

- Pilih variabel numerik.
- Pilih variabel kategorik 2 kelompok.

Interpretasi:

Jika  $p\text{-value} < 0.05$ , maka terdapat perbedaan varians yang signifikan antar kelompok.

4. Uji ANOVA (Lebih dari 2 Kelompok)

a. ANOVA Satu Arah

Digunakan untuk menguji apakah terdapat perbedaan rata-rata antar lebih dari dua kelompok berdasarkan satu faktor.

Langkah:

- Pilih variabel numerik.
- Pilih satu variabel kategorik (>2 kategori).

Interpretasi:

Jika  $p\text{-value} < 0.05$ , maka terdapat perbedaan rata-rata antar kelompok.

b. ANOVA Dua Arah

Memungkinkan uji pengaruh dua faktor terhadap satu variabel respon, serta interaksi keduanya.

Langkah:

- Pilih variabel numerik (respon).
- Pilih dua variabel kategorik sebagai faktor.

Interpretasi:

Dashboard menampilkan hasil uji untuk masing-masing faktor dan interaksi:

- Efek utama faktor A
- Efek utama faktor B
- Interaksi A\*B

Jika salah satu  $p\text{-value} < 0.05$ , maka faktor tersebut berpengaruh signifikan terhadap variabel respon.

#### 4.6 Tab Regresi Linier Berganda

Menu Regresi Linear Berganda pada dashboard ini ditujukan untuk menganalisis pengaruh beberapa variabel prediktor (independen) terhadap satu variabel respon (dependen) secara simultan. Menu ini juga dilengkapi dengan uji asumsi regresi dan interpretasi otomatis dari hasil analisis.

Fitur Utama

1. Pengaturan Model:

- Pengguna memilih satu variabel numerik sebagai variabel Y (respon).

- Pengguna dapat memilih lebih dari satu variabel numerik sebagai variabel X (prediktor).

## 2. Hasil Model Regresi:

- Menampilkan ringkasan model (koefisien, nilai  $R^2$ , F-statistic, dan p-value).
- Interpretasi disediakan secara otomatis berdasarkan nilai koefisien dan signifikansi.

## Uji Asumsi Regresi

Agar model regresi valid dan dapat digunakan untuk prediksi atau inferensi, sejumlah asumsi perlu diuji. Dashboard ini menyediakan uji asumsi berikut secara otomatis:

### 1. Uji Normalitas Residual (Shapiro-Wilk Test)

Tujuan: Memastikan bahwa sisa (residual) model regresi berdistribusi normal.

Interpretasi:

- $p\text{-value} > 0.05 \rightarrow$  Residual berdistribusi normal.
- $p\text{-value} \leq 0.05 \rightarrow$  Residual tidak normal (pelanggaran asumsi).

### 2. Uji Homoskedastisitas (Breusch-Pagan Test)

Tujuan: Memastikan bahwa residual memiliki varians yang sama (homoskedastisitas).

Interpretasi:

- $p\text{-value} > 0.05 \rightarrow$  Tidak terdapat heteroskedastisitas.
- $p\text{-value} \leq 0.05 \rightarrow$  Terdapat heteroskedastisitas (varian residual tidak homogen).

### 3. Uji Autokorelasi (Durbin-Watson Test)

Tujuan: Menguji apakah residual saling berkorelasi (khususnya dalam data spasial atau runtun waktu).

Interpretasi:

- Nilai DW  $< 1.5 \rightarrow$  Terdapat autokorelasi positif.
- Nilai DW  $> 2.5 \rightarrow$  Terdapat autokorelasi negatif.
- Nilai DW antara 1.5 – 2.5  $\rightarrow$  Tidak ada autokorelasi signifikan.

#### 4. Uji Multikolinieritas (VIF - Variance Inflation Factor)

Tujuan: Mengukur apakah antar variabel X saling berkorelasi tinggi.

Interpretasi:

- $VIF < 5 \rightarrow$  Tidak ada multikolinieritas.
- $5 \leq VIF < 10 \rightarrow$  Indikasi multikolinieritas sedang.
- $VIF \geq 10 \rightarrow$  Indikasi multikolinieritas kuat.

Penggunaan dan Interpretasi

Langkah-langkah analisis:

- 1) Pilih variabel respon Y dan satu atau beberapa prediktor X.
- 2) Dashboard secara otomatis menampilkan hasil model dan asumsi.
- 3) Interpretasi diberikan secara langsung, sehingga memudahkan pengambilan keputusan.

Contoh Keputusan:

Jika model menunjukkan bahwa variabel Elderly dan NoElectric berpengaruh signifikan terhadap Poverty, maka intervensi kebijakan dapat difokuskan pada perbaikan akses listrik dan perlindungan lansia.

Jika terjadi multikolinieritas, maka pengguna dapat mengurangi atau menggabungkan variabel prediktor.

Dengan model regresi yang valid, dashboard ini tidak hanya menyajikan data, tetapi juga menghubungkan data dengan pengambilan keputusan berbasis bukti (evidence-based decision making).

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Dashboard Business Intelligence yang dikembangkan dalam proyek ini berhasil menyajikan informasi kerentanan sosial wilayah di Indonesia secara interaktif dan sistematis. Melalui tab Manajemen Data, pengguna dapat melakukan pengelolaan dan transformasi data, seperti mengkategorikan variabel numerik menjadi kategorik secara otomatis, yang menjadi dasar penting untuk keperluan analisis statistik lanjutan.

Pada bagian Eksplorasi Data, dashboard menyediakan berbagai visualisasi seperti grafik batang, scatter plot, tabel statistik deskriptif, hingga peta interaktif yang menggambarkan persebaran nilai variabel maupun hasil klasterisasi spasial. Visualisasi ini dilengkapi dengan interpretasi otomatis yang memberikan gambaran cepat mengenai wilayah-wilayah dengan kerentanan sosial tinggi maupun rendah.

Tab Uji Asumsi dan Statistik Inferensia memungkinkan pengguna melakukan analisis statistik lebih mendalam. Uji normalitas dan homogenitas dilakukan untuk memastikan validitas model statistik, sedangkan berbagai uji inferensia—termasuk uji beda rata-rata, uji proporsi, dan uji varians untuk satu atau dua kelompok, serta ANOVA—memudahkan dalam mengidentifikasi perbedaan signifikan antar kelompok wilayah berdasarkan indikator sosial tertentu.

Akhirnya, tab Regresi Linear Berganda digunakan untuk mengevaluasi pengaruh simultan berbagai faktor terhadap kerentanan sosial. Uji asumsi regresi seperti normalitas residual, autokorelasi, homoskedastisitas, dan multikolinieritas juga disediakan agar hasil regresi valid secara statistik. Dengan interpretasi otomatis yang tersedia di setiap tahap, dashboard ini tidak hanya mempermudah analisis, tetapi juga membantu dalam pengambilan keputusan berbasis data.

Secara keseluruhan, dashboard ini mampu menjawab rumusan masalah yang telah dirumuskan sebelumnya, serta memberikan kontribusi signifikan dalam mendukung pengambilan kebijakan sosial yang lebih tepat sasaran. Keberadaan fitur-fitur yang lengkap dan interpretatif menjadikan dashboard ini sebagai alat bantu analisis spasial-statistik yang sangat berguna.



## 5.2 Saran

1. Pengembangan dashboard lebih lanjut dapat dilakukan dengan menambahkan fitur prediksi atau machine learning sederhana untuk klasifikasi tingkat kerentanan wilayah secara otomatis.
2. Sumber data sebaiknya diperbarui secara berkala, agar hasil analisis tetap relevan dan sesuai dengan dinamika sosial terbaru.
3. Diperlukan pemahaman pengguna terhadap konteks kebijakan, agar hasil dashboard tidak hanya menjadi informasi pasif, melainkan dasar bagi perencanaan strategis, terutama dalam mitigasi risiko sosial dan kebencanaan.
4. Kolaborasi dengan pemangku kepentingan lokal sangat disarankan dalam menginterpretasi hasil peta dan klaster, agar intervensi lebih tepat sasaran berdasarkan kondisi sosial budaya setempat.

## DAFTAR PUSTAKA

- BMLMCMC. (2021). Distance Matrix for SoVI Dataset. Diakses dari: <https://raw.githubusercontent.com/bmlmcmc/naspaclust/main/data/distance.csv>
- BMLMCMC. (2021). SoVI Dataset – Social Vulnerability Index Indonesia. Diakses dari: [https://raw.githubusercontent.com/bmlmcmc/naspaclust/main/data/sovi\\_data.csv](https://raw.githubusercontent.com/bmlmcmc/naspaclust/main/data/sovi_data.csv)
- Kurniawan, R., Nasution, B. I., Agustina, N., & Yuniarto, B. (2022). Revisiting social vulnerability analysis in Indonesia data. *Data in Brief*, 40, 107743. <https://doi.org/10.1016/j.dib.2021.107743>
- Sudirman, S., Ambarwati, A., & Lubis, L. (2022). Identifikasi Pengukuran Indeks Kerentanan Sosial Daerah Rentan Bencana Di Jawa Timur. *Jurnal Ilmiah Manajemen Publik dan Kebijakan Sosial*, 6(2), 171-184.
- Yulia. (2022). Navigasi Penelitian Sosial. Medan: Media Penerbit Indonesia. [http://repository.mediapenerbitindonesia.com/401/1/1.%20T%20155%20-%20%28FINISH%20LAYOUT%29Navigasi\\_Penelitian\\_Sosial%20%281%29.pdf](http://repository.mediapenerbitindonesia.com/401/1/1.%20T%20155%20-%20%28FINISH%20LAYOUT%29Navigasi_Penelitian_Sosial%20%281%29.pdf)

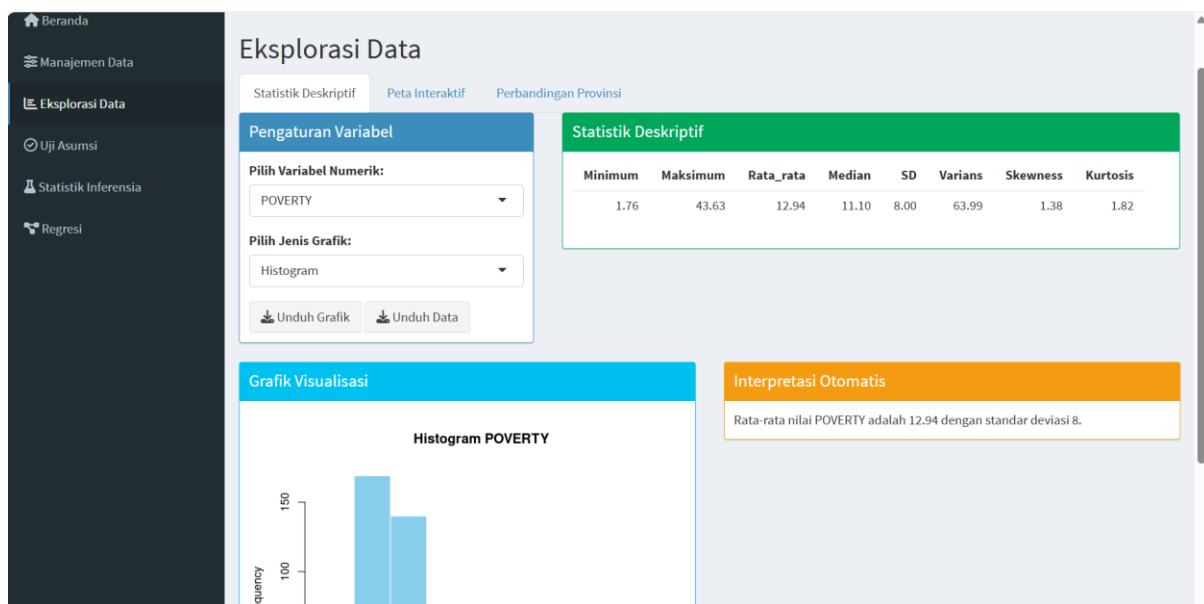
## LAMPIRAN

### Lampiran A – Struktur Data

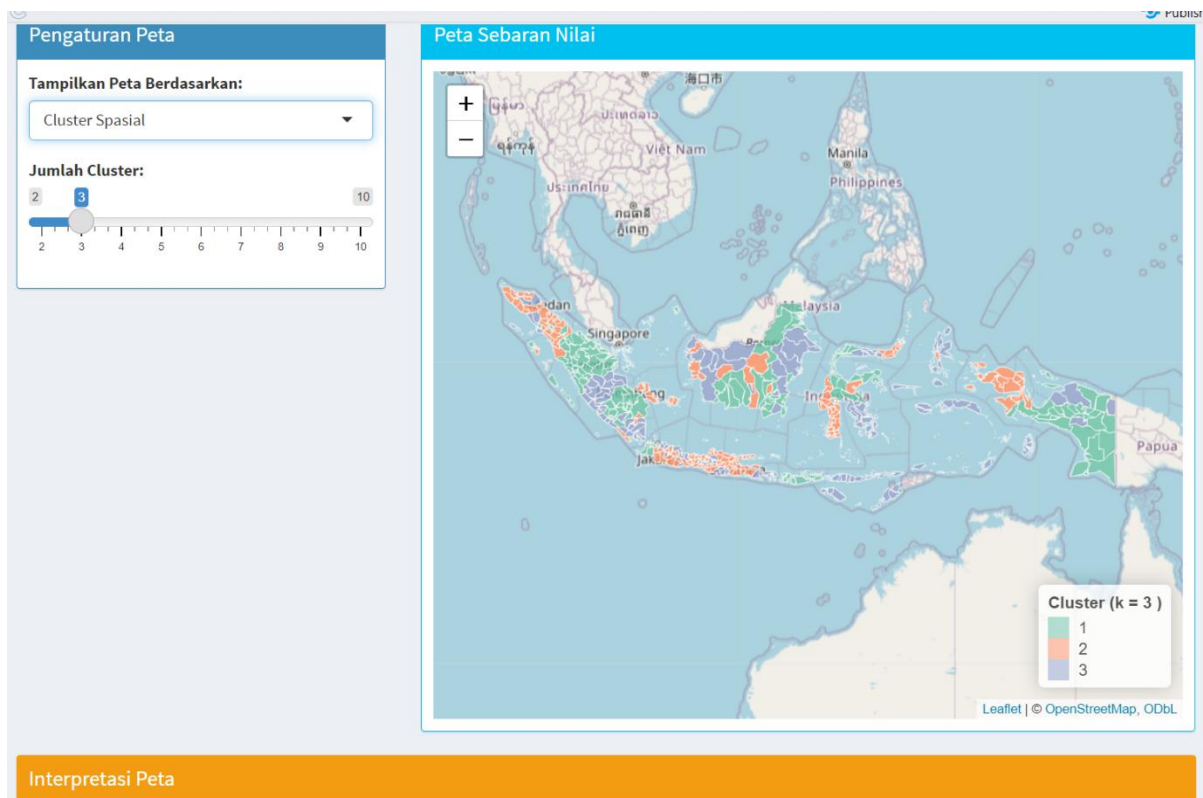
Metadata Dinamis Dataset SoVI							
Show	20	entries	Search: <input type="text"/>				
Label	Variabel	Deskripsi	Tipe_Data	Jumlah_Nilai_Unik	Data_Kosong	Contoh_Nilai	
DISTRICTCODE	Kode Wilayah	Kode dari wilayah/kabupaten	character	511	0	1101, 1102, 1103	
CHILDREN	Anak-anak	Persentase penduduk usia di bawah lima tahun	numeric	511	0	7.999956086, 13.51717522, 9.43677942	
FEMALE	Perempuan	Persentase penduduk perempuan	numeric	511	0	48.77590901, 49.68611677, 50.77554271	
ELDERLY	Lansia	Persentase penduduk usia 65 tahun ke atas	numeric	508	0	2.184700509, 2.296480805, 4.903403386	
FHEAD	Kepala Keluarga Perempuan	Persentase rumah tangga dengan kepala keluarga perempuan	numeric	511	0	13.11146752, 13.16721954, 20.73676104	
FAMILYSIZE	Ukuran Keluarga	Rata-rata jumlah anggota rumah tangga per wilayah	numeric	511	0	4.058094984, 4.479568758, 4.236907731	
NOELECTRIC	Tidak Listrik	Persentase rumah tangga yang tidak menggunakan listrik sebagai penerangan	numeric	400	0	1.425643767, 1.07433655, 0.50242042	
LOWEDU	Pendidikan Rendah	Persentase penduduk usia ≥15 tahun dengan pendidikan rendah	numeric	511	0	25.65208668, 28.7247149, 29.7794412	
GROWTH	Pertumbuhan Penduduk	Persentase perubahan jumlah penduduk	numeric	511	0	1.249634166, 2.287937103, 1.521355784	
POVERTY	Kemiskinan	Persentase penduduk miskin	numeric	457	0	20.2, 22.11, 14.07	
ILLITERATE	Buta Huruf	Persentase penduduk yang tidak bisa membaca dan menulis	numeric	511	0	5.019033186, 10.98267021, 7.717708677	
NOTTRAINING	Tidak Pelatihan Bencana	Persentase rumah tangga yang tidak mengikuti pelatihan bencana	numeric	458	0	92.71585138, 97.90033172, 98.76228546	
DPRONE	Rawan Bencana	Persentase rumah tangga yang tinggal di daerah rawan bencana	numeric	492	0	48.81047848, 73.08504222, 77.13620361	

RENTED	Sewa Rumah	Persentase rumah tangga yang menyewa rumah	numeric	505	0	4.882829903, 6.67973462, 3.337245122
NOSEWER	Tanpa Drainase	Persentase rumah tangga tanpa sistem drainase	numeric	511	0	22.89049274, 20.01281665, 31.78817662
TAPWATER	Sumber Air Pipa	Persentase rumah tangga yang menggunakan air perpipaan	numeric	468	0	5.595651787, 13.40470446, 6.98070999
POPULATION	Populasi	Jumlah total penduduk	numeric	511	0	91372, 119490, 231893
PROVINCE_NAME	Nama Provinsi	Nama provinsi tempat wilayah berada	character	34	0	Aceh, Sumatera Utara, Sumatera Barat
CITY_NAME	Nama Kota/Kabupaten	Nama kabupaten/kota tempat wilayah berada	character	511	0	Simeulue, Aceh Singkil, Aceh Selatan

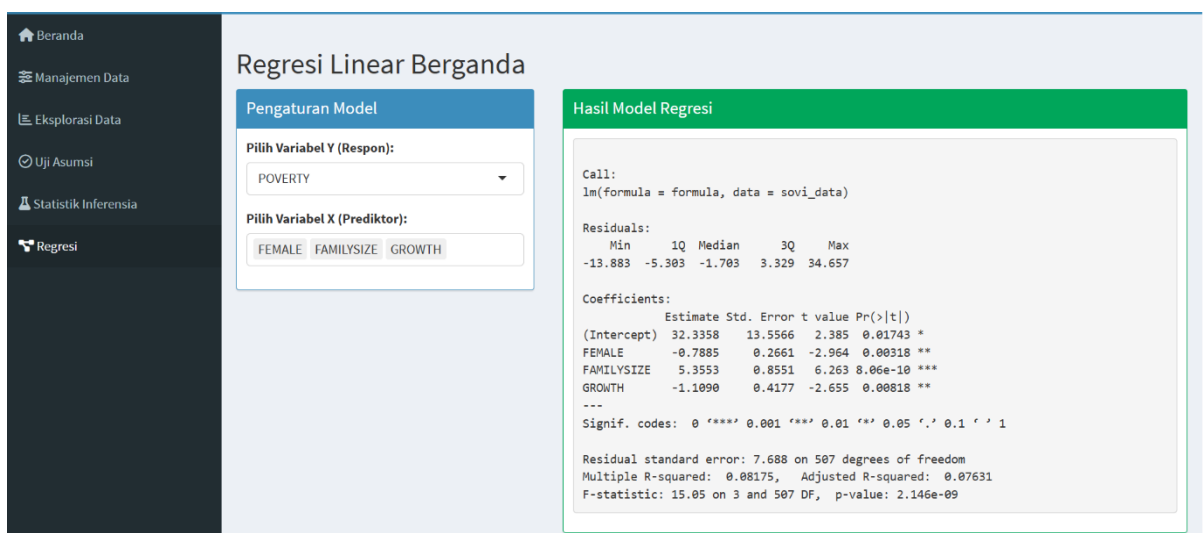
## Lampiran B – Tabel Statistik Deskriptif

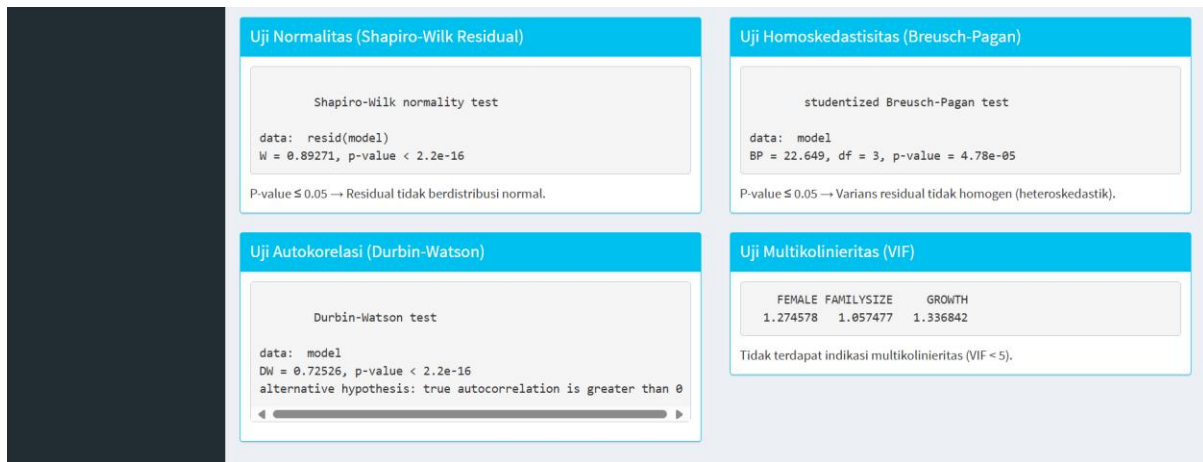


## Lampiran C – Peta Interaktif



## Lampiran D – Hasil Output Regresi Linear Berganda





## Lampiran E – Tampilan Dashboard

