**ANALISIS FAKTOR DEMOGRAFI DAN SOSIAL-EKONOMI TERHADAP TINGKAT KERENTANAN EKONOMI DAN SOSIAL DI INDONESIA**

**(STUDI KASUS DATA SUSENAS 2017)**



**Disusun Oleh:**

**Fadzilla Kusuma Ningrum**

**222313071**

**2KS3**

**Dosen Pengampu:**

**Yuliagnis Tranver Wijaya, S.ST., M.Sc.**

**PROGRAM STUDI D4 KOMPUTASI STATISTIK**

**POLITEKNIK STATISTIKA STIS**

**2024/2025**

Tahapan Pengolahan Data

1. Identifikasi Kebutuhan Data/Business Understanding
2. Masalah

Kerentanan sosial menunjukkan seberapa rentan individu atau wilayah terhadap dampak negatif dari guncangan, baik alamiah seperti bencana alam maupun non-alamiah seperti krisis ekonomi atau wabah penyakit. Dampak ini bisa memperburuk kondisi sosial-ekonomi masyarakat, memperluas kesenjangan, dan menghambat pembangunan. Indeks Kerentanan Sosial (Social Vulnerability Index – SVI) dikembangkan sebagai alat untuk mengukur, memetakan, dan memahami dimensi-dimensi kerentanan ini. SVI mengintegrasikan berbagai indikator demografi, sosial, ekonomi, dan lingkungan untuk memberikan gambaran tentang tingkat kerentanan suatu wilayah secara menyeluruh.

Berdasarkan permasalahan mengenai SVI di atas, topik "Analisis Faktor Demografi dan Sosial-Ekonomi terhadap Tingkat Kerentanan Ekonomi dan Sosial di Indonesia (Studi Kasus Data SUSENAS 2017)" berfokus pada tiga hal utama:

1. Identifikasi Faktor-Faktor yang Mendorong Kerentanan

Menganalisis bagaimana indikator seperti usia anak/lansia, pendidikan, kemiskinan, dan akses dasar berkontribusi terhadap kerentanan di tingkat kabupaten/kota.

1. Menyediakan Bukti Empiris untuk Pengambilan Kebijakan

Memberikan jawaban berbasis data terhadap pertanyaan, seperti apakah kabupaten dengan proporsi lansia yang tinggi cenderung lebih rentan? Apakah rendahnya tingkat pendidikan berpengaruh signifikan terhadap kemiskinan?

1. Melihat Keragaman Wilayah di Indonesia

Menyajikan gambaran yang lebih jelas dan mendalam tentang kerentanan sosial dan ekonomi di berbagai kabupaten/kota di Indonesia melalui visualisasi data.

Beberapa permasalahan yang ingin dianalisis dalam proyek ini sebagai berikut.

1. Bagaimana sebaran dan karakteristik indikator demografi, sosial-ekonomi, serta tingkat kemiskinan di berbagai kabupaten/kota di Indonesia?
2. Apakah data yang digunakan telah memenuhi asumsi dasar statistik untuk analisis lanjutan?
3. Apakah terdapat perbedaan atau hubungan signifikan secara statistik antar kelompok berdasarkan tingkat kerentanan ekonomi dan indikator lainnya?
4. Seberapa besar pengaruh faktor demografi dan sosial-ekonomi terhadap tingkat kerentanan ekonomi (kemiskinan)?
5. Tujuan proyek

Tujuan proyek ini adalah untuk:

1. Memberikan gambaran awal karakteristik indikator demografi, sosial-ekonomi, serta tingkat kemiskinan berdasarkan variabel-variabel seperti proporsi anak-anak, lansia, tingkat pendidikan, kepala rumah tangga perempuan, dan kemiskinan.
2. Memeriksa asumsi-asumsi statistik dasar seperti normalitas sebaran data dan homogenitas varians antar kelompok agar hasil analisis dapat dipercaya.
3. Melihat apakah perbedaan kondisi tersebut memang secara statistik berhubungan dengan tingkat kemiskinan atau kerentanan sosial lainnya.
4. Melihat hubungan antar berbagai variabel independen (misalnya proporsi lansia, tingkat pendidikan, dan akses infrastruktur) dengan variabel dependen yaitu tingkat kemiskinan.
5. Rancangan solusi dan tahapan untuk mencapai tujuan

Untuk mencapai tujuan proyek, tahapan-tahapan yang akan dilakukan adalah sebagai berikut.

1. Menggunakan visualisasi grafik, seperti diagram batang dan boxplot untuk mengetahui gambaran awal karakteristik berbagai variabel.
2. Melakukan uji normalitas dan homogenitas sebagai langkah awal sebelum melakukan statistik inferensia seperti uji t dan regresi linear. Jika asumsi tidak terpenuhi, maka perlu dipertimbangkan penggunaan transformasi data.
3. Membandingkan rata-rata, proporsi, atau varians dari kelompok-kelompok demografi atau sosial-ekonomi yang berbeda, seperti kelompok kabupaten dengan pendidikan tinggi dan rendah.
4. Melakukan analisis regresi linear berganda untuk melihat hubungan antar berbagai variabel independen (misalnya proporsi lansia, tingkat pendidikan, dan akses infrastruktur) dengan variabel dependen yaitu tingkat kemiskinan. Hasil dari model ini dievaluasi melalui signifikansi masing-masing variabel serta validitas model berdasarkan uji asumsi dan kebaikan model.
5. Mengambil Data

Sumber data:

Data diunduh melalui tautan:

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

1. Mengintegritaskan Data

Data yang diperoleh dapat dipertanggungjawabkan karena berasal dari sumber yang terpercaya, yaitu dataset yang dipublikasikan dalam jurnal ilmiah "Revisiting social vulnerability analysis in Indonesia data" melalui platform ScienceDirect. Dataset ini telah melalui validasi metodologi penelitian yang ketat sehingga memiliki kredibilitas tinggi untuk keperluan analisis statistik. Untuk memastikan kelayakan data setelah diimpor ke dalam dashboard, dilakukan serangkaian pemeriksaan integritas data yang mencakup validasi struktur dataset, konsistensi tipe data, dan pemeriksaan nilai-nilai anomali.

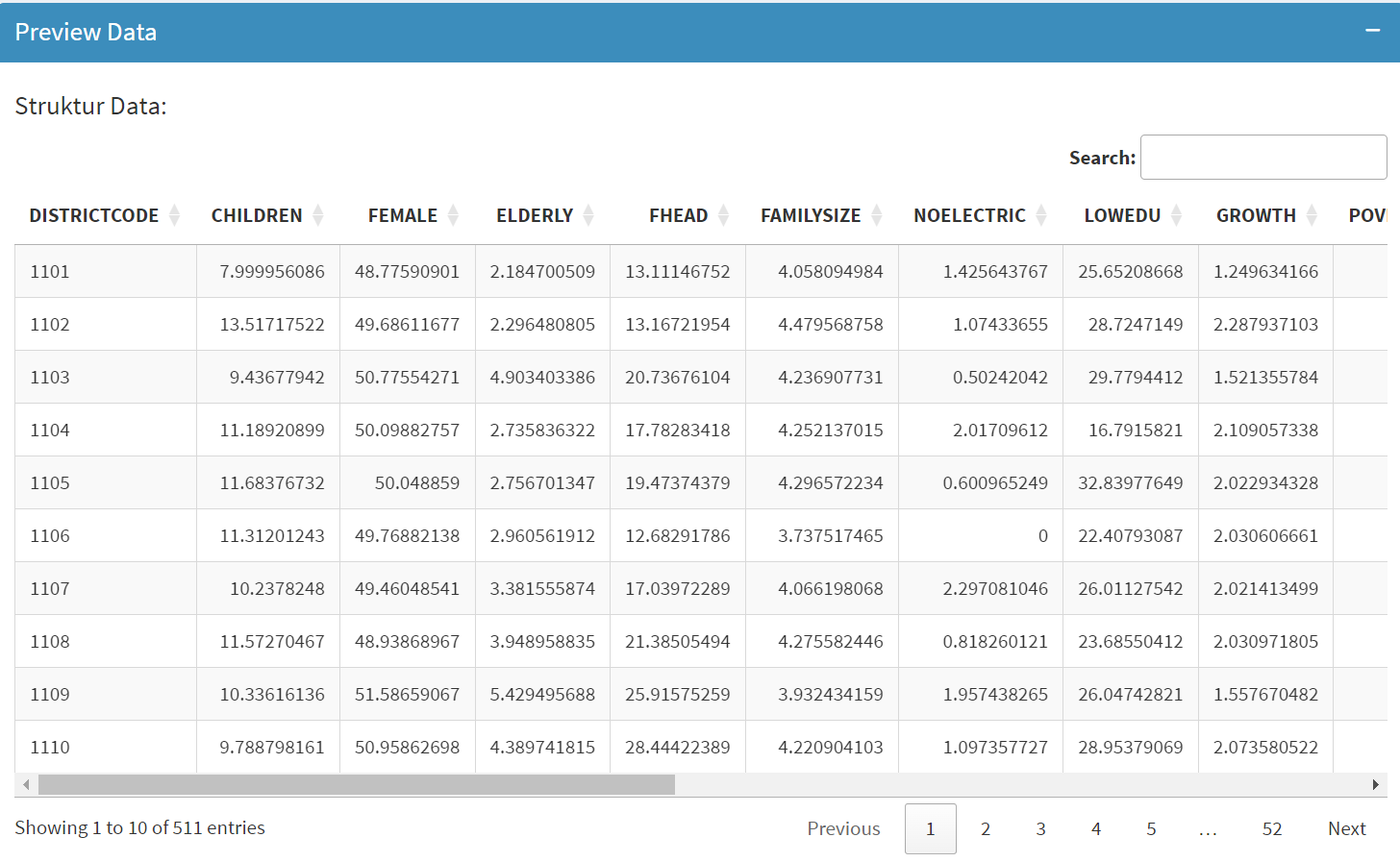
1. Menelaah Data
2. Identifikasi tipe data

Data SUSENAS 2017 memiliki 17 variabel dengan tambahan 2 variabel nama provinsi dan nama kabupaten/kota sesuai kode distrik. Berikut variabel yang ada di Data SUSENAS 2017 beserta tipe datanya.



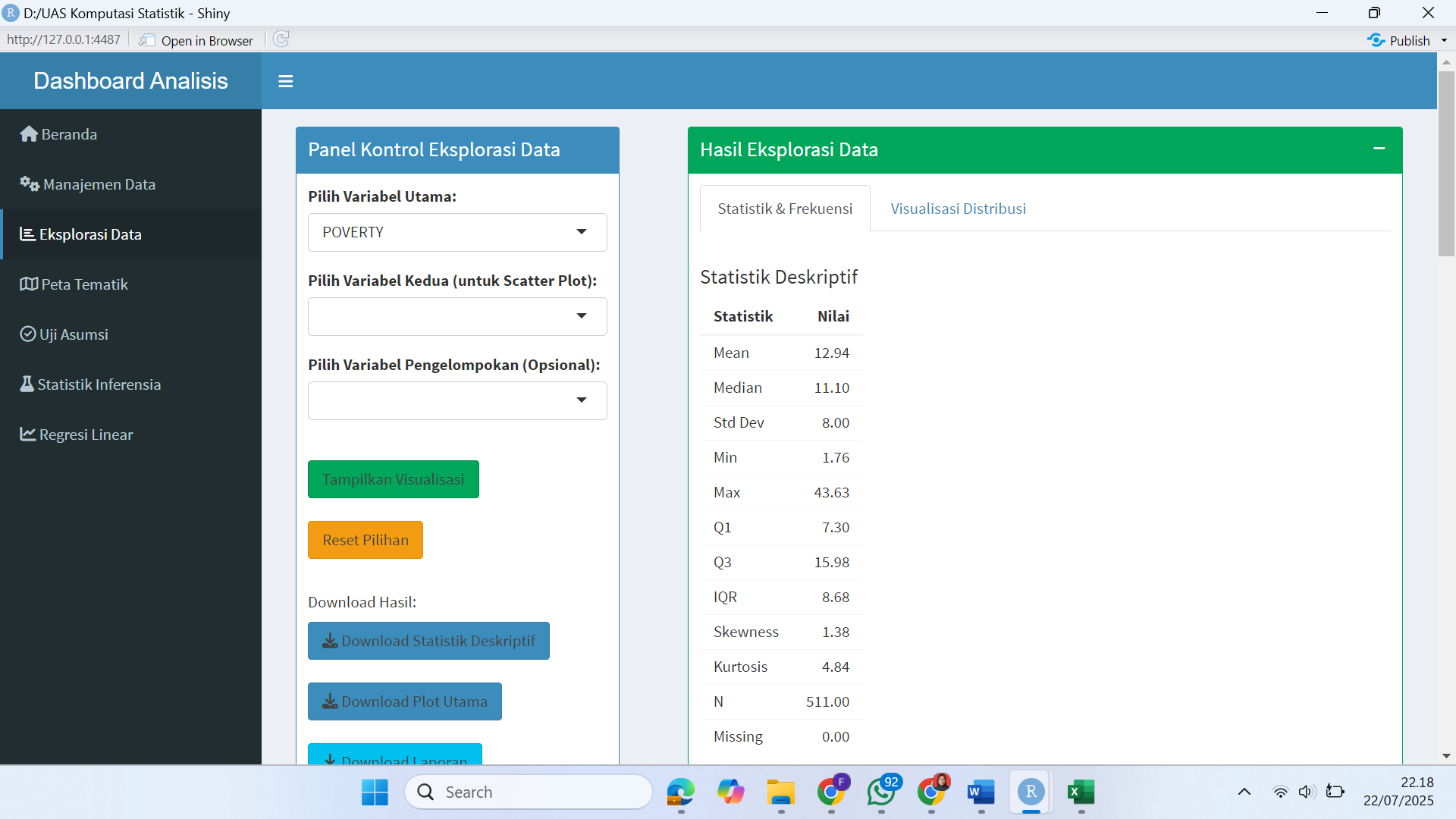
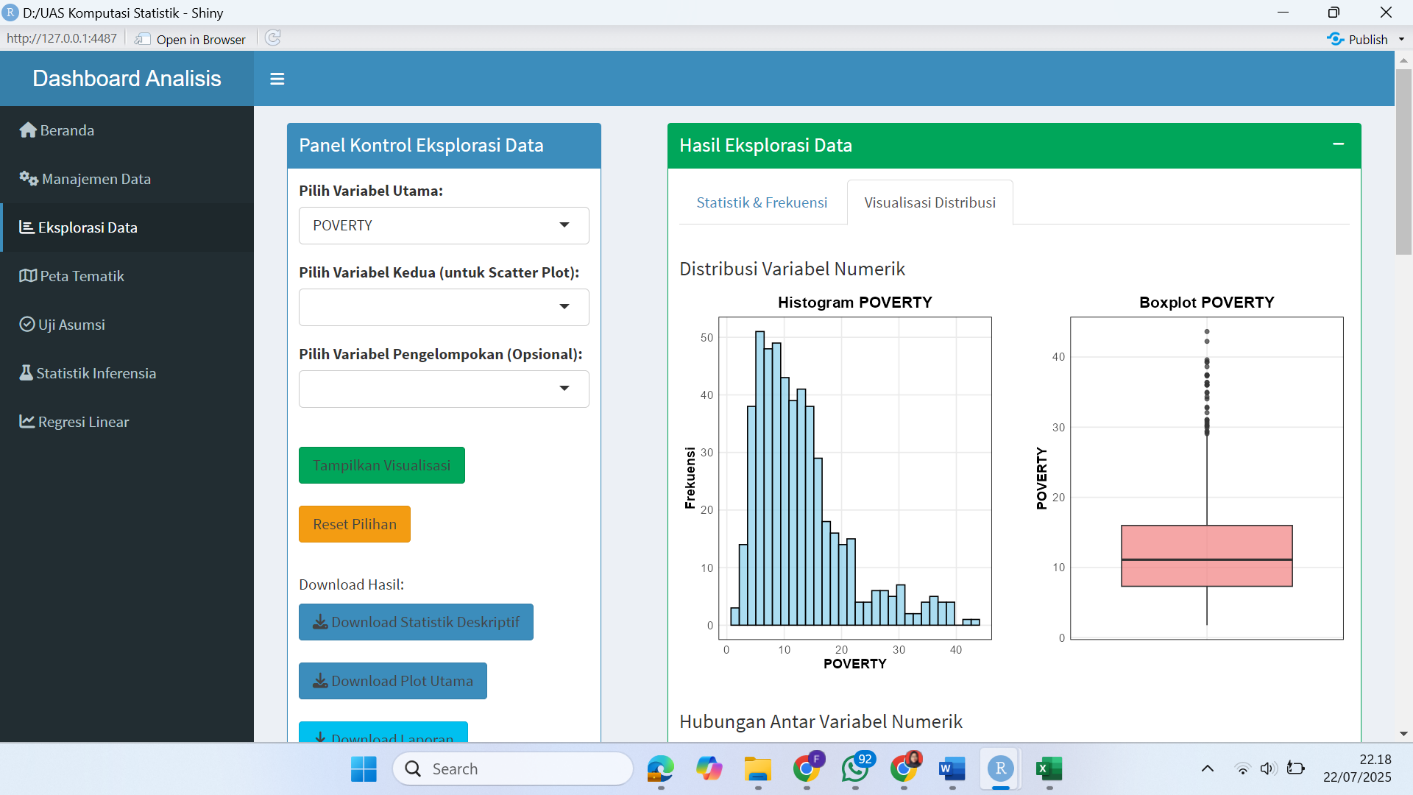
1. Cleaning data
2. Pengecekan *Missing Values*

Berdasarkan pemeriksaan yang dilakukan, ditemukan bahwa tidak ada missing values pada seluruh variabel dalam dataset Indeks SVI SUSENAS 2017. Hal ini mengindikasikan bahwa data yang tersedia lengkap dan tidak memerlukan proses imputasi atau penghapusan data akibat nilai yang hilang. Kelengkapan data ini memastikan bahwa analisis statistik yang akan dilakukan dapat menggunakan seluruh observasi tanpa mengurangi ukuran sampel atau memperkenalkan bias akibat penanganan missing values. Dalam dashboard ini, pemeriksaan missing values dilakukan secara otomatis untuk setiap variabel dan menampilkan ringkasan berupa jumlah dan persentase missing values per variabel. Jumlah total record (baris) dalam data adalah 511 baris



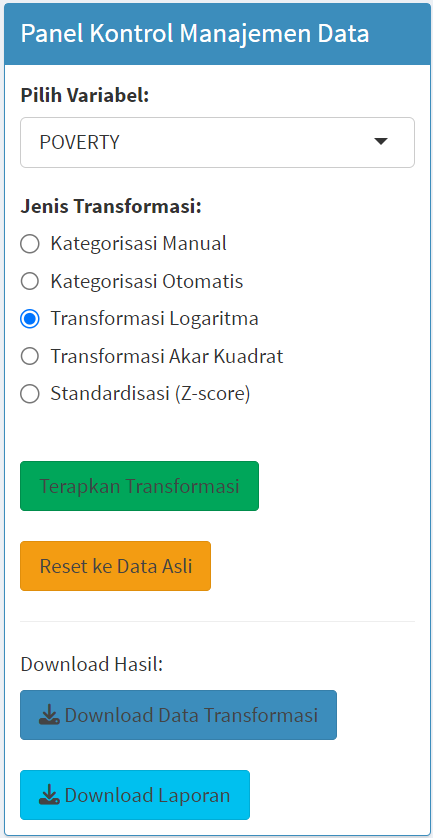
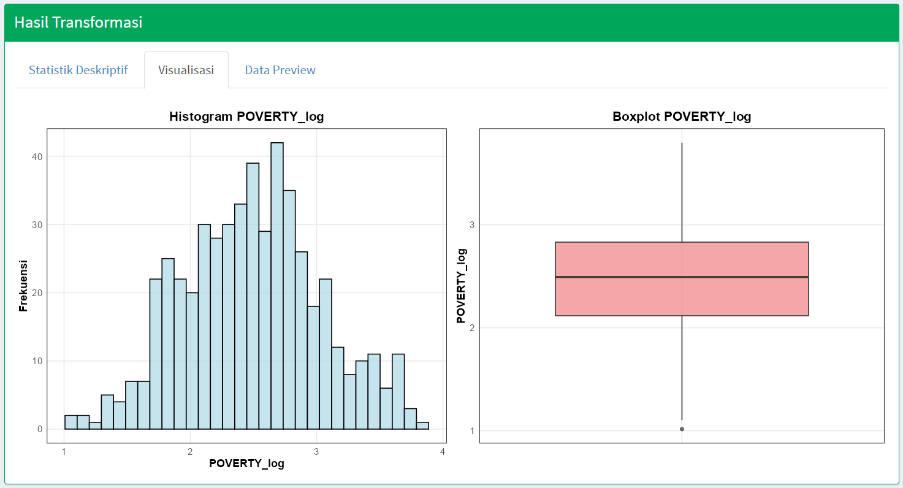
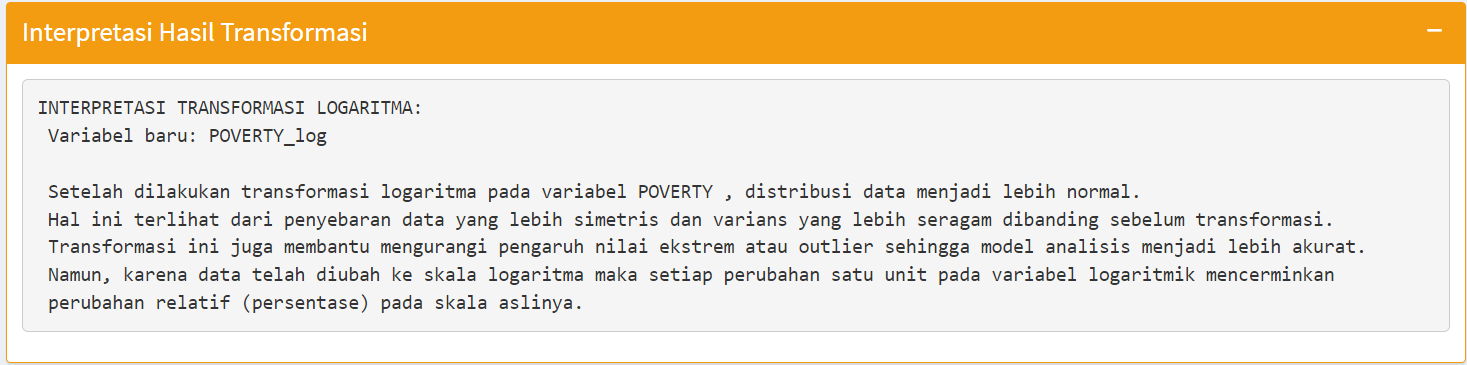
1. Pengecekan *Outlier*

Pengecekan outlier (nilai pencilan) dilakukan untuk mengidentifikasi observasi yang secara signifikan menyimpang dari pola umum data, yang dapat memengaruhi asumsi statistik dan hasil model. Dashboard menyediakan fitur identifikasi dan visualisasi outlier untuk membantu pengguna dalam mengevaluasi apakah outlier tersebut merupakan data yang valid atau perlu ditangani lebih lanjut.



1. Penanganan *Outlier*

Dashboard menyediakan beberapa metode penanganan outlier yang dapat dipilih sesuai dengan karakteristik data dan tujuan analisis. Penanganan outlier dalam data SOVI memerlukan pertimbangan khusus karena outlier dapat mencerminkan kondisi sosial ekonomi yang memang ekstrem di wilayah tertentu, bukan kesalahan data. Metode penanganan outlier yang tersedia dalam dashboard meliputi transformasi data menggunakan logaritma natural atau akar kuadrat untuk mengurangi pengaruh nilai ekstrem. Setiap metode penanganan outlier dilengkapi dengan visualisasi yang menunjukkan distribusi data sebelum dan sesudah penanganan, serta penjelasan interpretasi yang membantu pengguna dalam memahami dampak dari setiap metode terhadap analisis selanjutnya.



1. Memvalidasi Data

Sebelum melakukan analisis data lebih lanjut, dilakukan validasi data untuk memastikan bahwa dataset telah siap dan layak digunakan untuk berbagai jenis analisis statistik yang tersedia dalam dashboard. Proses validasi data dalam dashboard meliputi pemeriksaan struktur data untuk memastikan tipe data setiap variabel sesuai dengan kebutuhan analisis, verifikasi rentang nilai untuk memastikan tidak ada nilai yang di luar batas logis (misalnya persentase tidak melebihi 100%), dan konfirmasi konsistensi antar variabel yang memiliki hubungan logis.

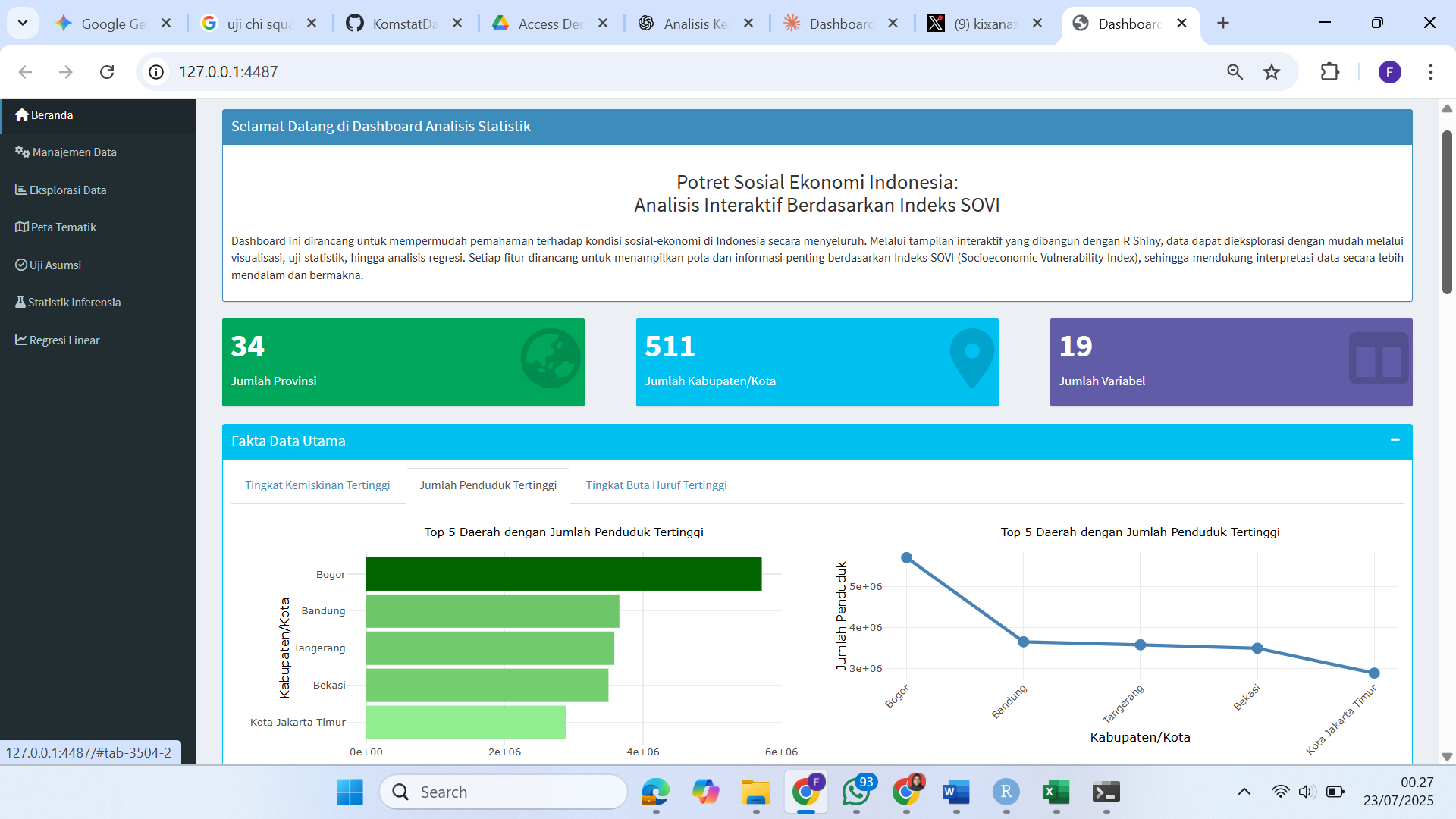
1. Menentukan Objek Data

Proyek dashboard ini bertujuan untuk menganalisis data Indeks Kerentanan Sosial Ekonomi (SOVI) Indonesia yang mencakup 511 kabupaten/kota di seluruh Indonesia. Dashboard menyediakan fleksibilitas bagi pengguna untuk memilih kombinasi variabel yang sesuai dengan fokus analisis yang diinginkan. Langkah penentuan objek data dilakukan melalui menu filter dan seleksi variabel dalam dashboard. Pengguna dapat memilih variabel independen dan dependen berdasarkan hipotesis atau pertanyaan penelitian yang ingin dijawab. Misalnya, untuk analisis faktor-faktor yang mempengaruhi tingkat kemiskinan, variabel POVERTY dapat dipilih sebagai variabel dependen, sedangkan variabel seperti CHILDREN, LOWEDU, dan ILLITERACY dapat dipilih sebagai variabel independen.

1. Membuat Business Intelligence
2. Menu Beranda

Menu beranda berfungsi sebagai pintu masuk yang informatif sebelum pengguna melakukan analisis di menu-menu lainnya. Menu beranda menyediakan fitur-fitur utama sebagai berikut:

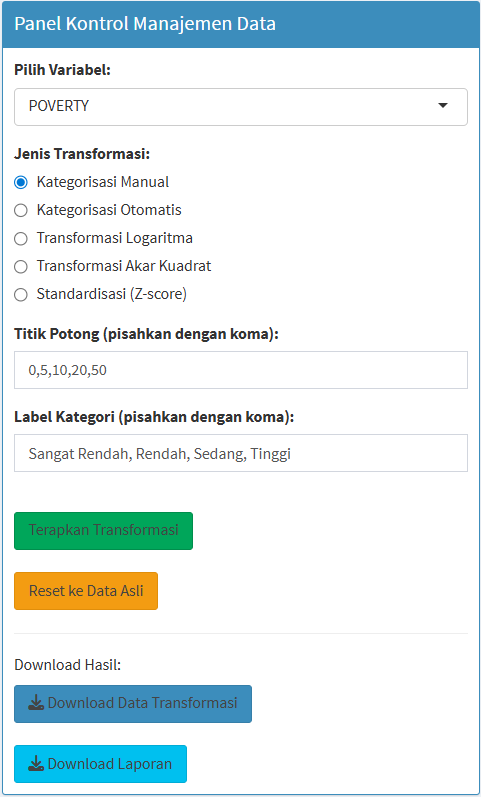
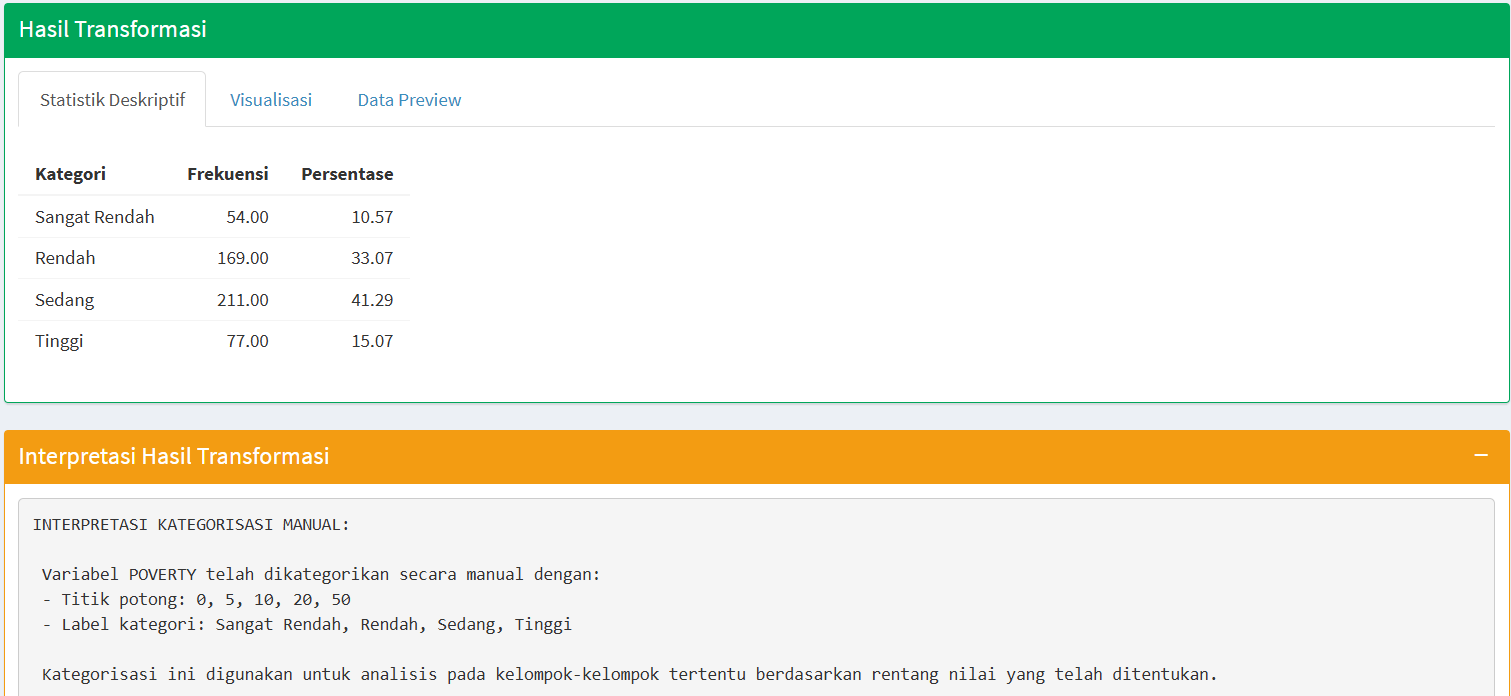
* Informasi Dataset: Menampilkan metadata, deskripsi variabel, dan sumber data SOVI Indonesia
* Statistik Ringkasan: Value boxes yang menunjukkan jumlah provinsi, total observasi, dan jumlah variabel dalam dataset
* Fakta Data Utama: Visualisasi interaktif daerah dengan tingkat kemiskinan tertinggi, jumlah penduduk terbesar, dan tingkat buta huruf tertinggi
* Preview Data: Tabel interaktif untuk melihat struktur dan isi dataset
* Fitur Download: Akses untuk mengunduh data lengkap dalam format CSV



1. Menu Manajemen Data

Menu manajemen data menyediakan fitur-fitur untuk preprocessing dan transformasi data sebagai berikut:

* Kategorisasi Data: Mengubah variabel kontinyu menjadi kategorik menggunakan kategorisasi manual yang diisikan pengguna atau kategori otomatis dengan metode equal interval, quantile, atau Jenks natural breaks
* Transformasi Variabel: Fitur untuk melakukan transformasi logaritma, akar kuadrat, dan standardisasi pada variabel numerik
* Statistik Deskriptif dan Visualisasi Before-After: Menampilkan perbandingan distribusi data sebelum dan sesudah transformasi melalui histogram dan boxplot

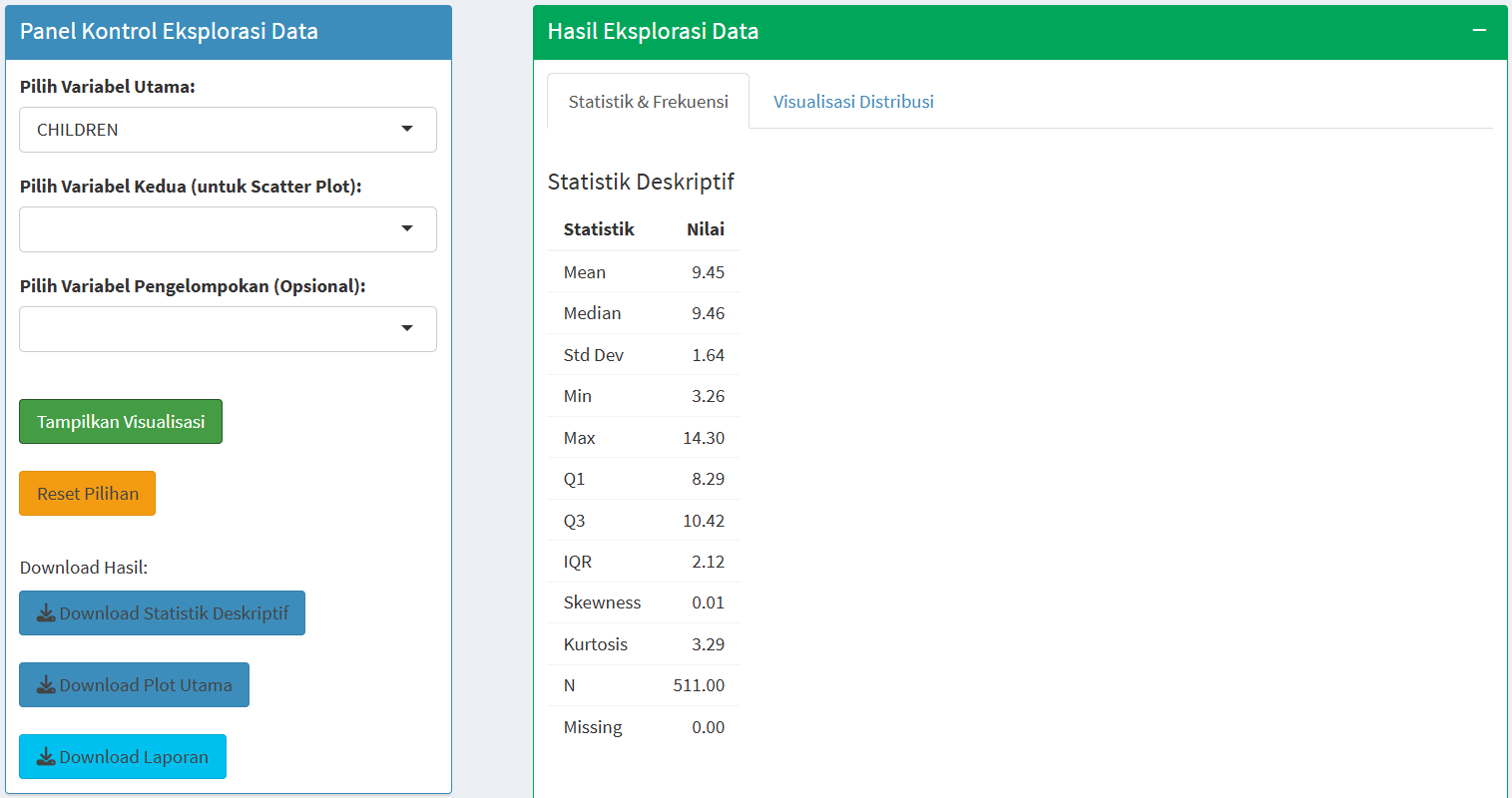


1. Menu Eksplorasi Data

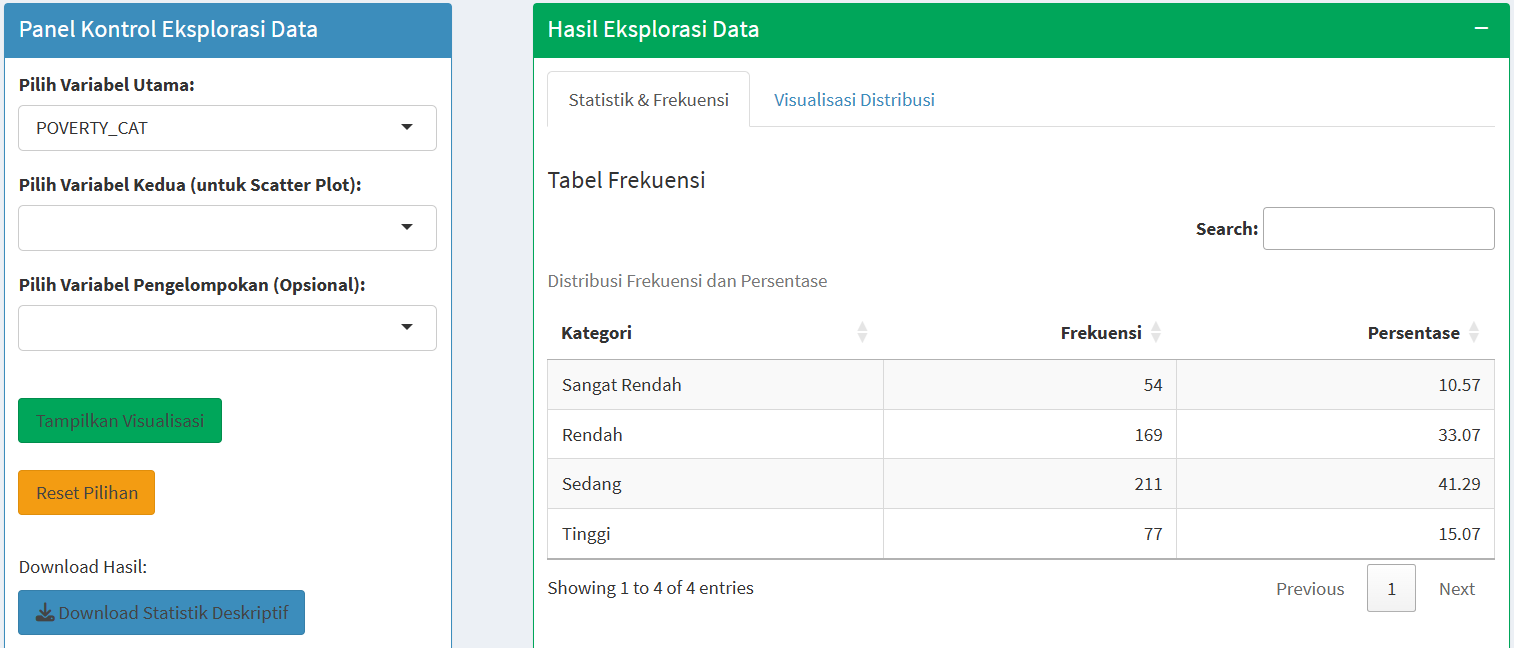
Menu eksplorasi data menyediakan fitur-fitur untuk analisis deskriptif dan visualisasi data sebagai berikut:

* Statistik Deskriptif: Menampilkan ringkasan statistik lengkap meliputi mean, median, standar deviasi, skewness, dan kurtosis untuk setiap variabel

Variabel numerik

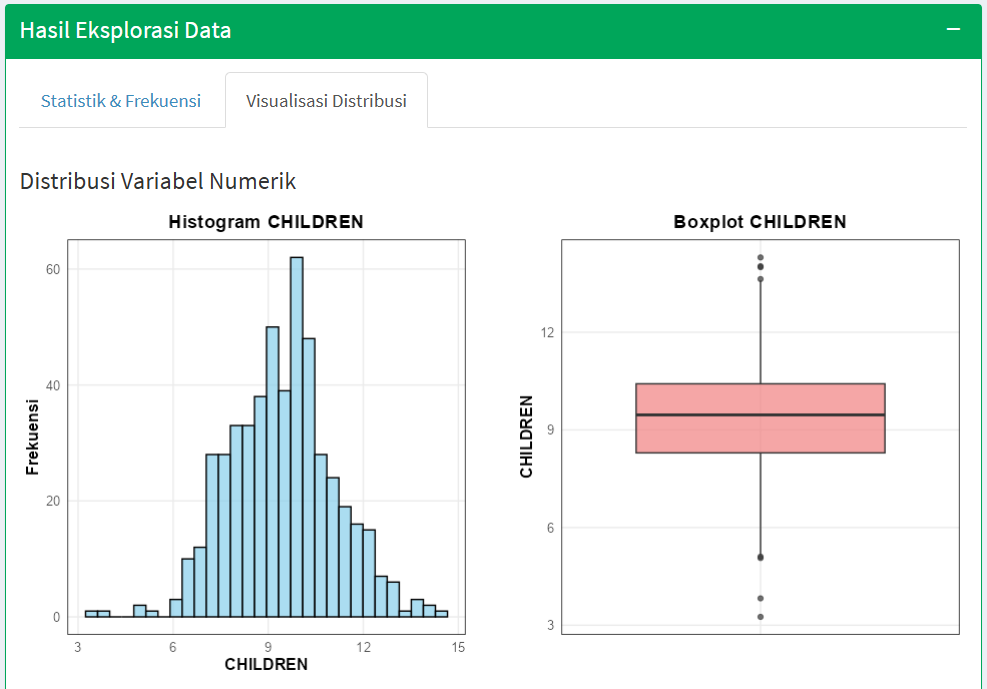


Variabel kategorik

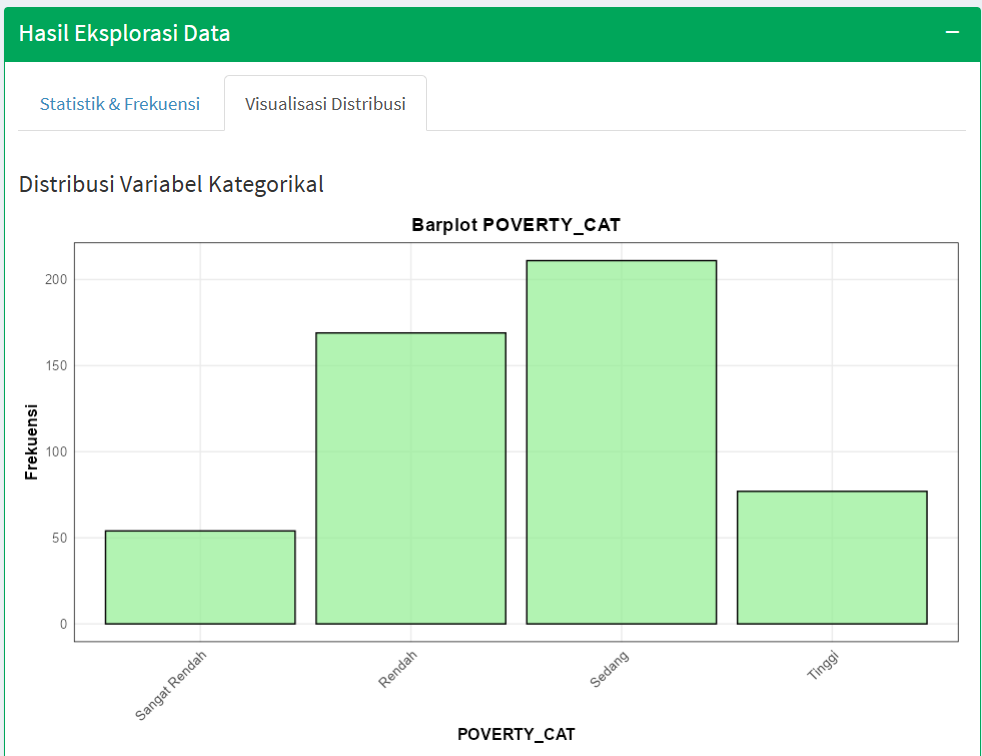


* Visualisasi Distribusi: Histogram, boxplot, dan barplot untuk memahami distribusi dan pola data pada variabel tunggal

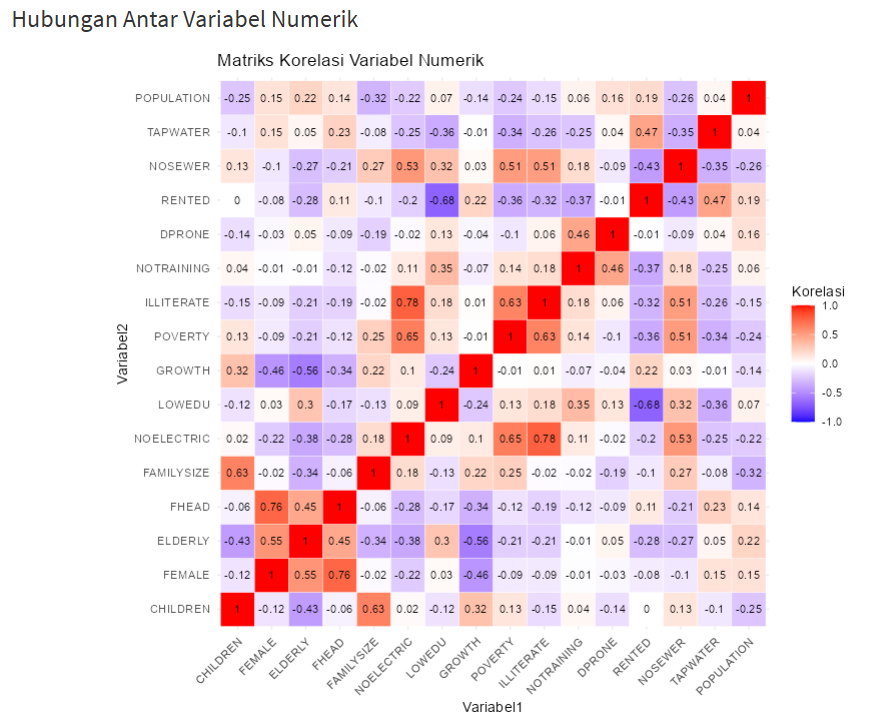
Variabel numerik



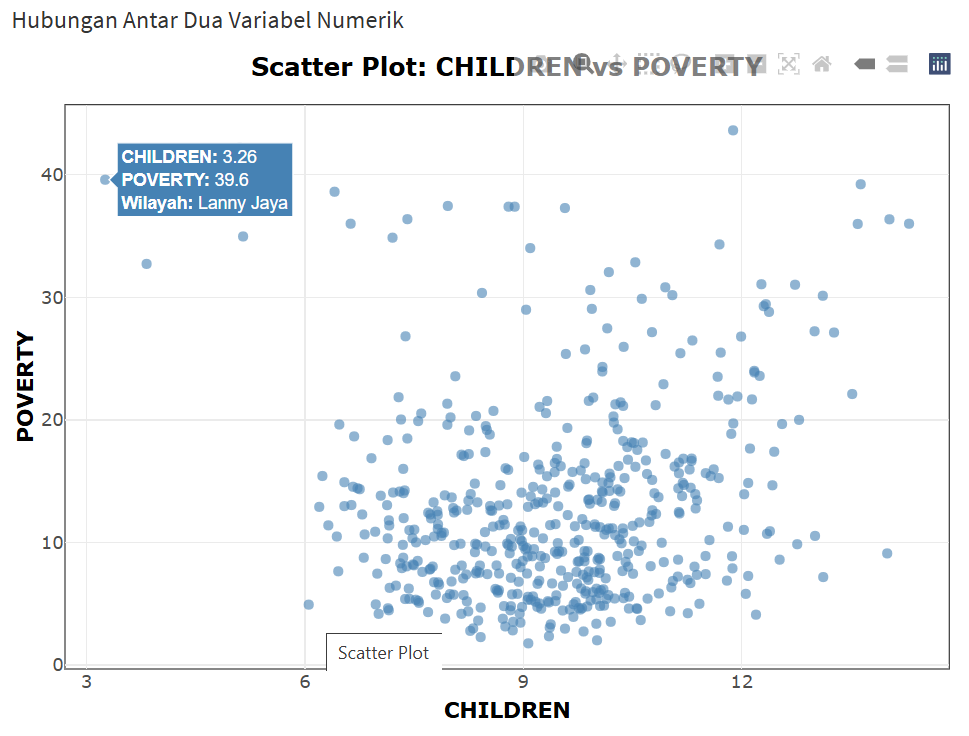
Variabel Kategorik



* Analisis Korelasi: Heatmap korelasi interaktif untuk mengidentifikasi hubungan antar variabel numerik



* Scatter Plot: Visualisasi hubungan bivariat dengan fitur interaktif untuk eksplorasi detail



* Boxplot dan Histogram Berkelompok: Perbandingan distribusi variabel numerik berdasarkan kategori tertentu

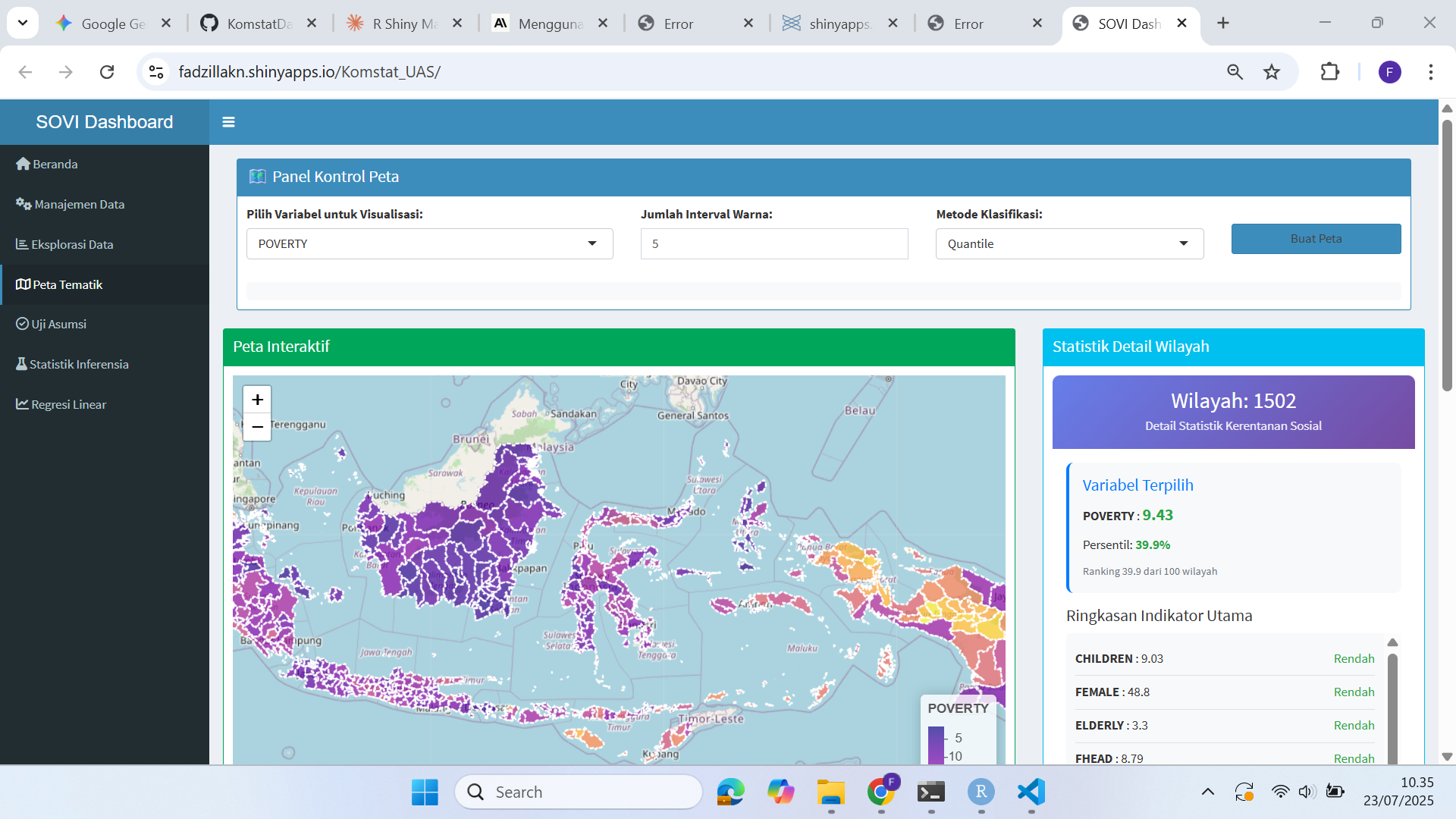


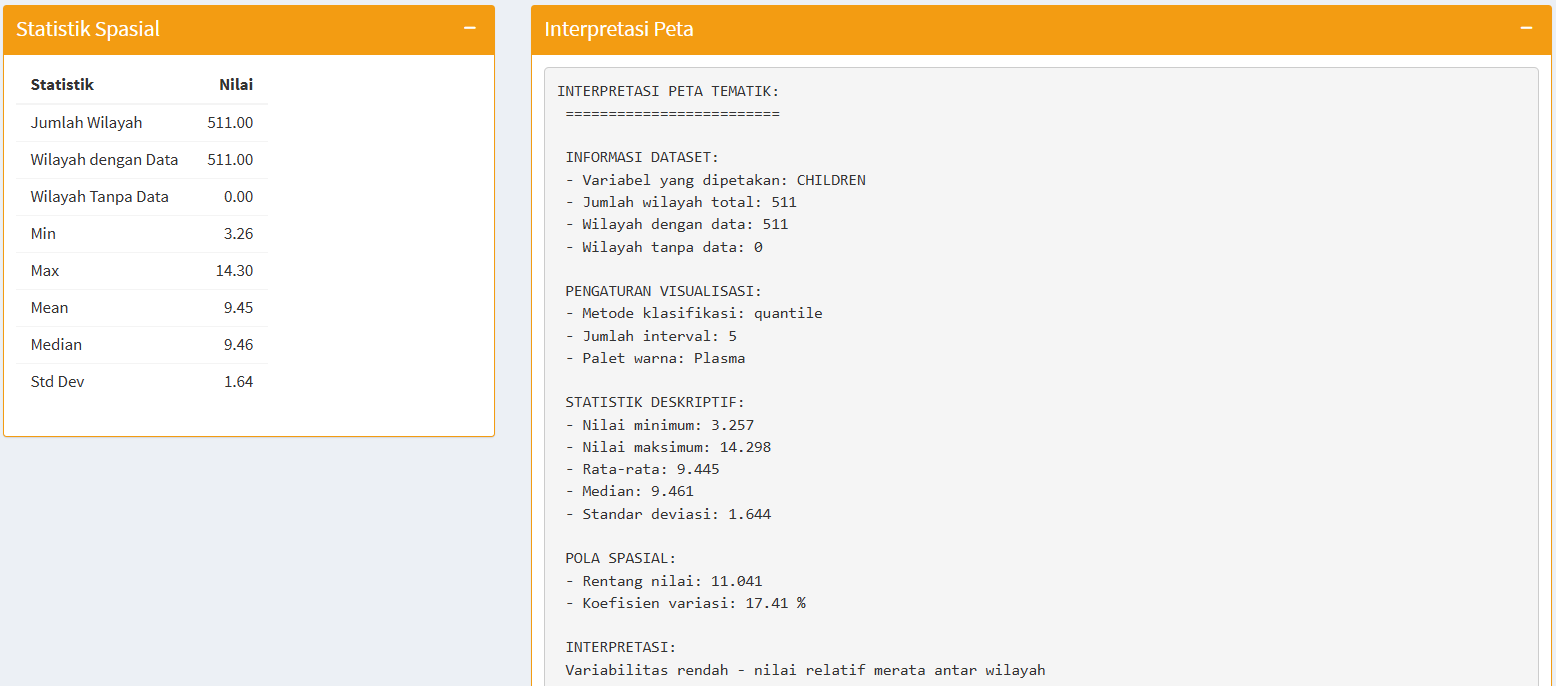
* Interpretasi Otomatis: Penjelasan hasil analisis untuk setiap output visualisasi
* Download Output: Download tabel statistik, plot, dan interpretasi

1. Menu Peta Tematik

Menu peta tematik menyediakan fitur-fitur untuk visualisasi spasial data SOVI sebagai berikut:

* Peta Choropleth: Visualisasi distribusi spasial variabel SOVI dengan pewarnaan berdasarkan nilai pada peta Indonesia
* Pemilihan Variabel: Dropdown untuk memilih variabel yang akan divisualisasikan pada peta
* Klasifikasi Data: Metode kategorisasi seperti equal interval, quantile, atau Jenks natural breaks untuk pembagian kelas
* Interaktivitas: Fitur hover dan klik pada wilayah untuk menampilkan informasi detail
* Legenda Dinamis: Legenda yang menyesuaikan dengan variabel dan metode klasifikasi yang dipilih
* Interpretasi Otomatis: Penjelasan hasil analisis untuk setiap output visualisasi

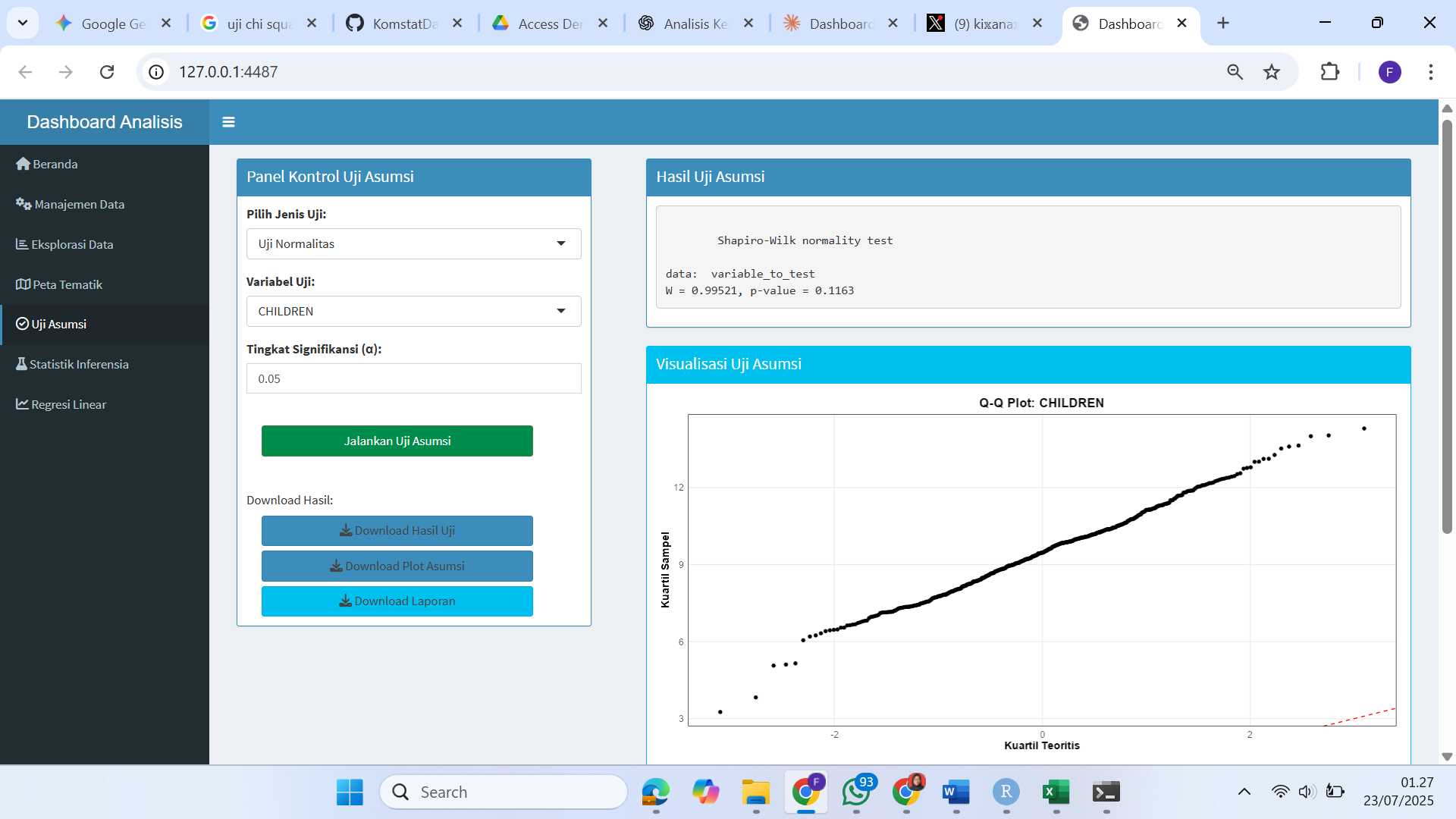




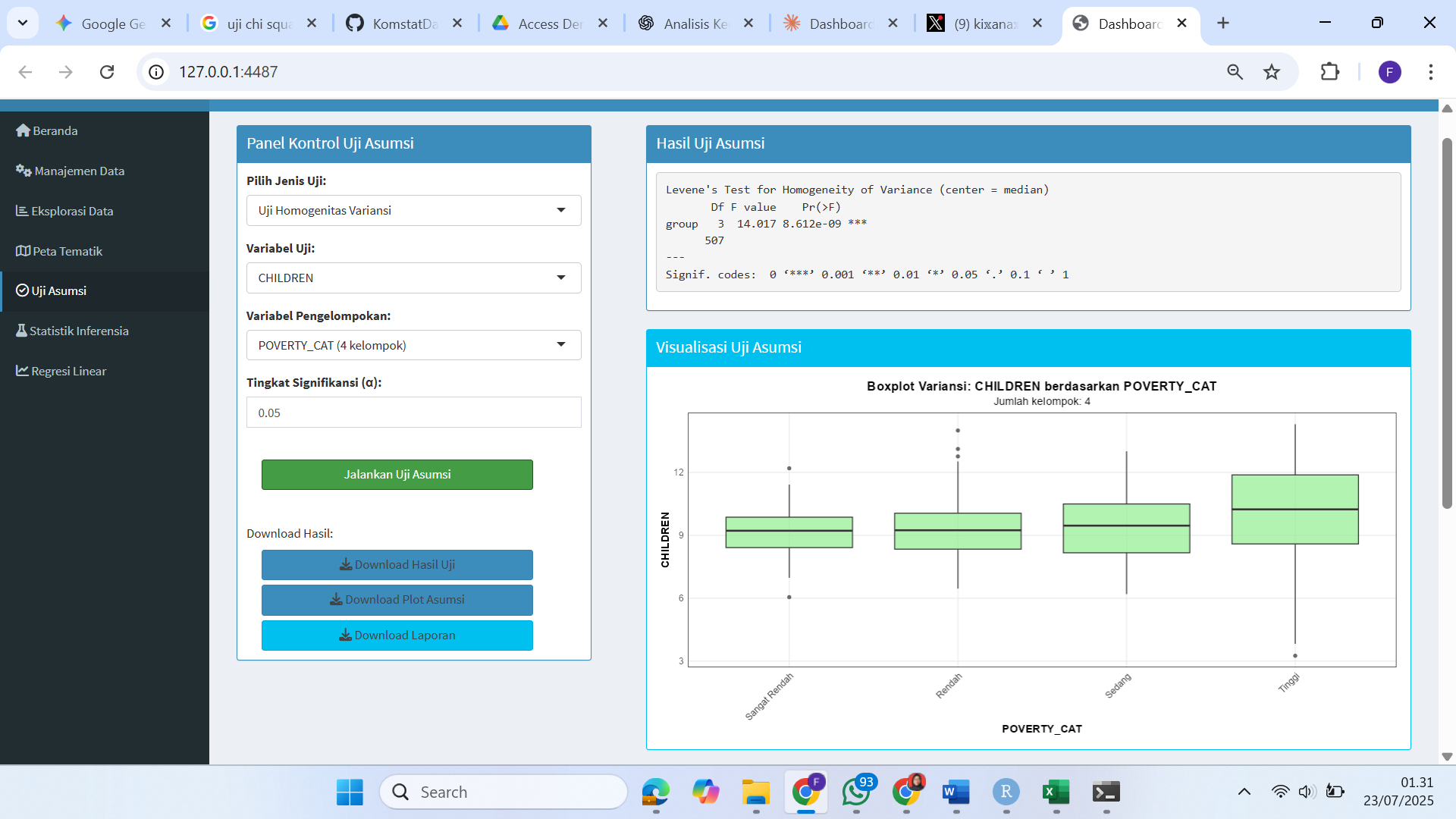
1. Menu Uji Asumsi

Menu uji asumsi menyediakan fitur-fitur untuk pengujian asumsi statistik sebagai berikut:

* Uji Normalitas: Pengujian distribusi normal menggunakan Shapiro-Wilk test dan Kolmogorov-Smirnov test



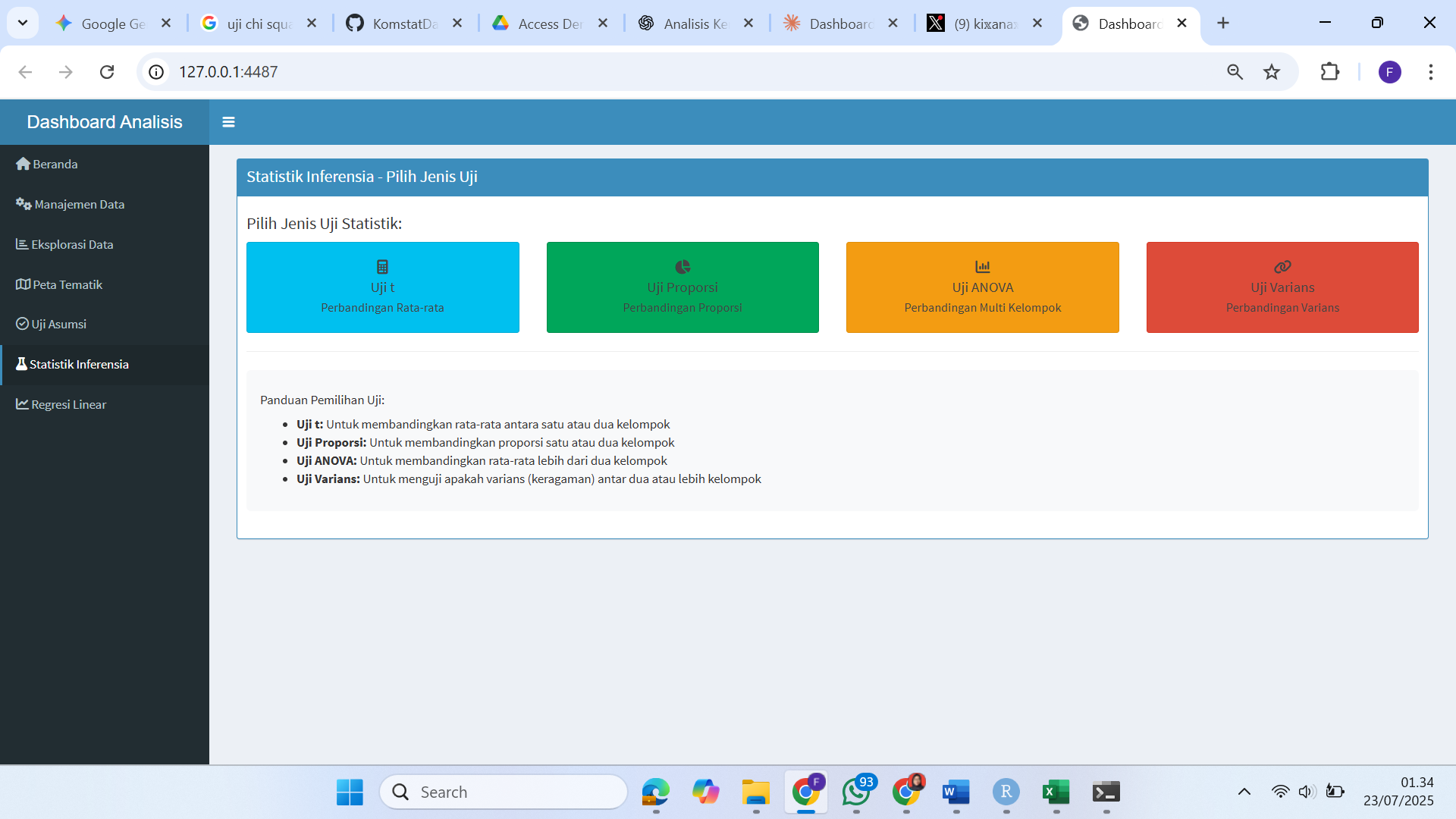
* Uji Homogenitas: Pengujian kesamaan varians menggunakan Levene test dan Bartlett test



* Interpretasi Hasil: Penjelasan otomatis hasil uji statistik dengan kesimpulan penerimaan atau penolakan hipotesis
* Download Hasil: Download output uji asumsi dalam format tabel dan visualisasi

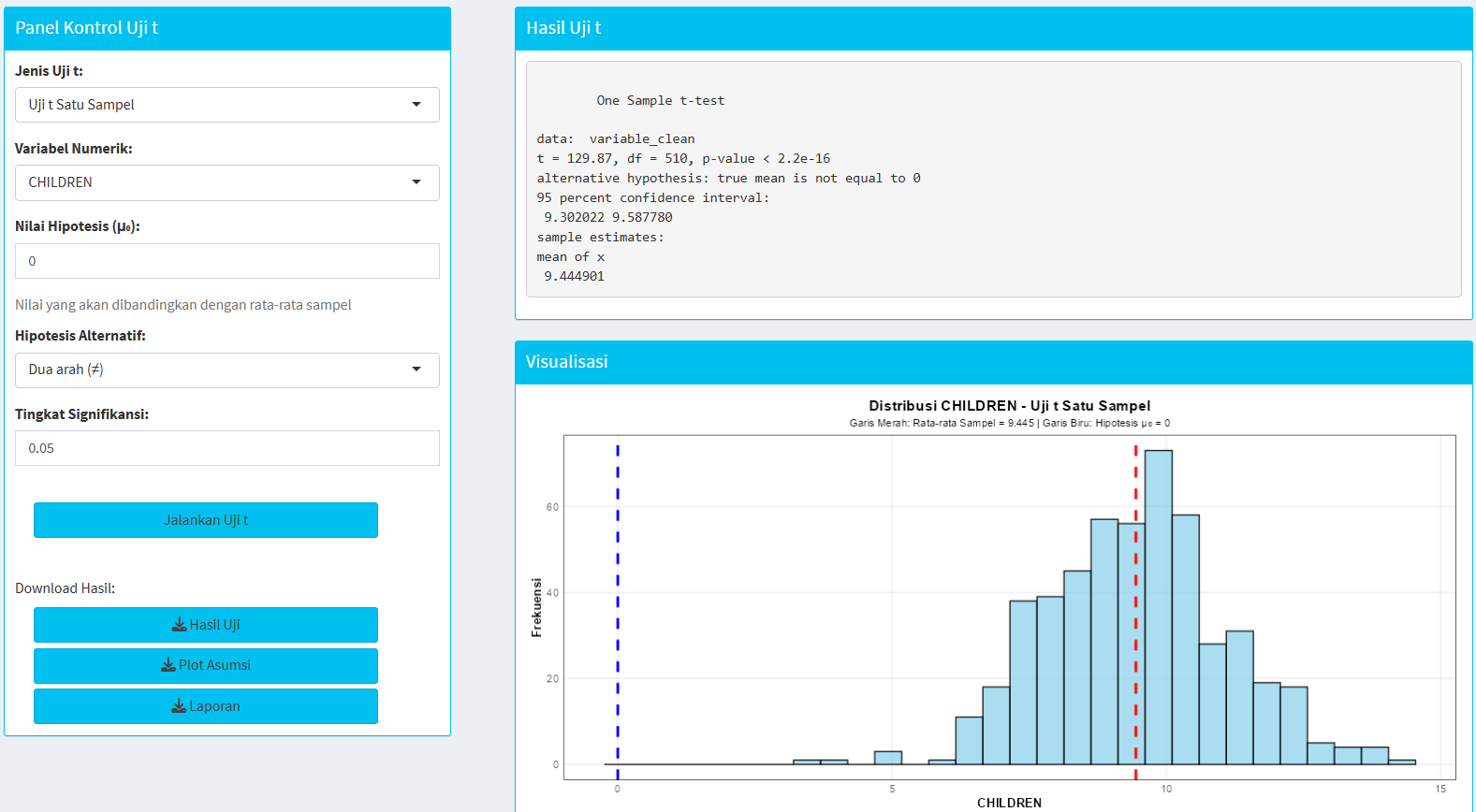
1. Menu Statistik Inferensia

Menu statistik inferensia menyediakan berbagai jenis pengujian hipotesis dan analisis perbandingan untuk menarik kesimpulan tentang populasi berdasarkan sampel data SOVI.

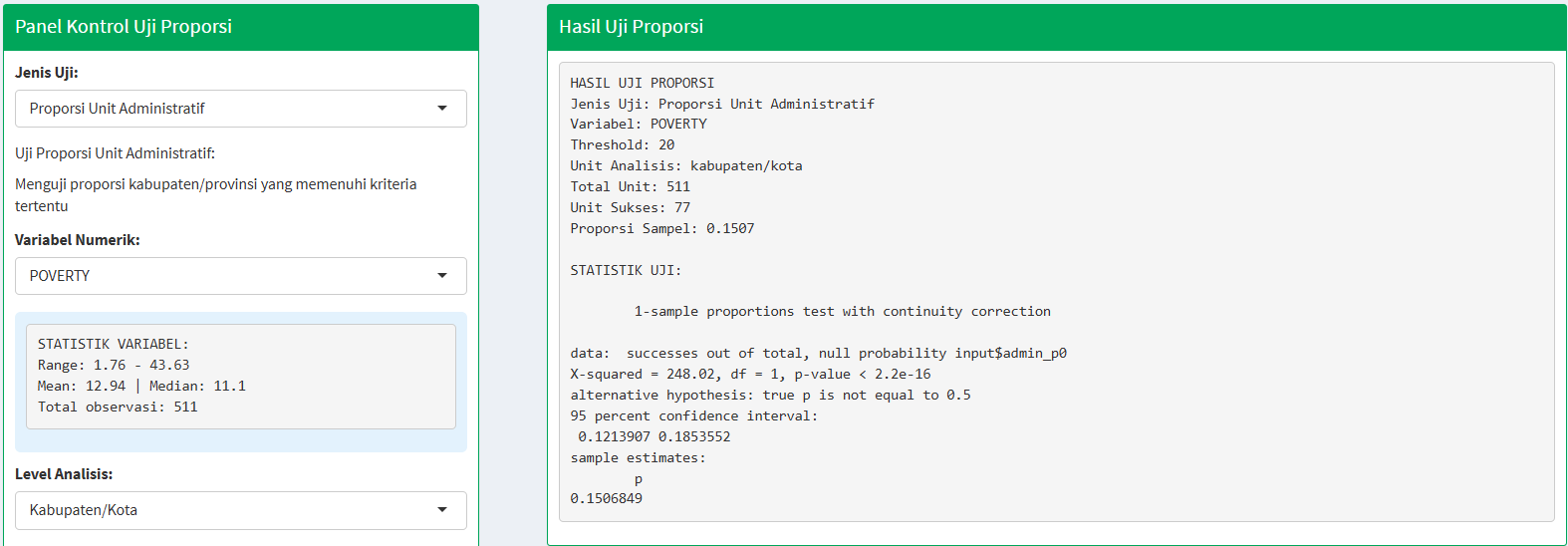


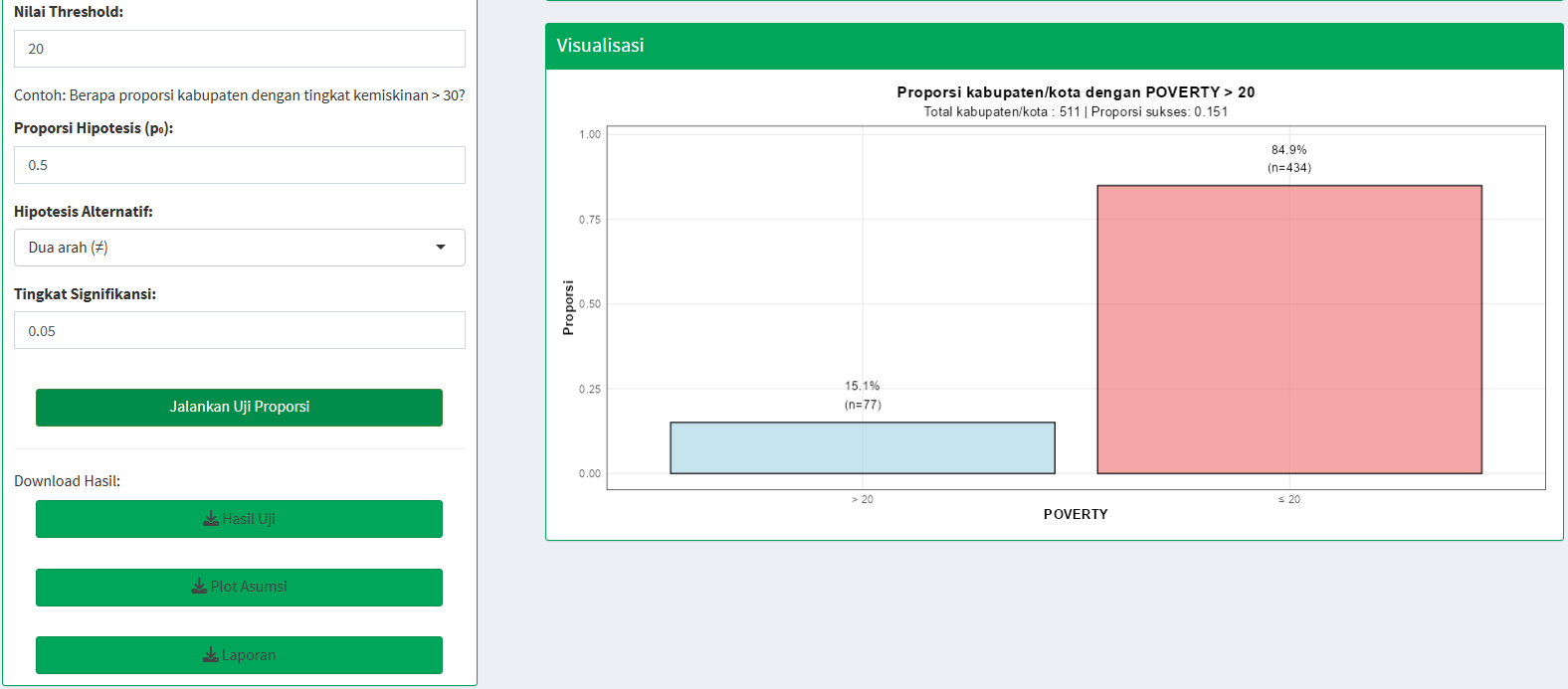
Fitur-fitur yang tersedia meliputi:

* Uji t Satu Sampel: One-sample t-test untuk membandingkan rata-rata sampel dengan nilai hipotesis tertentu
* Uji t Dua Sampel: Independent t-test untuk perbandingan rata-rata antar dua kelompok independent

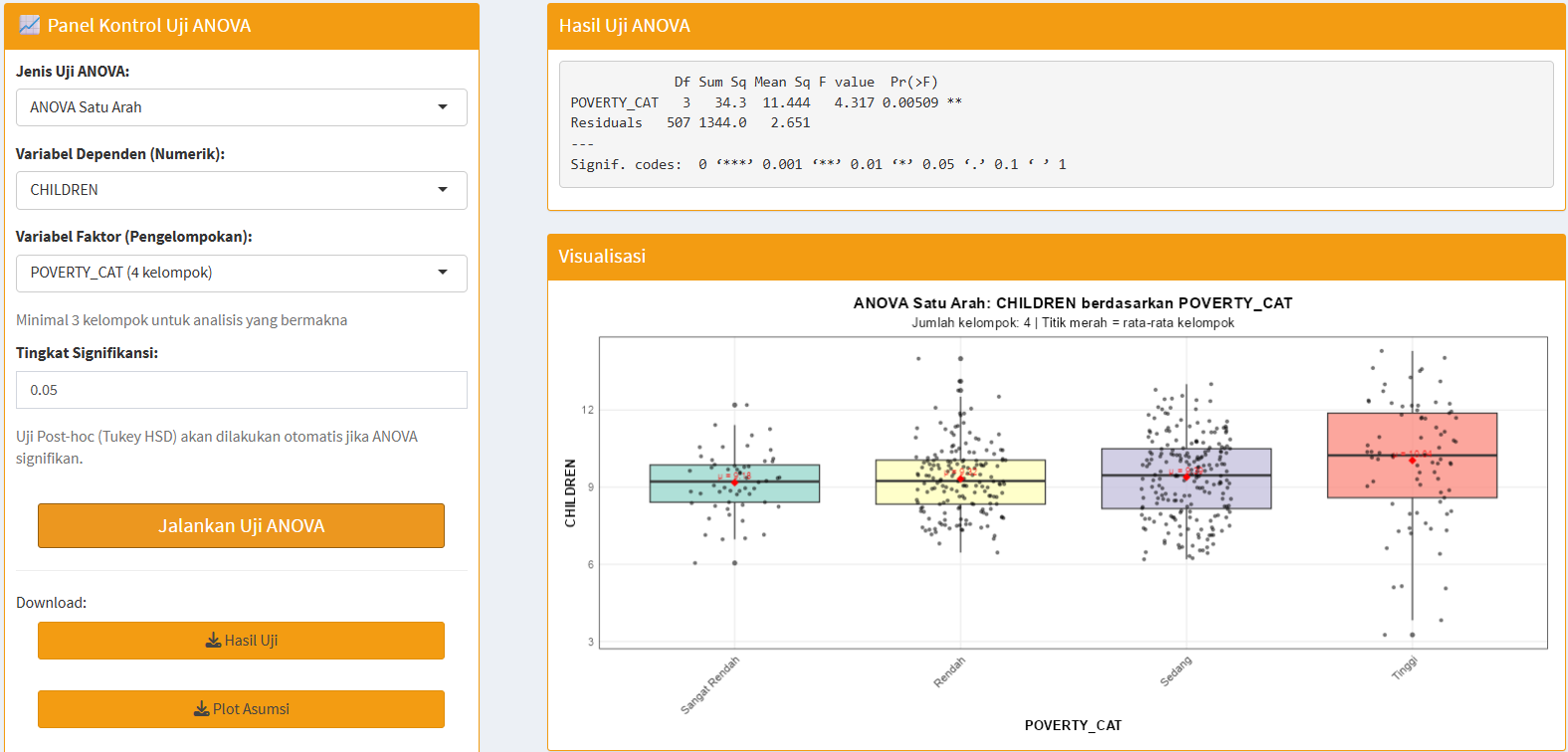


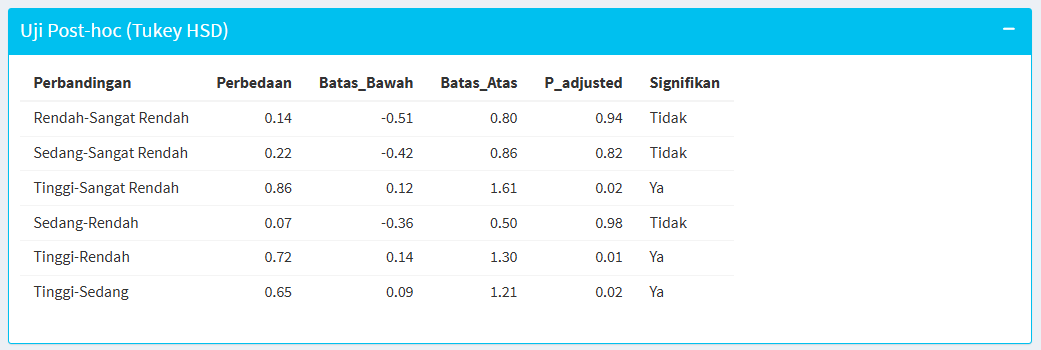
* Uji Proporsi Satu Sampel: One-sample proportion test untuk menguji proporsi populasi terhadap nilai hipotesis
* Uji Proporsi Dua Sampel: Two-sample proportion test untuk membandingkan proporsi antar dua kelompok



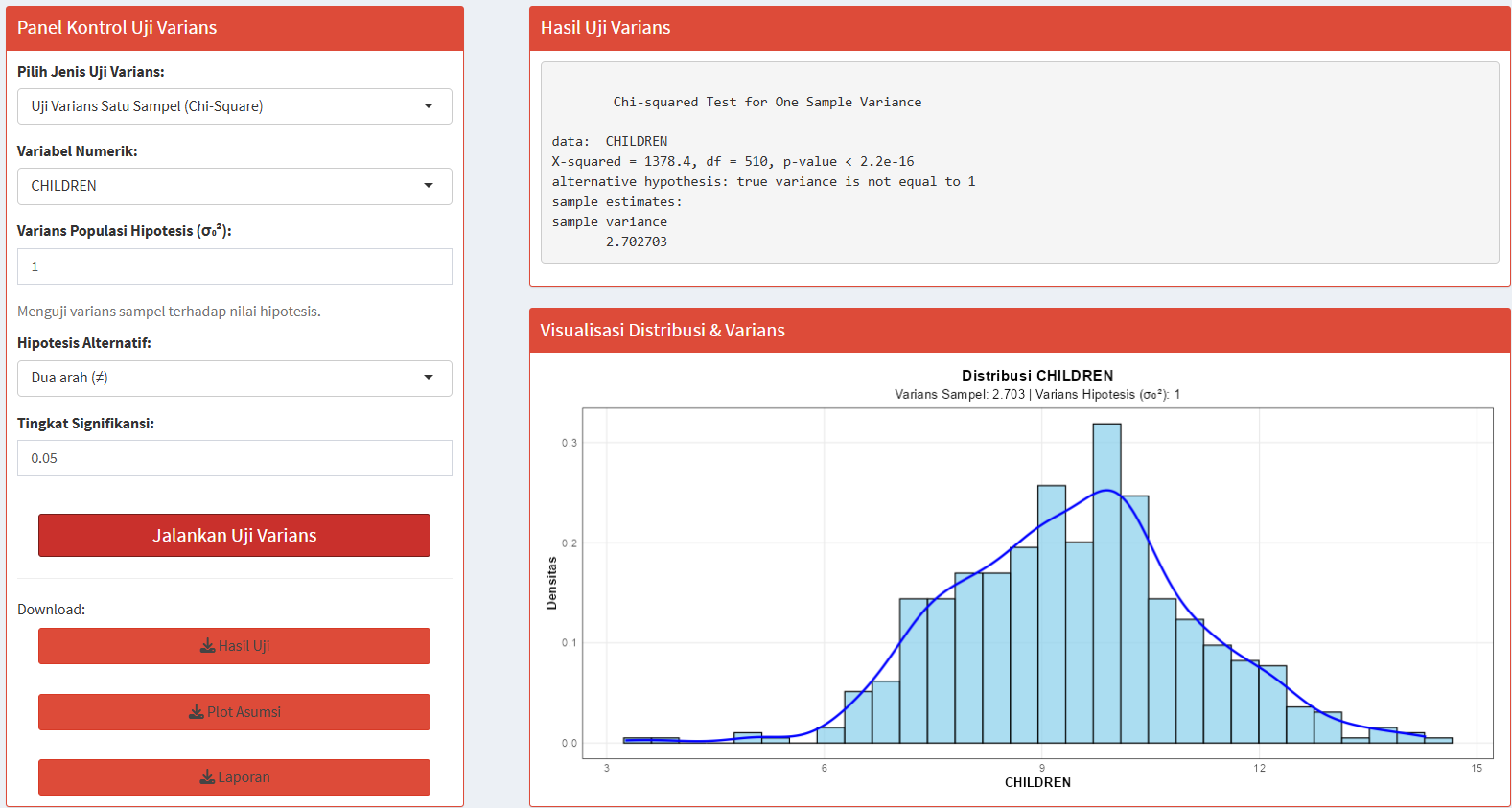


* ANOVA Satu Arah: Analisis varians untuk perbandingan rata-rata lebih dari dua kelompok
* ANOVA Dua Arah: Analisis varians dengan dua faktor untuk menguji efek utama dan interaksi





* Uji Varians Satu Kelompok: Chi-square test untuk menguji varians populasi terhadap nilai hipotesis
* Uji Varians Dua Kelompok: F-test untuk membandingkan varians antar dua kelompok



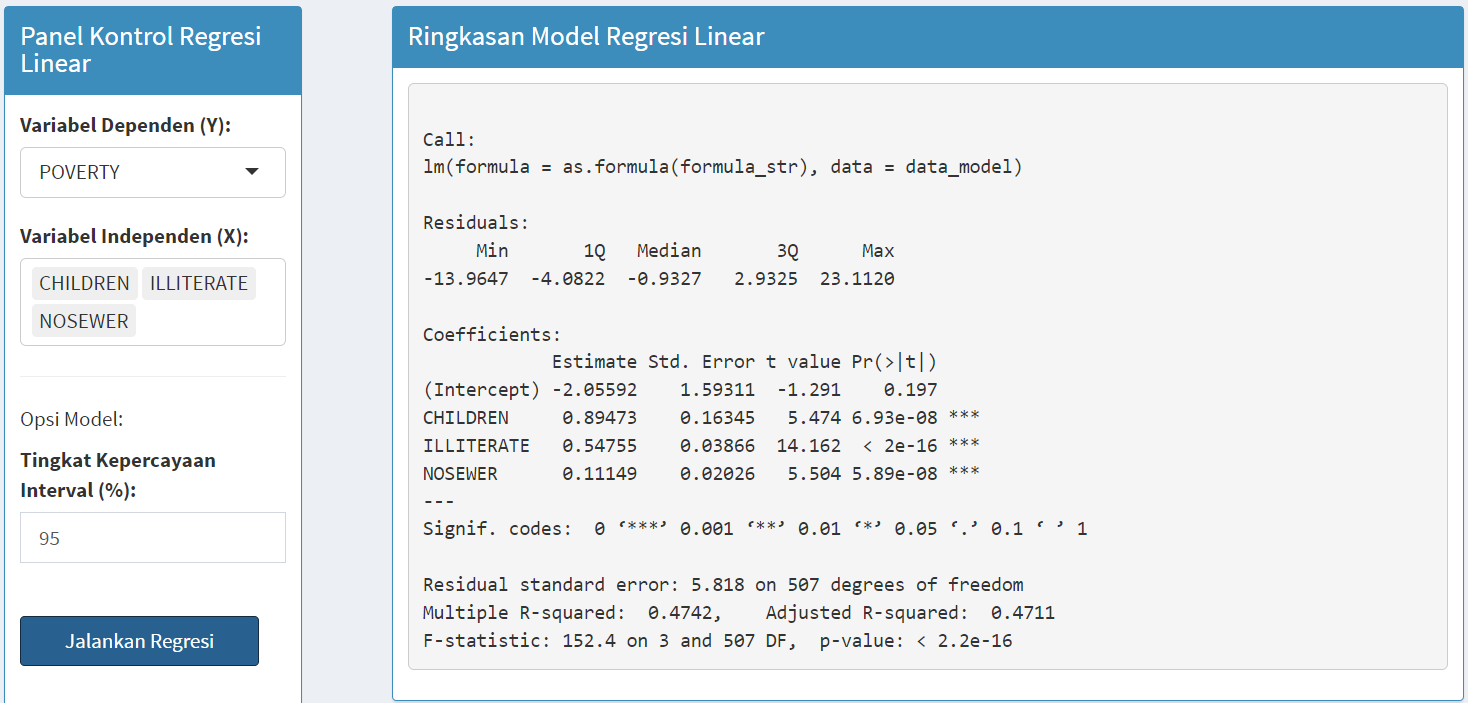
* Interpretasi Statistik: Penjelasan hasil uji dengan nilai p-value, confidence interval, dan kesimpulan hipotesis
* Download Hasil: Download output analisis dalam format tabel, visualisasi, dan laporan

1. Menu Regresi Liniear

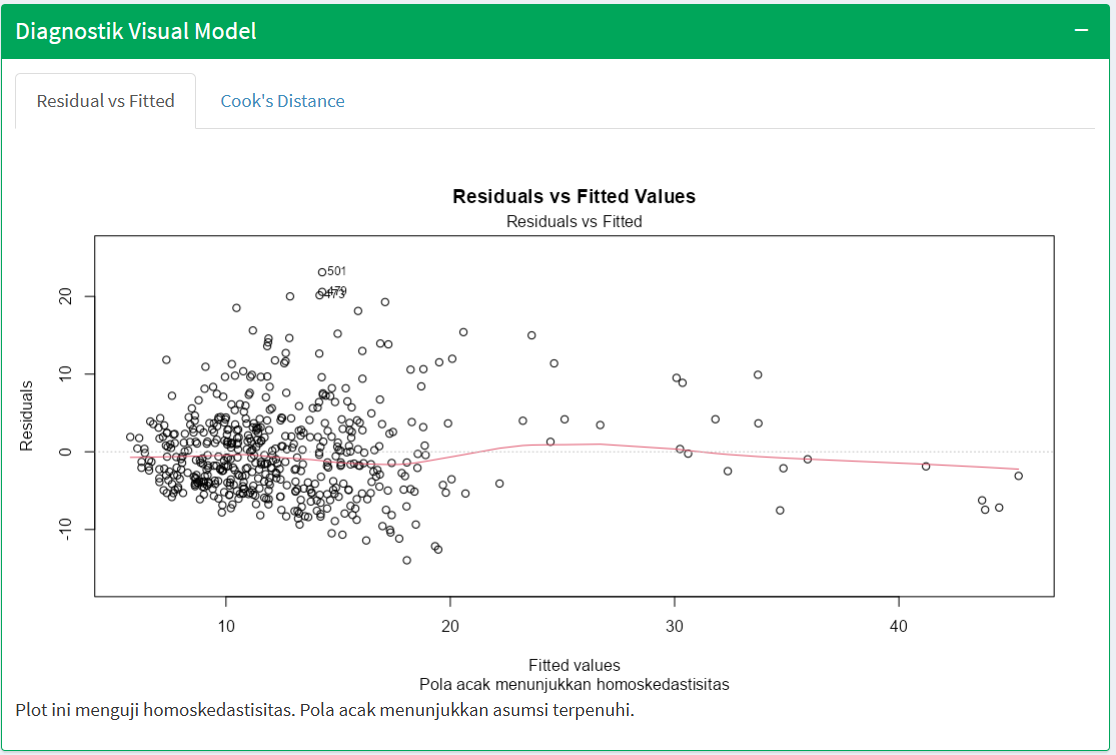
Menu regresi linear menyediakan fitur analisis pemodelan hubungan antar variabel untuk prediksi dan pemahaman pola dalam data SOVI. Menu ini memungkinkan pengguna membangun model regresi dari yang sederhana hingga kompleks dengan berbagai fitur diagnostik untuk memastikan validitas model.

Fitur-fitur yang tersedia meliputi:

* Regresi Linear Sederhana: Pemodelan hubungan antara satu variabel independen dengan satu variabel dependen
* Regresi Linear Berganda: Pemodelan dengan multiple variabel independen untuk prediksi yang lebih akurat

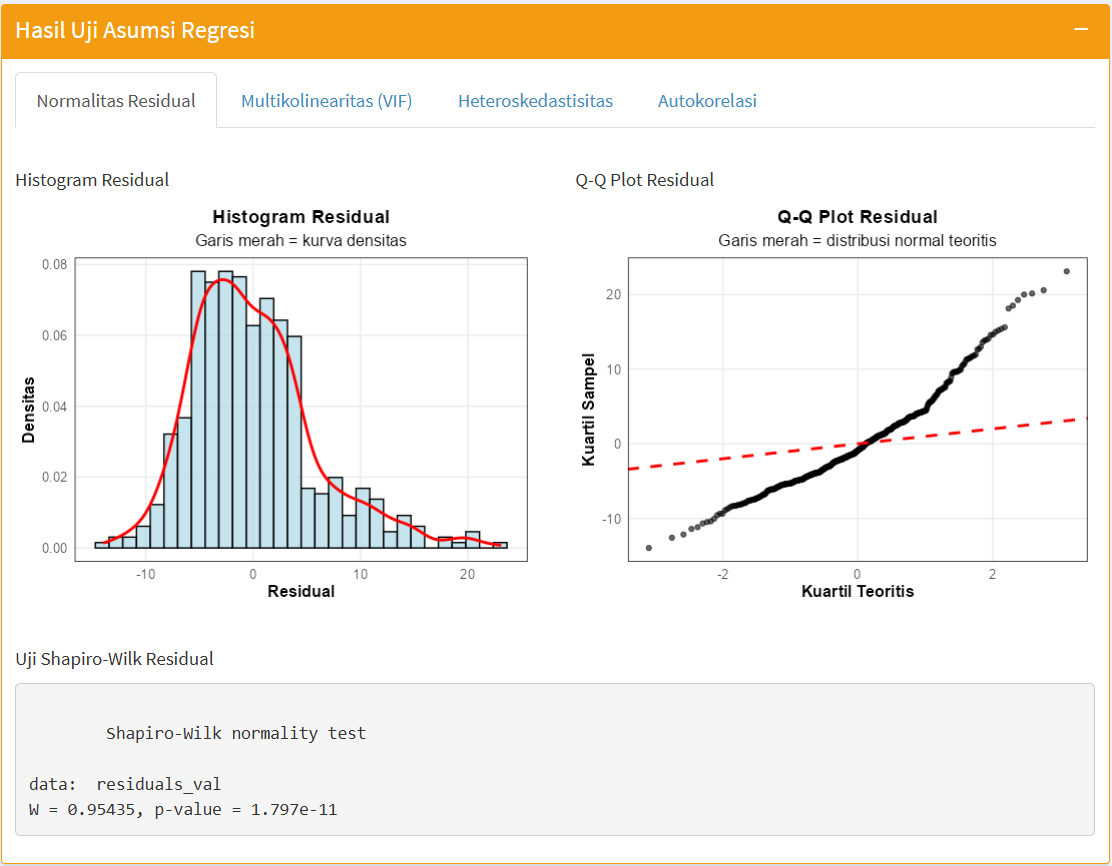


* Diagnostik Model: Analisis residual, uji linearitas, dan evaluasi goodness of fit



* Uji Asumsi Regresi:

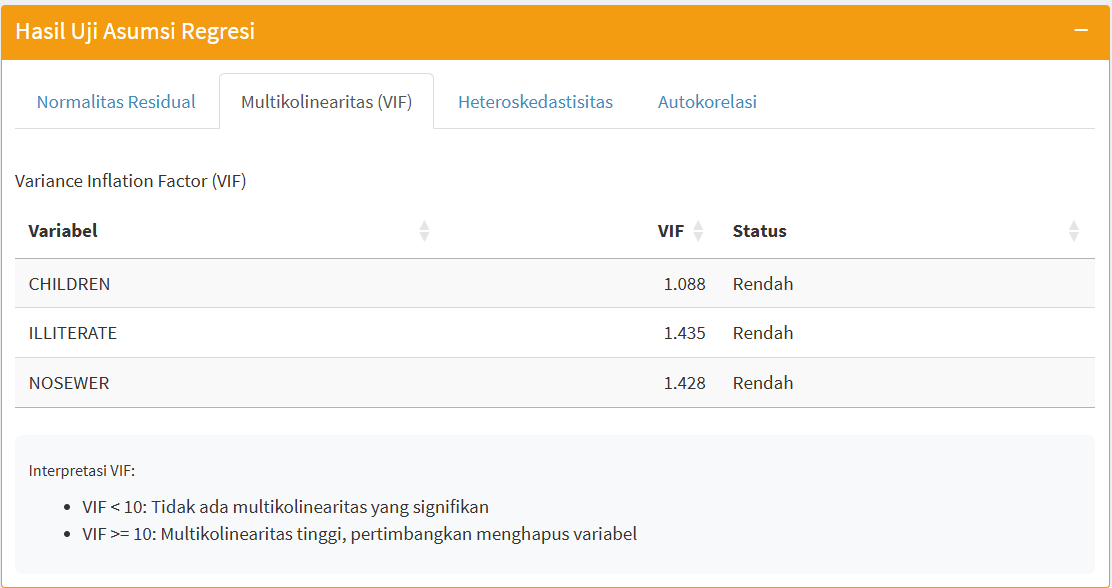
Pengujian normalitas residual



Homoskedastisitas



Multikolinearitas (VIF)



Autokorelasi



* Interpretasi Koefisien: Penjelasan makna dan signifikansi setiap koefisien regresi
* Download Model: Download hasil regresi, plot diagnostik, dan summary model