

KOMPUTASI STATISTIK

**SOVI EXPLORE : DASHBOARD INTERAKTIF UNTUK ANALISIS KERENTANAN
SOSIAL INDONESIA**



Disusun Oleh :

Sabila Bintang Kusuma Dewi

222313363

2KS2

Dosen Pengampu :

Yuliagnis Transver Wijaya, S.ST, M.Sc

PROGRAM STUDI D4 KOMPUTASI STATISTIK

POLITEKNIK STATISTIKA STIS

2024/2025

Dalam pengerjaan dashboard ini, dilakukan beberapa langkah, yaitu :

1. Identifikasi Kebutuhan Data

a. Masalah yang akan dianalisis :

- 1) Belum tersedianya platform terintegrasi untuk eksplorasi data dan analisis terhadap data kerentanan sosial (SOVI) tahun 2017 untuk seluruh kabupaten/kota di Indonesia
- 2) Tantangan dalam memahami distribusi dan karakteristik setiap indikator statistik secara komprehensif baik secara regional maupun nasional.
- 3) Belum diketahui perbedaan kondisi kerentanan sosial antar kelompok pulau secara statistik.
- 4) Belum diketahui faktor-faktor sosial ekonomi yang paling berpengaruh terhadap tingkat kemiskinan di Indonesia melalui model statistik yang interaktif.

b. Tujuan yang ingin dicapai :

- 1) Menyajikan gambaran umum tentang kerentanan sosial melalui statistik deskriptif dan visualisasi data yang informatif.
- 2) Menyediakan fitur visualisasi data berupa grafik dan peta untuk memudahkan interpretasi data.
- 3) Melakukan analisis statistik inferensia untuk mengetahui keterbandingan data antar kelompok.
- 4) Mengidentifikasi pengaruh variabel sosial-ekonomi terhadap tingkat kemiskinan menggunakan analisis regresi linear.

c. Solusi untuk mencapainya :

- 1) Mengembangkan dashboard interaktif menggunakan R-Shiny
- 2) Mengimplementasikan menu eksplorasi data yang menampilkan statistik deskriptif, visualisasi data, dan peta interaktif.
- 3) Menyediakan fitur uji asumsi untuk mengetahui normalitas dan homogenitas data sebagai kontrol kualitas data.

- 4) Mengintegrasikan fungsi untuk analisis inferensia, yaitu uji kesamaan rata-rata, uji proporsi, uji variance, dan ANOVA.
- 5) Menyediakan menu analisis regresi linear berganda yang dilengkapi uji asumsi dan interpretasi hasil secara otomatis.

2. Pengambilan Data

Data didapatkan dari github yang telah tersedia di soal UAS

Link github:

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

Metadata :

<https://www.sciencedirect.com/science/article/pii/S2352340921010180>

Data secara otomatis dimuat dalam dashboard dengan proses validasi dan cleaning yang terintegrasi. Dashboar melakukan preprocessing data secara otoamtis termasuk konversi tipe data dan pembersihan format. Data yang telah didapatkan melalui URL tersebut, kemudian di-*import* ke dalam dashboard R Shiny.

3. Integritas Data

Data berasal sumber yang terpercaya, yaitu dari data Survei Sosial dan Ekonomi Nasioanal tahun 2017 (SUSENAS 2017). Hal ini memastikan validitas dan keandalan data.

4. Telaah Data

a. Analisis tipe data

Dashboard secara otomatis mengidentifikasi tipe data dari tiap vairabel berdasarkan metadata yang tersedia.

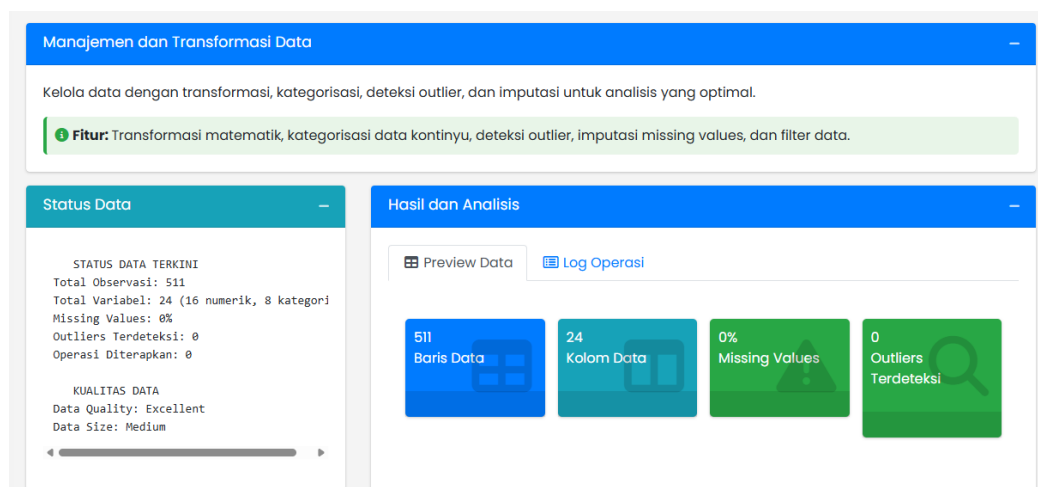
Label	Variabel	Tipe_Data	Deskripsi
DISTRICTCODE	Kode Distrik	String	Kode unik untuk wilayah/distrik di Indonesia
CHILDREN	Anak-anak	Numerik	Persentase populasi di bawah lima tahun

FEMALE	Perempuan	Numerik	Persentase populasi perempuan
ELDERLY	Lansia	Numerik	Persentase populasi 65 tahun ke atas
FHEAD	Kepala Rumah Tangga Perempuan	Numerik	Persentase rumah tangga dengan kepala rumah tangga perempuan
FAMILYSIZE	Anggota Rumah Tangga	Numerik	Rata-rata jumlah anggota rumah tangga dalam satu distrik
NOELECTRIC	Rumah Tangga Tanpa Listrik	Numerik	Persentase rumah tangga yang tidak menggunakan listrik sebagai sumber penerangan
LOWEDU	Pendidikan Rendah	Numerik	Persentase populasi 15 tahun ke atas dengan pendidikan rendah
GROWTH	Pertumbuhan Populasi	Numerik	Persentase perubahan populasi
POVERTY	Kemiskinan	Numerik	Persentase penduduk miskin
ILLITERATE	Buta Huruf	Numerik	Persentase populasi yang tidak bisa membaca dan menulis
NOTRAINING	Pelatihan Kesiapsiagaan Bencana	Numerik	Persentase rumah tangga yang tidak mendapatkan pelatihan bencana
DPRONE	Rentan Bencana	Numerik	Persentase rumah tangga yang tinggal di daerah rawan bencana

RENTED	Kepemilikan Rumah (Sewa)	Numerik	Persentase rumah tangga yang menyewa rumah
NOSEWER	Drainase	Numerik	Persentase rumah tangga yang tidak memiliki sistem drainase
TAPWATER	Sumber Air	Numerik	Persentase rumah tangga yang menggunakan air pipa
POPULATION	Populasi	Numerik	Jumlah total populasi
PROVINCE_NAME	Provinsi	String	Nama provinsi sesuai periode data
CITY_NAME	Kabupaten/Kota	String	Nama kabupaten/kota sesuai periode data

b. Pemeriksaan missing values

Dashboard melakukan pemeriksaan missing values secara otomatis untuk semua variabel. Hasil pemeriksaan menunjukkan missing values dan persentase kelengkapan data untuk setiap variabel. Hal ini akan ditampilkan pada menu manajemen data pada dashboard. Untuk data SOVI_2017, dapat dilihat bahwa tidak ada missing value, dengan kata lain data lengkap.



c. Pemeriksaan Outlier

Dashboard mengimplementasikan deteksi outlier menggunakan beberapa metode, yaitu :

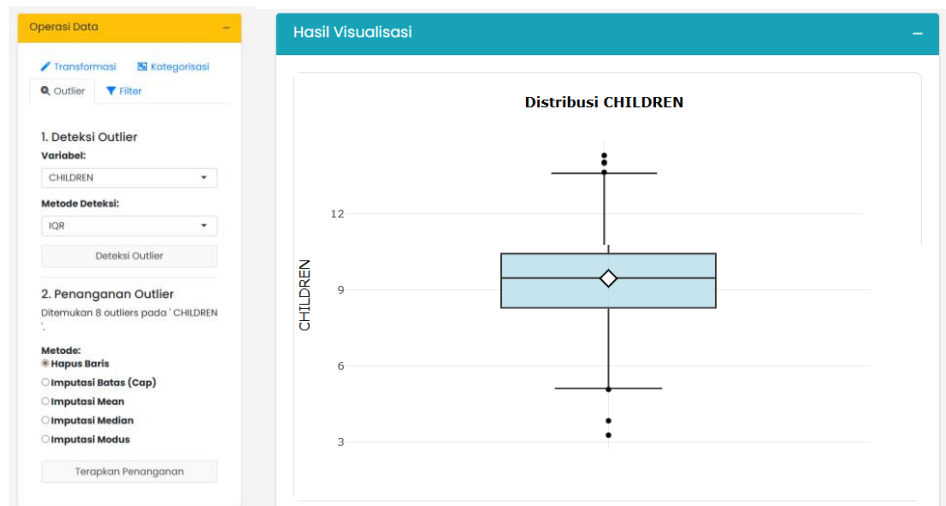
1) Metode IQR

Mengidentifikasi nilai yang berada di luar rentang $Q1 - 1.5IQR$ sampai $Q3 + 1.5IQR$.

2) Visualisasi Boxplot

Menyediakan representasi visual outlier untuk setiap variabel numerik dengan boxplot

Selanjutnya, dashboard ini menyediakan pilihan untuk menangani outlier, yaitu dengan menghapusnya atau mengisinya dengan median, mean, atau modus.



5. Validasi Data

Dashboard melakukan validasi data yang komprehensif untuk menjamin reliabilitas data. Validasi dilakukan secara bertahap. Pertama, validasi struktural dilakukan untuk memastikan konsistensi format, kewajaran rentang nilai (range) indikator, dan eliminasi data duplikat. Kedua, validasi geografis diterapkan untuk memverifikasi kesesuaian kode wilayah dengan standar BPS dan data referensi spasial, serta memastikan konsistensi hierarki administratif. Ketiga, validasi statistik mencakup analisis distribusi variabel, deteksi pencilan (outlier) yang tidak plausibel, dan pemeriksaan konsistensi hubungan antar-variabel yang berkorelasi secara teoretis. Kesimpulan dari proses validasi ini adalah dataset berada dalam kondisi prima, dengan kelengkapan dan konsistensi yang memadai untuk analisis statistik yang robust.

6. Menentukan Objek Data

Objek data yang digunakan dalam analisis ini adalah seluruh 511 kabupaten/kota yang terdapat dalam dataset `sovi_data.csv`. Analisis ini mencakup keseluruhan wilayah Indonesia untuk memberikan gambaran yang komprehensif. Namun, dasbor menyediakan fitur interaktif yang memungkinkan pengguna untuk memfilter dan memfokuskan analisis pada wilayah atau kelompok pulau tertentu sesuai kebutuhan. Untuk menentukan objek data, user dapat menggunakan menu Manajemen data pada dashboard ini.

Manajemen dan Transformasi Data

Kelola data dengan transformasi, kategorisasi, deteksi outlier, dan imputasi untuk analisis yang optimal.

Fitur: Transformasi matematik, kategorisasi data kontinyu, deteksi outlier, imputasi missing values, dan filter data.

Status Data

STATUS DATA TERKINI

Total Observasi: 511
Total Variabel: 24 (16 numerik, 8 kategorikal)
Missing Values: 0%
Outliers Terdeteksi: 0
Operasi Diterapkan: 0

KUALITAS DATA

Data Quality: Excellent
Data Size: Medium

Operasi Data

Transformasi
Kategorisasi
Outlier
Filter

Variabel:
CHILDREN

Metode:
Logaritma (log)

Suffix nama variabel baru:
Contoh: _log

Terapkan Transformasi

Kontrol Umum

Reset Semua Perubahan

Export Data (.csv)

Hasil dan Analisis

Preview Data Log Operasi

511 Baris Data 24 Kolom Data 0% Missing Values 0 Outliers Terdeteksi

Show 5 entries Search:

	DISTRICTCODE	CHILDREN	FEMALE	ELDERLY	
	All	All	All	All	All
1	1101	7.999956086	48.77590901	2.184700509	13
2	1102	13.51717522	49.68611677	2.296480805	13.
3	1103	9.43677942	50.77554271	4.903403386	20.7
4	1104	11.18920899	50.09882757	2.735836322	17.7
5	1105	11.68376732	50.048859	2.756701347	19.4

Showing 1 to 5 of 511 entries

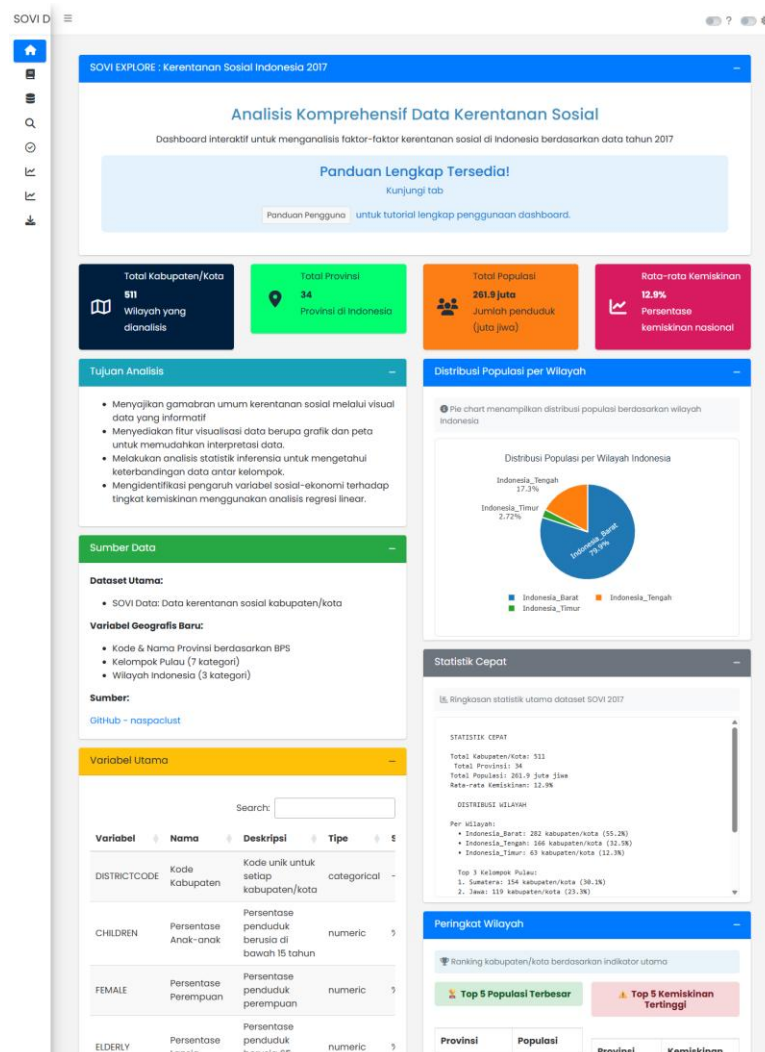
Previous 1 2 3 4 5 ... 103 Next

7. Membuat Bussiness Intelligence

Terdapat beberapa fitur-fitur pada dahsboard ini, yaitu :

a. Halaman Beranda

Halaman beranda berfungsi sebagai tampilan awal bagi pengguna. Halaman ini menyajikan deskripsi singkat mengenai tujuan dashboard, yaitu menyajikan alat analisis interaktif terhadap data kerentanan sosial di Indoneia. Sleain itu, halaman ini juga memberikan informasi mengenai sumber data yang digunakan (SUSENAS 2017), sehingga memberikan konteks dan kredibilitas pada analisis yang disajikan. Selain itu, pada beranda juga terdapat informasi komprehensif mengenai kerentanan sosial di Indonesia, mulai dari metadata lengkap dataset SUSENAS 2017, deskripsi rinci setiap variabel, ringkasan statistik utama, hingga grafik visual yang menggambarkan distribusi tingkat kerentanan di seluruh wilayah.



b. Menu Manajemen Data

Menu Manajemen Data merupakan komponen fundamental dalam alur kerja analisis data, berfungsi sebagai jembatan antara data mentah dan analisis yang bermakna. Dashboard ini menyediakan serangkaian alat interaktif yang memungkinkan pengguna untuk melakukan pra-pemrosesan, transformasi, dan penyiapan data secara mandiri tanpa memerlukan penulisan kode.

- **Transformasi Variabel**

Modul Transformasi Variabel ini memungkinkan pengguna untuk menerapkan fungsi matematis, seperti logaritma atau akar kuadrat, pada variabel numerik. Proses ini sangat penting untuk memperbaiki sebaran data yang miring dan menstabilkan varians, sehingga memenuhi asumsi-asumsi dasar untuk analisis statistik yang lebih lanjut.

- **Kategorisasi Variabel**

- **Kategorisasi Variabel Kontinu**

Fitur ini memungkinkan pengguna untuk mengubah variabel numerik (kontinu) menjadi variabel kategorik. Contohnya, variabel POVERTY dapat diubah menjadi tiga kategori: 'Rendah', 'Sedang', dan 'Tinggi'. Ini sangat esensial untuk analisis tabulasi silang atau sebagai variabel pengelompokan dalam uji statistik.

- **Pembuatan Variabel Baru**

Pengguna dapat membuat kolom baru berdasarkan kalkulasi dari variabel yang sudah ada. Misalnya, membuat variabel rasio 'Lansia terhadap Anak-anak' dengan membagi variabel ELDERLY dengan CHILDREN.

- **Penanganan Outlier**

Meskipun dataset sumber yang digunakan dalam proyek ini bersih, dasbor yang robust harus memiliki kapabilitas untuk menangani data yang hilang. Fitur ini menyediakan beberapa opsi standar untuk menangani *missing values*, seperti:

- **Seleksi Variabel**

Studi Kasus : Melakukan pengkategorian data presentase jumlah anak menjadi Rendah, Sedang, Tinggi, berdasarkan kuantilnya.

	PROVINCE_NAME	ISLAND_GROUP	REGION	POVERTY_cat
All	All	All	All	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Sedang	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Tinggi	
Aceh	Sumatera	Indonesia_Barat	Rendah	
Aceh	Sumatera	Indonesia_Barat	Tinggi	

c. Menu Eksplorasi Data

Menu ini merupakan jantung dari analisis deskriptif, menyediakan berbagai alat untuk menelaah dan memahami data secara mendalam.

- **Tampilan Data**

Fitur ini menyajikan keseluruhan dataset dalam bentuk tabel interaktif. Pengguna dapat melakukan pencarian, penyortiran (sorting), dan paginasi (mengatur jumlah data yang ditampilkan per halaman).

SOVI D

Statistik Deskriptif

Halaman ini menyajikan ringkasan statistik untuk setiap variabel numerik. Gunakan ini untuk memahami tendensi sentral, sebaran, dan bentuk distribusi data Anda.

Tabel Statistik Deskriptif

Show 10 entries Search:

Variabel	N	Mean	Median	SD	Min	Max
CHILDREN	511	9.44	9.46	1.64	3.26	14.3
FEMALE	511	49.49	49.54	1.44	44.63	59.38
ELDERLY	511	4.5	4.08	2.15	0	12.64
FHEAD	511	14.27	13.95	4.92	0.04	31.66
FAMILYSIZE	511	3.96	3.92	0.41	2.92	5.77
NOELECTRIC	511	6.16	0.85	15.03	0	97.1
LOWEDU	511	32.45	33.3	9.22	9.3	55.42
GROWTH	511	1.41	1.25	0.94	0.01	7.19
POVERTY	511	12.94	11.1	8	1.76	43.63
ILLITERATE	511	8.32	6.37	7.98	1.54	66.1

Showing 1 to 10 of 16 entries Previous 1 2 Next

Unduh Tabel

Pilih Format:

☒ PDF ☐ Word

Unduh

Interpretasi Hasil

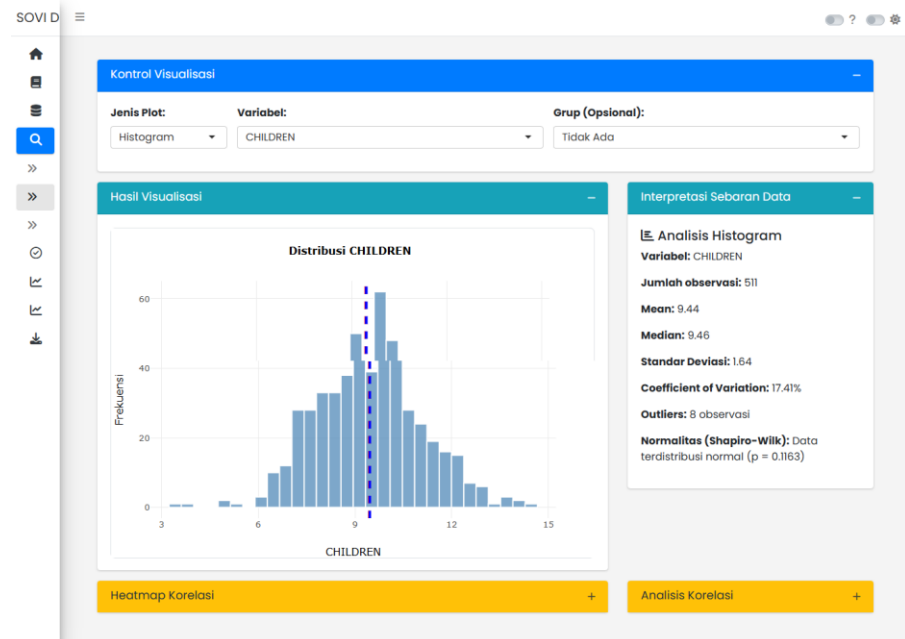
Interpretasi:

- Mean & Median: Perbandingan nilai mean dan median dapat memberikan indikasi awal kemiringan (skewness) distribusi data.
- Standar Deviasi (SD): Nilai SD yang besar menandakan sebaran data yang luas.

- **Visualisasi Data:**

- **Histogram:** Fitur ini memungkinkan pengguna untuk membuat histogram dari variabel numerik mana pun. Histogram berguna untuk memvisualisasikan bentuk distribusi data, mengidentifikasi apakah sebarannya simetris, miring (skewed), atau memiliki modus ganda.

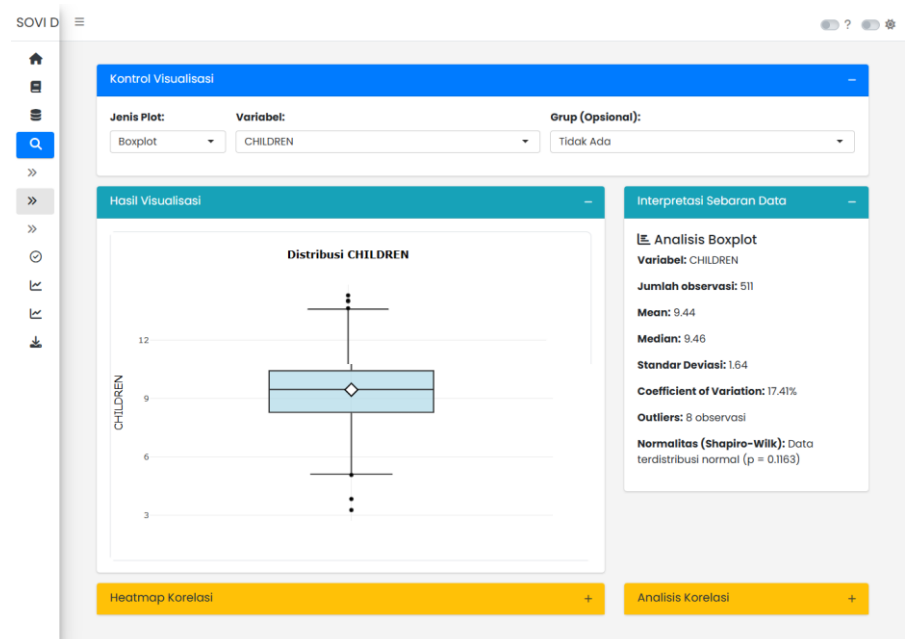
Studi kasus : Melakukan pembuatan histogram untuk variabel children.



Berdasarkan grafik tersebut, terlihat bahwa variabel children berdistribusi normal.

- **Boxplot:** Boxplot memberikan ringkasan lima angka (minimum, Q1, median, Q3, maksimum) dan sangat efektif dalam mengidentifikasi adanya nilai pencilan (outlier) dalam data.

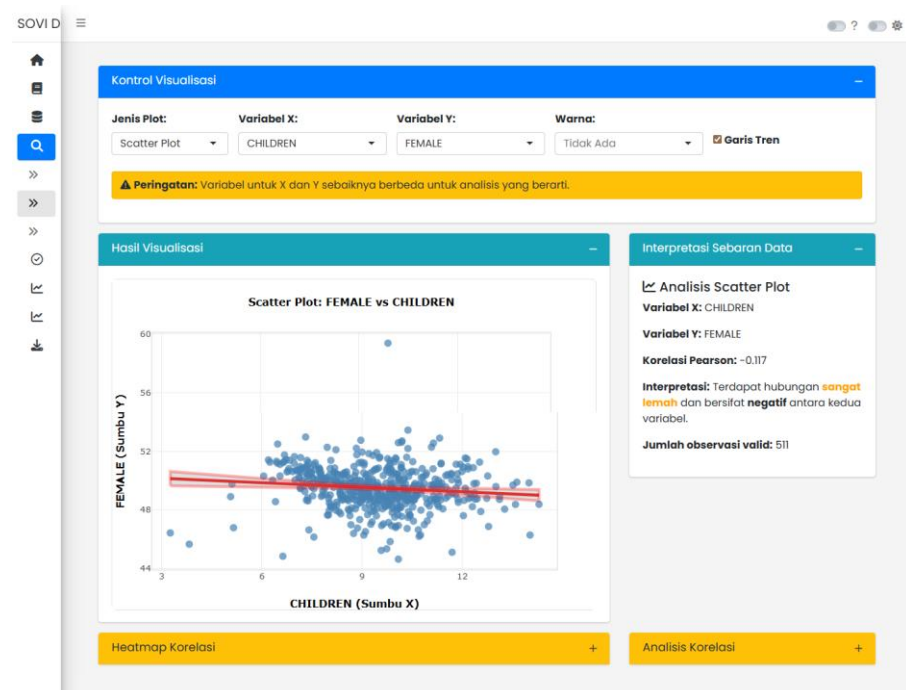
Studi kasus : membuat boxplot untuk variabel children untuk mendeteksi outlier



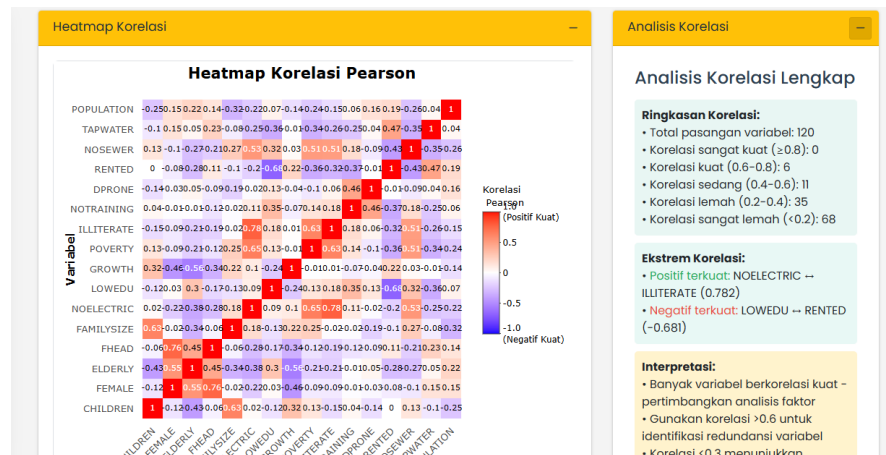
Berdasarkan grafik tersebut, diketahui bahwa terdapat outlier pada data children, tetapi untuk kasus ini tidak diberi perlakuan apapun karena variabel tersebut sudah normal.

- **Scatter Plot:** Fitur ini memungkinkan pengguna untuk memvisualisasikan hubungan antara dua variabel numerik. Pengguna dapat memilih variabel X, variabel Y, dan juga warna yang diinginkan. Fitur ini juga sangat membantu pengguna untuk melakukan indentifikasi hubungan antarvariabel sebelum membentuk model regresi.

Studi kasus : Scatter Plot antara Female dan Children

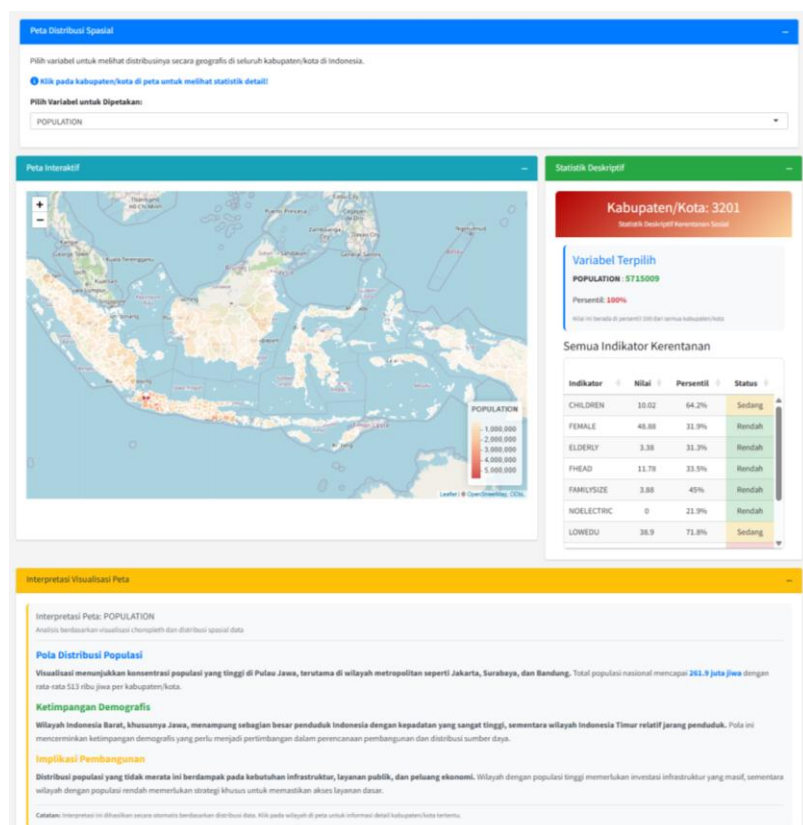


- **Matriks Korelasi:** Fitur ini menghasilkan peta panas (heatmap) dari matriks korelasi antara variabel-variabel numerik yang dipilih. Fitur ini adalah alat analisis data eksploratif yang fundamental dan sangat kuat, yang diimplementasikan dalam dasbor ini untuk memberikan wawasan mendalam mengenai hubungan antar variabel. Berbeda dengan *scatter plot* yang hanya membandingkan dua variabel, matriks korelasi memungkinkan pengguna untuk mengevaluasi jaringan hubungan antara banyak variabel numerik secara serentak dalam satu visualisasi yang ringkas.



- Peta Tematik:** Ini adalah salah satu fitur visualisasi paling kuat. Berbeda dengan grafik konvensional yang menampilkan data secara abstrak, peta tematik mampu menyajikan data dalam konteks geografisnya yang nyata. Hal ini memungkinkan pengguna untuk tidak hanya melihat nilai dari sebuah variabel, tetapi juga lokasi dan distribusi spasialnya di seluruh wilayah Indonesia. Tab ini juga memberikan fitur interpretasi yang sangat memudahkan pengguna untuk menganalisis data.

Studi kasus : Peta presentase populasi penduduk di Indonesia.



d. Uji Asumsi Data

Fitur Uji Asumsi Data merupakan komponen fundamental dalam dasbor ini, yang dirancang untuk memvalidasi prasyarat statistik sebelum melakukan analisis parametrik seperti T-test atau ANOVA. Fitur ini menyediakan dua jenis pengujian:

- 1) **Uji Normalitas:** Untuk mengevaluasi apakah sebaran data dari suatu variabel mengikuti distribusi normal, dasbor ini mengadopsi pendekatan ganda yang komprehensif. Secara kuantitatif, **Uji Shapiro-Wilk** diimplementasikan untuk menghasilkan nilai p-value sebagai dasar pengambilan keputusan statistik. Sebagai pelengkap visual, **grafik Q-Q (Quantile-Quantile)** juga disediakan, memungkinkan pengguna untuk menilai normalitas secara intuitif dengan mengamati kedekatan titik-titik data terhadap garis diagonal referensi.
- 2) **Uji Homogenitas:** Ketika melakukan perbandingan antar kelompok, asumsi kesamaan varians menjadi krusial. Untuk memvalidasi asumsi ini, dasbor menggunakan **Uji Levene (Levene's Test)**, sebuah metode statistik yang robust untuk menentukan apakah varians dari suatu variabel sama atau berbeda secara signifikan di antara kelompok-kelompok yang sedang dibandingkan.

Pada halaman ini juga terdapat fitur untuk mendownload hasil analisisnya. Apabila data atau variabel yang digunakan user tidak memenuhi asumsi kenormalan dan homogenitas, user dapat melakukan transformasi data di halaman Manajemen Data.

Studi Kasus : Melihat Normalitas dan Kehomogenan Variabel Children.

[illegible]

3) Uji Variance

Fitur ini memungkinkan pengguna untuk menguji apakah terdapat perbedaan yang signifikan dalam hal varians (sebaran data) antara dua atau lebih kelompok. Terdapat fitur untuk menguji variance satu populasi dan dua populasi yang saling independen. Pada halaman ini terdapat fitur untuk mendownload hasil analisis yang telah dilakukan.

Uji Varians

Uji ini digunakan untuk menguji varians (ragam) dari satu populasi atau membandingkan varians dua populasi.

Catatan:

- Uji 1 kelompok menggunakan Chi-square test, sedangkan uji 2 kelompok menggunakan F-test.
- Asumsi utama untuk uji ini adalah data harus berdistribusi normal.

Pengaturan Uji

Pilih Jenis Uji:

Satu Kelompok (Chi-square)

Pilih Variabel Numerik:

CHILDREN

Varians Hipotesis (σ^2_0):

3

Hipotesis Alternatif:

Dua Sisi (Two-sided)

Tingkat Signifikansi (α):

0.05

Jalankan Uji Variance

Hasil Analisis Uji Varians

Ringkasan Hasil Uji Statistik

=== HASIL UJI VARIANS ===

Uji: Chi-square test for variance

$H_0: \sigma^2 = 3$

$H_1: \sigma^2 \neq 3$

statistik Chi-square: 459.4595

Derajat bebas: 510

P-value: 0.105589

Statistik Deskriptif

Statistik	Nilai
N	511
Mean	9.4449
Variance	2.7027
Std Dev	1.644

Interpretasi, Visualisasi, dan Laporan

Interpretasi Uji

Visualisasi

=== INTERPRETASI HASIL UJI STATISTIK ===

KEPUTUSAN: Karena p-value (0.1059) > dari α (0.05), maka H_0 GAGAL DITOLAK.

KESIMPULAN: Dengan tingkat kepercayaan 95%, data tidak memberikan cukup bukti untuk menyatakan bahwa varians populasi (σ^2) berbeda dari nilai hipotesis (3.00).

Unduh Laporan Lengkap

Unduh semua hasil analisis (teks dan gambar) dalam satu dokumen.

Pilih Format:

PDF Word

Unduh Laporan

4) ANOVA

Sebuah uji statistik parametrik yang kuat untuk membandingkan rata-rata dari suatu variabel di antara lebih dari dua kelompok. Pengguna dapat menggunakan fitur ini jika data telah memenuhi asumsi normalitas dan homogenitas varians. Selain ANOVA satu arah, fitur ini juga menyediakan pengujian ANOVA dua arah. Selain menyediakan hasil uji ANOVA, tab ini juga menyediakan interpretasi data hasil pengujian ANOVA dan juga visualisasinya.

Analysis of Variance (ANOVA)

ANOVA digunakan untuk membandingkan rata-rata dari tiga atau lebih kelompok.

Pastikan asumsi **normalitas** dan **homogenitas variansi** terpenuhi sebelum melakukan analisis.

Tips Penggunaan

- One-Way: 1 faktor, 3+ grup
- Two-Way: 2 faktor + interaksi
- Cek asumsi di tab 'Uji Asumsi'

Pengaturan ANOVA

Pilih Jenis ANOVA:

Satu Arah (One-Way)

Pilih Variabel Dependen (Numerik):

CHILDREN

Pilih Variabel Independen 1 (Kategorik):

POVERTY_LEVEL

Jalankan ANOVA

Asumsi ANOVA

- Normalitas:** Data berdistribusi normal
- Homogenitas:** Variansi antar grup sama
- Independensi:** Observasi saling independen

Hasil Analisis ANOVA

Tabel ANOVA

Interpretasi

Visualisasi

=== HASIL ANALISIS ANOVA ===

Formula yang digunakan: CHILDREN ~ POVERTY_LEVEL

Jumlah observasi: 511

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
POVERTY_LEVEL	1	16.3	16.309	6.095	0.0139 *
Residuals	509	1362.1	2.676		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=== RINGKASAN MODEL ===

R-squared: 0.0118

Adjusted R-squared: 0.0099

Pada tab ini terdapat juga uji tukey HSD untuk mengetahui kelompok mana yang memiliki rata – rata berbeda dan interpretasinya. Selain menyajikan hasil uji Tukey HSD, submenu ini juga menyajikan interpretasi serta visualisasi dari hasil pengujian yang telah dilakukan.

Uji Post Hoc Tukey HSD

Hasil ANOVA menunjukkan perbedaan signifikan antar grup. Uji post hoc Tukey HSD akan mengidentifikasi grup mana yang berbeda secara spesifik.

Uji Post Hoc Tukey HSD

Metode: Tukey HSD
Metode paling umum dan andal untuk perbandingan multiple

Jalankan Tukey HSD

Hasil

Interpretasi

Visualisasi

=== HASIL UJI POST HOC ===

Metode: Tukey HSD

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = as.formula(formula_str), data = data_for_anova)

\$POVERTY_LEVEL

	diff	lwr	upr	p adj
Tinggi-Rendah	0.3573816	0.07295825	0.641645	0.0138867

f. Menu Analisis Regresi

Menu ini merupakan fitur analisis paling canggih dalam dasbor, yang memungkinkan pengguna untuk membangun model regresi linear berganda. Pengguna dapat memilih satu variabel dependen dan beberapa variabel independen untuk memodelkan hubungan di antara mereka. Dasbor akan menampilkan ringkasan hasil model, termasuk nilai koefisien, standar error, t-statistik, p-value, dan nilai R-squared, yang sangat penting untuk interpretasi kekuatan dan signifikansi pengaruh setiap variabel.

Regresi Linear Berganda

Bangun model regresi untuk memprediksi sebuah variabel dependen (Y) berdasarkan beberapa variabel independen (X).

Tips: Pastikan hubungan linear antar variabel dan periksa asumsi regresi pada tab 'Uji Asumsi'.

Pengaturan Model Regresi

Pilih Variabel Dependen (Y):
CHILDREN

Pilih Variabel Independen (X):
FEMALE LOWEDU FAMILYSIZE ILLITERATE POVERTY

Bangun Model

Status: Siap untuk membangun model regresi

Hasil Model Regresi

Ringkasan Model

Uji Asumsi

Interpretasi

Cara Membaca Hasil:

- R-squared: Persentase variabilitas Y yang dijelaskan model
- P-value koefisien < 0.05: Variabel berpengaruh signifikan
- Estimate: Perubahan Y per unit perubahan X

Regression model created successfully

Call:
lm(formula = as.formula(formula_str), data = processed_data\$current)

Residuals:

	Min	1Q	Median	3Q	Max
	-5.2338	-0.7921	-0.0225	0.6474	5.7387


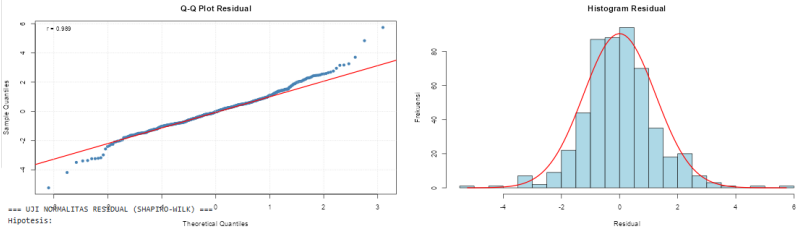
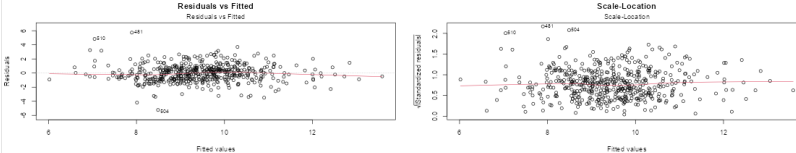
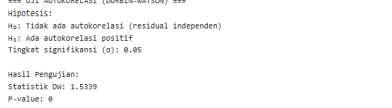
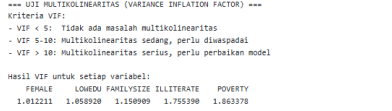

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.558361	1.997389	3.288	0.001111 **
FEMALE	-0.128806	0.038337	-3.339	0.000903 ***
LOWEDU	-0.063225	0.006143	-0.525	0.599752
FAMILYSIZE	2.378147	0.144248	16.487	< 2e-16 ***
ILLITERATE	-0.044046	0.009128	-4.828	1.9e-06 ***
POVERTY	0.022382	0.009393	2.383	0.017556 *

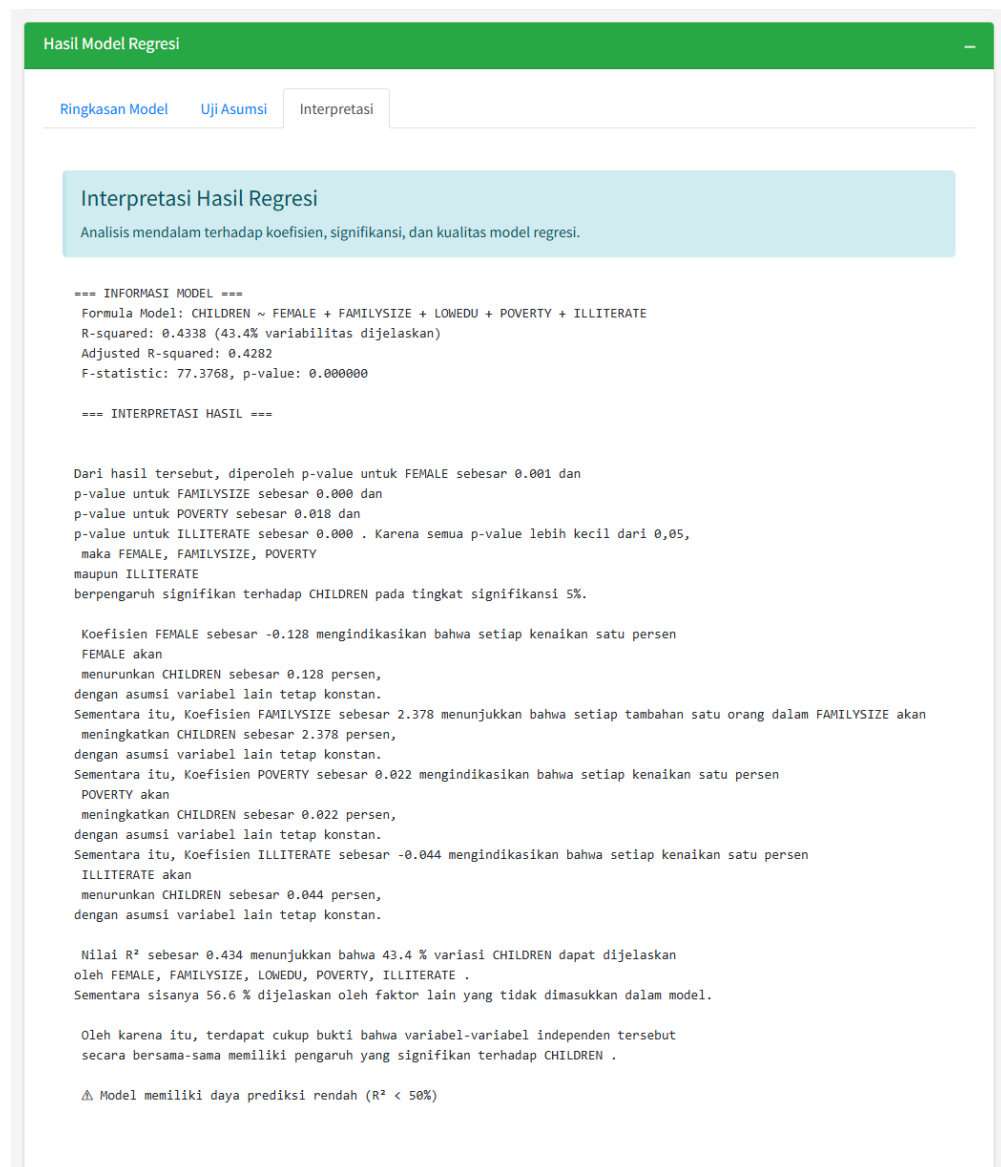
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.243 on 585 degrees of freedom
Multiple R-squared: 0.4338, Adjusted R-squared: 0.4282
F-statistic: 77.38 on 5 and 585 Df, p-value: < 2.2e-16

Lalu terdapat juga tab untuk menguji apakah model yang dihasilkan memenuhi asumsi klasik Gauss-Marcoff.

Hasil Model Regresi		
Ringkasan Model	Uji Asumsi	Interpretasi
<div> <div>  Asumsi Regresi Linear </div> <div> Model regresi linear memiliki asumsi utama yang harus dipenuhi untuk validitas hasil: </div> </div>		
<div> <div> 1. Normalitas Residual </div> <div> Residual harus berdistribusi normal. Gunakan Q-Q plot dan uji Shapiro-Wilk. </div> <div>  </div> <div> <pre> === UJI NORMALITAS RESIDUAL (SHAPIRO-WILK) === H0: Residual berdistribusi normal H1: Residual tidak berdistribusi normal Tingkat signifikansi (α): 0.05 Hasil Pengujian: Statistik W: 0.978698 P-value: 1e-06 === INTERPRETASI === Berdasarkan uji Shapiro-Wilk yang telah dilakukan, diketahui bahwa residual model regresi TIDAK berdistribusi normal karena nilai p-value (0.000001) < α (0.05), sehingga keputusan yang diambil adalah TOLAK H0. Kesimpulan: Asumsi normalitas residual TIDAK terpenuhi. </pre> </div> </div>		
<div> <div> 2. Homoskedastisitas (Variansi Konstan) </div> <div> Variansi residual harus konstan. Gunakan plot Residuals vs Fitted dan uji Breusch-Pagan. </div> <div>  </div> <div> <pre> === UJI HOMOSKEDASTISITAS (BREUSCH-PAGAN) === H0: Variansi residual konstan (homoskedastisitas) H1: Variansi residual tidak konstan (heteroskedastisitas) Tingkat signifikansi (α): 0.05 Hasil Pengujian: Statistik BP: 105.0945 P-value: < 2.22e-16 === INTERPRETASI === Berdasarkan uji Breusch-Pagan, model regresi mengalami masalah heteroskedastisitas karena nilai p-value (0.000000) < α (0.05). Keputusan yang diambil adalah TOLAK H0. Kesimpulan: Asumsi homoskedastisitas TIDAK terpenuhi. </pre> </div> </div>		
<div> <div> 3. Autokorelasi </div> <div> Residual tidak boleh berkorelasi satu sama lain. Gunakan uji Durbin-Watson. </div> <div>  </div> <div> <pre> === UJI AUTOKORELASI (DURBIN-WATSON) === H0: Tidak ada autokorelasi (residual independen) H1: Ada autokorelasi positif Tingkat signifikansi (α): 0.05 Hasil Pengujian: Statistik DW: 1.5339 P-value: 0 === INTERPRETASI === Berdasarkan uji Durbin-Watson, terindikasi adanya masalah autokorelasi karena nilai p-value (0.000000) < α (0.05), sehingga keputusan yang diambil adalah TOLAK H0. Kesimpulan: Asumsi independensi residual TIDAK terpenuhi. </pre> </div> </div>		
<div> <div> 4. Multikolinearitas </div> <div> Antar variabel independen tidak boleh saling berkorelasi tinggi. Gunakan Variance Inflation Factor (VIF). </div> <div>  </div> <div> <pre> === UJI MULTIKOLINEARITAS (VARIANCE INFLATION FACTOR) === Kriteria VIF: - VIF < 5: Tidak ada masalah multikolinearitas - VIF 5-10: Multikolinearitas sedang, perlu diwaspadai - VIF > 10: Multikolinearitas serius, perlu perbaikan model Hasil VIF untuk setiap variabel: FEMALE LOWEDU FAMILYSIZE ILLITERATE POVERTY 1.012211 1.058920 1.150909 1.755290 1.063378 === INTERPRETASI === Tidak ditemukan masalah multikolinearitas yang signifikan pada model (semua VIF < 5). Kesimpulan: Asumsi bebas multikolinearitas TERPENUHI. </pre> </div> </div>		
<div> <div>  Penilaian Keseluruhan Asumsi </div> <div> Catatan: Jika asumsi dilanggar, pertimbangkan transformasi data atau metode regresi alternatif. </div> </div>		

Selanjutnya, terdapat juga tab interpretasi dari persamaan yang dihasilkan :



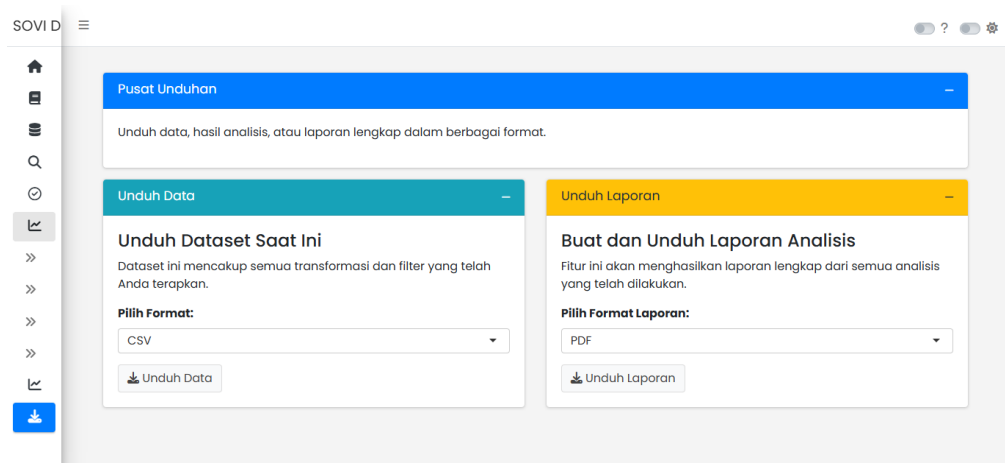
g. Download Center

Dasbor ini dilengkapi dengan fitur ekspor yang lengkap dan sistem interpretasi otomatis untuk mendukung pengambilan keputusan. Pengguna dapat dengan mudah mengeksport berbagai aset dari dasbor, termasuk:

- **Visualisasi:** Gambar beresolusi tinggi (PNG, JPEG), grafik vektor (SVG), dan widget interaktif (HTML).
- **Hasil Analisis:** Data mentah dan output statistik dapat diunduh dalam format Excel, CSV, Word, atau PDF, lengkap dengan kode R untuk memastikan transparansi dan reproduktifitas penelitian.

- **Laporan Otomatis:** Sistem dapat secara otomatis menghasilkan laporan profesional yang menggabungkan ringkasan eksekutif, analisis mendalam, dan lampiran teknis mengenai metodologi yang digunakan.

Pada tab Download center, terdapat fitur untuk mendownload dataset yang dipakai dan juga download laporan secara agregat.



Selain pada tab ini, user dapat melakukan unduh data analisis secara diasgregated pada setiap tab analisis, dan bisa melakukan unduh gambar dengan cara klik kanan pada gambar.