

**LAPORAN UJIAN AKHIR SEMESTER : PEMBUATAN
DASHBOARD RSHINY *SOCIA*alyze: *Analyzing Social
Vulnerability in Indonesia***



**Disusun untuk memenuhi Ujian Akhir Semester
Mata kuliah Komputasi Statistik
yang diampu oleh Yuliagnis Transver Wijaya, S.ST, M.Sc.**

Nuzul Athaillah
28
222313305

**KELAS 2KS3
PROGRAM STUDI D-IV KOMPUTASI STATISTIK
POLITEKNIK STATISTIKA STIS
JAKARTA
2025**

BAB I

PERENCANAAN DAN PERANCANGAN DASHBOARD

A. Latar Belakang dan Tujuan Project

Indonesia merupakan negara kepulauan yang terletak di kawasan Cincin Api Pasifik dan menjadi titik pertemuan tiga lempeng tektonik utama dunia yakni Lempeng Indo-Australia, Eurasia, dan Pasifik. Posisi geologis ini menjadikan Indonesia sangat rawan terhadap berbagai bencana alam seperti gempa bumi, tsunami, dan erupsi gunung berapi. Tingginya frekuensi kejadian bencana ini menimbulkan tantangan besar bagi upaya perlindungan dan ketahanan masyarakat, terutama di wilayah-wilayah yang memiliki kapasitas adaptasi dan pemulihan yang rendah.

Dalam konteks tersebut, konsep Kerentanan Sosial (Social Vulnerability Index/SoVI) menjadi sangat relevan. SoVI merupakan indeks yang mengukur sejauh mana kondisi sosial suatu masyarakat memengaruhi kemampuannya untuk menghadapi dan pulih dari bencana. Indeks ini disusun berdasarkan sejumlah indikator yang mencerminkan berbagai aspek kehidupan masyarakat, seperti komposisi demografi (anak-anak, lansia), kondisi sosial-ekonomi (tingkat kemiskinan, pendidikan), kondisi hunian (akses listrik, kepadatan rumah tangga), hingga kesiapsiagaan terhadap bencana (akses pelatihan atau sistem peringatan dini).

Namun, data SoVI yang tersedia bersifat kompleks dan multidimensional, mencakup puluhan indikator yang saling berinteraksi dan berbeda antar wilayah. Sebagai contoh, wilayah-wilayah seperti Maluku dan Papua menunjukkan tingkat kerentanan tinggi pada aspek kemiskinan dan buta huruf, sementara wilayah lain mungkin lebih rentan pada dimensi yang berbeda. Jumlah wilayah administratif yang dianalisis pun tidak sedikit. Lebih dari 500 kabupaten/kota di Indonesia yang semakin menambah kompleksitas analisis.

Kondisi ini menjadi tantangan tersendiri, terutama bagi pengguna non-teknis seperti pengambil kebijakan, staf pemerintah daerah, atau organisasi penanggulangan bencana. Mereka sering kali membutuhkan akses cepat terhadap informasi berbasis data untuk mendukung pengambilan keputusan, namun terbatas oleh kurangnya keahlian teknis dalam pengolahan data atau pemrograman statistik.

Untuk menjawab kebutuhan tersebut, proyek ini bertujuan mengembangkan sebuah dashboard interaktif berbasis R Shiny yang berfungsi sebagai alat bantu analisis *Business Intelligence*. Dashboard ini dirancang agar intuitif dan mudah digunakan, bahkan oleh pengguna tanpa latar belakang teknis. Melalui antarmuka yang ramah pengguna, dashboard memungkinkan eksplorasi data SoVI secara menyeluruh dari visualisasi spasial, identifikasi pola-pola kerentanan, perbandingan antar wilayah, hingga analisis hubungan antar variabel. Dengan demikian, aplikasi ini dapat menjadi alat strategis dalam memperkuat sistem manajemen risiko bencana dan perencanaan pembangunan yang tangguh terhadap bencana.

B. Penentuan Objek dan Ruang Lingkup Data

Langkah awal dalam pengembangan dashboard interaktif ini adalah merumuskan kebutuhan data yang akan menjadi fondasi utama analisis. Terdapat dua jenis data utama yang digunakan, yaitu data tabel dan data spasial, yang dirancang saling melengkapi agar menghasilkan visualisasi dan analisis yang komprehensif. Untuk data tabel, sumber utamanya adalah dataset *sovi_data.csv*, yang memuat

16 indikator kerentanan sosial (SoVI) dari 511 kabupaten/kota di Indonesia. Data ini berasal dari Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2017 yang diselenggarakan oleh Badan Pusat Statistik (BPS). Selain itu, ditambahkan pula data pendukung berupa jumlah penduduk (POPULATION) dan tingkat pertumbuhan penduduk (GROWTH), yang diperoleh dari Proyeksi Penduduk BPS tahun yang sama. Kedua variabel tambahan ini dimaksudkan untuk memperkaya konteks analisis sosial dan demografis di setiap wilayah.

Untuk melengkapi informasi spasial dalam dashboard, digunakan pula data geospasial berupa poligon batas wilayah kabupaten/kota dari Peta Geospasial Indonesia tahun 2013. Data spasial ini telah melalui proses penyederhanaan (*geometry simplification*) untuk mengurangi kompleksitas bentuk tanpa mengorbankan akurasi lokasi, dan disimpan dalam format `.rds` agar dapat dimuat secara efisien oleh aplikasi berbasis R Shiny. Seluruh proses integrasi dilakukan secara hati-hati, dengan mencocokkan antara kode wilayah dari dataset sosial dan spasial agar konsisten secara administratif. Akibat dari proses kalibrasi ini, cakupan analisis dibatasi hanya pada 511 kabupaten/kota yang memiliki kesesuaian data pada kedua sumber tersebut.

Selain kedua komponen utama tersebut, terdapat pula satu jenis data tambahan yang sangat penting, yakni file *distance.csv*. File ini memuat informasi jarak antar wilayah berdasarkan perhitungan jarak multivariat atau spasial, yang disajikan dalam bentuk matriks jarak (*distance matrix*). Data ini berperan besar dalam mendukung analisis lanjutan seperti *clustering*, *multidimensional scaling*, atau visualisasi hubungan antar wilayah berdasarkan kedekatan karakteristik sosial mereka. Dengan melibatkan data ini, dashboard tidak hanya menampilkan informasi deskriptif, tetapi juga membuka kemungkinan analisis pola dan struktur regional secara lebih mendalam.

Dengan seluruh komponen data ini baik tabular, spasial, maupun matriks jarak, dashboard yang dikembangkan dirancang untuk mendukung eksplorasi data yang menyeluruh. Semua indikator yang tercantum dalam dokumentasi SoVI dapat diakses dan dianalisis melalui fitur-fitur interaktif yang telah disiapkan, menjadikan aplikasi ini sebagai alat bantu strategis bagi pemangku kepentingan dalam memahami dan merespons risiko kerentanan sosial secara lebih tepat sasaran.

C. Perancangan Fitur dan Struktur Dashboard

Agar dashboard mudah digunakan oleh berbagai kalangan, termasuk pengguna non-teknis, perancangan alur navigasi dan tata letak antarmuka pengguna (UI) dirancang secara sistematis dan intuitif. Setiap menu mewakili tahapan logis dalam proses analisis data, mulai dari praproses hingga analisis statistik lanjutan.

Navigasi utama menggunakan sidebarMenu dari paket *shinydashboard*, yang mengorganisasi fitur-fitur utama sebagai berikut:

1. Beranda

Halaman awal yang berfungsi sebagai titik masuk ke dashboard. Di dalamnya terdapat ringkasan informasi terkait tujuan analisis, metadata dataset yang digunakan, dan petunjuk penggunaan aplikasi secara umum.

2. Manajemen Data

Menu ini berperan sebagai tahapan awal praproses data, sebelum dilakukan analisis. Terdiri dari dua tab fungsional:

- a) Kategorisasi Variabel: Mengubah variabel numerik menjadi kategorik menggunakan metode *kuantil* atau *interval sama lebar*. Hasilnya divisualisasikan dalam bentuk histogram dan diagram batang kategori.
- b) Transformasi Variabel: Menerapkan transformasi matematis seperti $\log(x+1)$, akar kuadrat (*sqrt*), atau *standardisasi (Z-score)* guna memperbaiki distribusi data atau menyamakan skala.

Fitur ini memastikan bahwa data telah memenuhi asumsi yang dibutuhkan sebelum analisis statistik dilakukan.

3. Eksplorasi Data

Menu ini menyediakan analisis eksploratif terhadap data melalui tiga pendekatan:

- a) Statistik Deskriptif: Menampilkan ringkasan numerik (mean, median, kuartil, min, max) untuk variabel yang dipilih, dilengkapi dengan interpretasi otomatis.
- b) Visualisasi Grafik: Menyediakan histogram, boxplot, dan scatter plot untuk membantu memahami sebaran dan hubungan antar variabel.
- c) Peta Interaktif: Menampilkan sebaran spasial variabel dalam bentuk peta tematik (*choropleth map*) berdasarkan wilayah administratif. Cocok untuk melihat pola geografis dalam data.

4. Uji Asumsi

Sebelum masuk ke statistik inferensial, menu ini menyediakan fitur untuk melakukan pengujian prasyarat statistik, seperti:

- a) Uji Normalitas
- b) Uji Homogenitas Varian

Uji ini penting untuk menentukan metode inferensial yang tepat.

5. Statistik Inferensia

Menu ini menyediakan alat analisis untuk mengambil kesimpulan dari data sampel ke populasi, antara lain:

- a) Uji Rata-Rata (t-test)
 - b) Uji Proporsi & Ragam
 - c) Uji ANOVA
- Setiap uji dilengkapi dengan formulir input parameter, hasil pengujian, serta interpretasi secara otomatis.

6. Regresi Linear

Tahapan ini digunakan untuk membangun model hubungan antar variabel numerik. Pengguna dapat memilih variabel independen dan dependen, melihat hasil estimasi koefisien, serta mengevaluasi kesesuaian model dengan asumsi regresi linear berganda.

BAB II

PERSIAPAN DAN PENGELOLAAN DATA UNTUK DASHBOARD

A. Sumber Data

Tahapan awal dalam pengembangan teknis dashboard adalah memastikan ketersediaan dan keterpaduan data yang menjadi dasar analisis. Proses ini mencakup pengambilan dan pemuatan dataset yang valid, sehingga aplikasi dapat berjalan dengan data yang akurat dan konsisten.

Data utama berupa dataset Kerentanan Sosial (SoVI) diperoleh secara daring dari sumber publik yang telah terverifikasi, yaitu melalui repositori GitHub. Dataset ini tersedia dalam format CSV (Comma-Separated Values) dan diakses langsung melalui tautan berikut: https://raw.githubusercontent.com/bmlmcmc/naspaclust/main/data/sovi_data.csv. Untuk memuat data tersebut ke dalam aplikasi, digunakan fungsi `readr::read_csv()` yang dikenal efisien dalam menangani file teks berukuran besar.

Sebagai tambahan, aplikasi ini juga memanfaatkan matriks penimbang jarak antar wilayah yang penting dalam analisis spasial seperti pengelompokan atau pemodelan berbasis jarak. Data ini diambil dari file `distance.csv` yang juga tersedia secara daring melalui URL: <https://raw.githubusercontent.com/bmlmcmc/naspaclust/main/data/distance.csv>. File ini dimuat dengan pendekatan serupa menggunakan `read_csv()` untuk memastikan integrasi langsung ke dalam pipeline analisis dashboard.

Selain data tabular, data spasial juga diperlukan untuk menampilkan informasi dalam bentuk peta. Data peta wilayah administratif Indonesia semula diperoleh dari Peta Geospasial Indonesia tahun 2013. Namun, untuk meningkatkan performa aplikasi dan mempercepat waktu pemuatan, file mentah seperti GeoJSON atau Shapefile tidak langsung digunakan. Sebaliknya, data spasial tersebut telah diproses terlebih dahulu melalui simplifikasi geometri dan kemudian disimpan dalam format `.rds`, yaitu format asli R yang lebih efisien.

Pemuatan file spasial ini dilakukan dengan fungsi `readRDS()`, yang secara signifikan lebih cepat dibandingkan dengan memuat ulang file geospasial mentah setiap kali aplikasi dijalankan. Dengan cara ini, proses inisialisasi aplikasi menjadi lebih ringan dan responsif.

B. Penelaahan dan Validasi Awal

Sebelum data dimanfaatkan lebih lanjut dalam proses analisis maupun visualisasi, dilakukan terlebih dahulu proses peninjauan awal untuk memastikan bahwa data yang digunakan layak, lengkap, dan konsisten. Langkah ini menjadi krusial untuk menghindari potensi kesalahan interpretasi serta menjaga keandalan hasil analisis yang ditampilkan dalam dashboard.

Proses peninjauan dimulai dari struktur dataset tabular SoVI, dengan memeriksa tipe data setiap variabel untuk memastikan kesesuaiannya. Seluruh variabel numerik dikonversi secara eksplisit agar dapat langsung digunakan dalam analisis kuantitatif. Sementara itu, kolom identifikasi wilayah seperti `DISTRICTCODE` dipastikan bertipe karakter agar mendukung proses penggabungan data dengan informasi spasial.

Langkah validasi juga mencakup pengecekan kesesuaian jumlah entitas wilayah (kabupaten/kota) antara data tabular dan data spasial. Hanya wilayah yang terdapat di kedua jenis data yang disertakan dalam proses pemetaan, agar tidak terjadi ketidaksesuaian saat visualisasi peta dilakukan. Dengan memastikan data telah melalui proses validasi awal, analisis selanjutnya dapat dilaksanakan secara lebih terstruktur, akurat, dan efisien.

C. Optimalisasi Data Spasial untuk Performa

Visualisasi spasial merupakan salah satu komponen utama dalam dashboard ini, sehingga efisiensi dalam pemrosesan dan pemuatan data geospasial menjadi sangat penting. Untuk menghindari beban komputasi yang tinggi akibat ukuran file geospasial yang kompleks, dilakukan proses penyederhanaan bentuk geometris pada batas administratif kabupaten/kota.

Data spasial yang awalnya berasal dari Peta Geospasial Indonesia tahun 2013 telah dikonversi ke dalam format GeoJSON yang sudah disederhanakan, sehingga lebih ringan dan dapat dimuat dengan lebih cepat di dalam lingkungan aplikasi R Shiny. File hasil konversi ini kemudian dimuat menggunakan fungsi `st_read()` dari paket `sf`, dan ditransformasi ke sistem koordinat WGS84 agar sesuai dengan kebutuhan pemetaan menggunakan paket `leaflet`.

Untuk menjaga konsistensi dalam pemetaan, kolom pengenalan wilayah `DISTRICTCODE` pada data spasial diselaraskan dengan kolom yang sama dari dataset tabular utama. Pendekatan ini terbukti mampu mengurangi waktu muat peta secara signifikan, sekaligus menjaga agar visualisasi tetap interaktif dan responsif, bahkan saat menampilkan seluruh wilayah Indonesia sekaligus.

Selain itu, metode ini juga memberikan keuntungan dari sisi skalabilitas aplikasi, karena memungkinkan pengembangan lebih lanjut seperti penambahan dimensi waktu atau indikator baru yang memerlukan layer spasial tambahan di masa mendatang.

BAB III

IMPLEMENTASI FITUR DAN VISUALISASI

A. Struktur Kode Aplikasi

Dashboard SOCIAalyze dibangun dengan pendekatan modular yang memisahkan antara komponen antarmuka pengguna (UI) dan logika server (server logic). Struktur ini dirancang agar mudah dibaca, dikelola, dan dikembangkan lebih lanjut. Aplikasi memanfaatkan berbagai pustaka R untuk manipulasi data, analisis statistik, visualisasi grafik dan peta, serta pembuatan laporan.

Secara umum, file `app.R` terdiri atas beberapa bagian utama:

1. Pemuatan Library

Pada bagian awal kode, dilakukan pemanggilan berbagai pustaka R penting yang digunakan untuk mendukung fungsi-fungsi utama dalam dashboard, antara lain:

- `shiny`, `shinydashboard`, dan `shinyWidgets` untuk membangun antarmuka interaktif.
- `dplyr` dan `broom` untuk manipulasi data dan peringkasan hasil analisis.
- `ggplot2`, `gridExtra` untuk visualisasi grafik.
- `DT` untuk menampilkan tabel interaktif.
- `car`, `lmtest`, dan `EnvStats` untuk uji asumsi dan analisis regresi.
- `sf`, `leaflet`, `sp`, `terra`, dan `spdep` untuk manajemen serta visualisasi data spasial.
- `rmarkdown`, `writexl`, `webshot2`, `mapview`, `htmlwidgets`, dan `zip` untuk pembuatan laporan dan fungsi unduh.

2. Definisi Variabel dan Label

Variabel-variabel yang digunakan dalam analisis, seperti `DISTRICTCODE`, `CHILDREN`, `FEMALE`, `ELDERLY`, `FHEAD`, `FAMILYSIZE`, `NOELECTRIC`, `LOWEDU`, `GROWTH`, `POVERTY`, `ILLITERATE`, `NOTRAINING`, `DPRONE`, `RENTED`, `NOSEWER`, `TAPWATER`, dan `POPULATION`, didefinisikan secara eksplisit dalam sebuah vektor tertentu (misalnya `kept_vars`). Selain itu, label yang lebih deskriptif disediakan dalam struktur `label_nama` untuk meningkatkan keterbacaan antarmuka pengguna.

3. Antarmuka Pengguna (UI)

Bagian `ui <- dashboardPage(...)` mendefinisikan tata letak tampilan utama aplikasi yang terdiri atas:

- 1) `dashboardHeader`: Menampilkan judul utama dashboard.
- 2) `dashboardSidebar`: Berisi struktur menu utama menggunakan `menuItem` dan `menuSubItem`, yang mengarahkan pengguna ke berbagai tab fitur seperti:
 - Beranda
 - Manajemen Data
 - Eksplorasi Data
 - Uji Asumsi
 - Statistik Inferensial
 - Regresi Linear
- 3) `dashboardBody`: Tempat menampilkan konten utama setiap tab melalui `tabItems`, di mana masing-masing `tabItem` menyesuaikan dengan fitur yang dipilih.

4. Logika Server

Bagian server <- function(input, output, session) {...} berfungsi sebagai inti dari aplikasi. Di dalamnya terdapat berbagai proses reaktif dan output yang menyesuaikan interaksi pengguna, meliputi:

- 1) Pemrosesan Data Reaktif: Data utama (baik tabular maupun spasial) dimuat dan dikelola secara reaktif untuk menjaga konsistensi hasil dengan input pengguna. Data spasial diolah menggunakan sf dan terra, lalu ditransformasikan untuk ditampilkan sebagai peta interaktif menggunakan leaflet.
- 2) Fungsi Per Tab: Setiap tab utama memiliki blok server masing-masing untuk menangani
 - Pembuatan elemen UI yang dinamis seperti dropdown atau checkbox.
 - Analisis statistik seperti uji normalitas (Shapiro-Wilk), homogenitas varians (Levene), uji varians (varTest), uji t, uji regresi (lm), dan lainnya.
 - Penyusunan visualisasi berupa grafik, tabel, dan peta.
 - Pembuatan interpretasi otomatis dari hasil analisis.
 - Fitur unduh laporan (Word, Excel, PNG) melalui downloadHandler, dengan bantuan rmarkdown::render dan writexl.

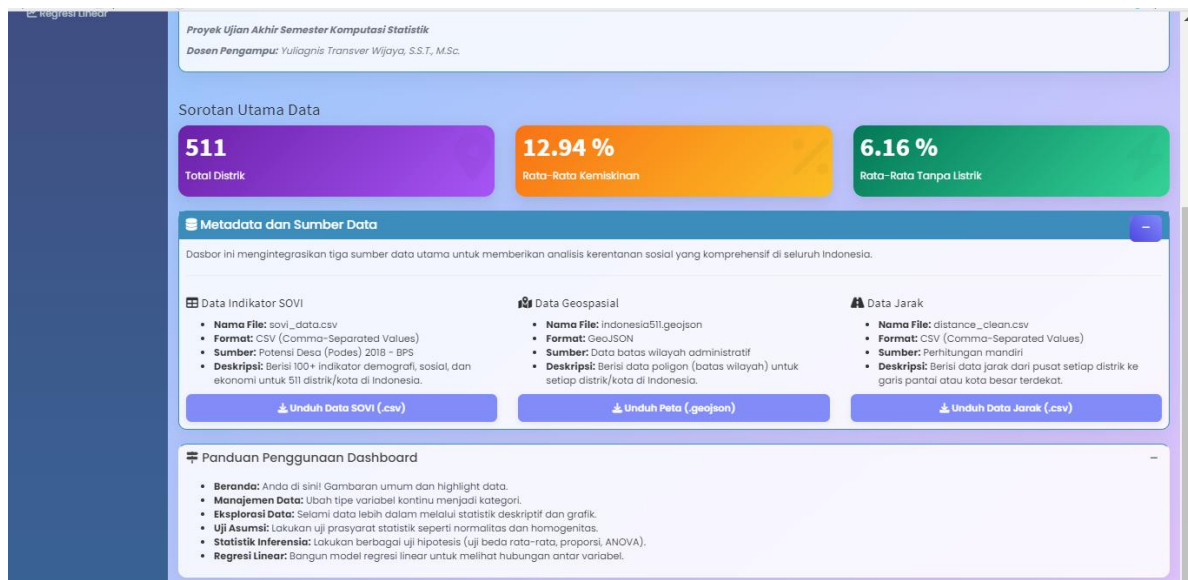
Struktur ini secara keseluruhan menjaga pemisahan yang jelas antara tampilan (UI) dan proses di balik layar (server logic), sehingga aplikasi lebih mudah untuk diperluas, di-debug, maupun dipelihara dalam pengembangan jangka panjang.

B. Pengembangan Fitur Inti

Fitur utama dalam dashboard SOCIAalyze dikembangkan untuk memberikan dukungan analisis yang menyeluruh, mencakup tahapan eksplorasi data awal hingga analisis statistik lanjutan. Setiap fitur dirancang secara interaktif dan dilengkapi dengan interpretasi otomatis guna mempermudah pemahaman pengguna terhadap hasil analisis.

1. Beranda

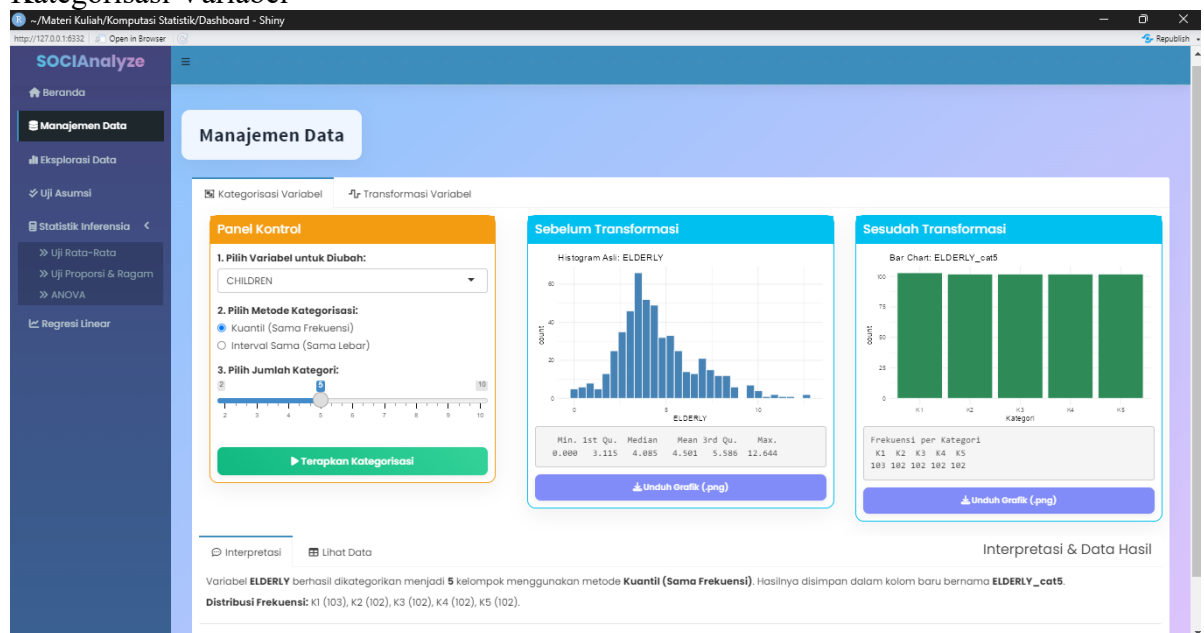




Dashboard SOCIAnalyze bagian Beranda saya berisi penjelasan singkat tentang keseluruhan dashboard, metadata yang bisa diunduh, sorotan utama data, dan panduan penggunaan dashboard.

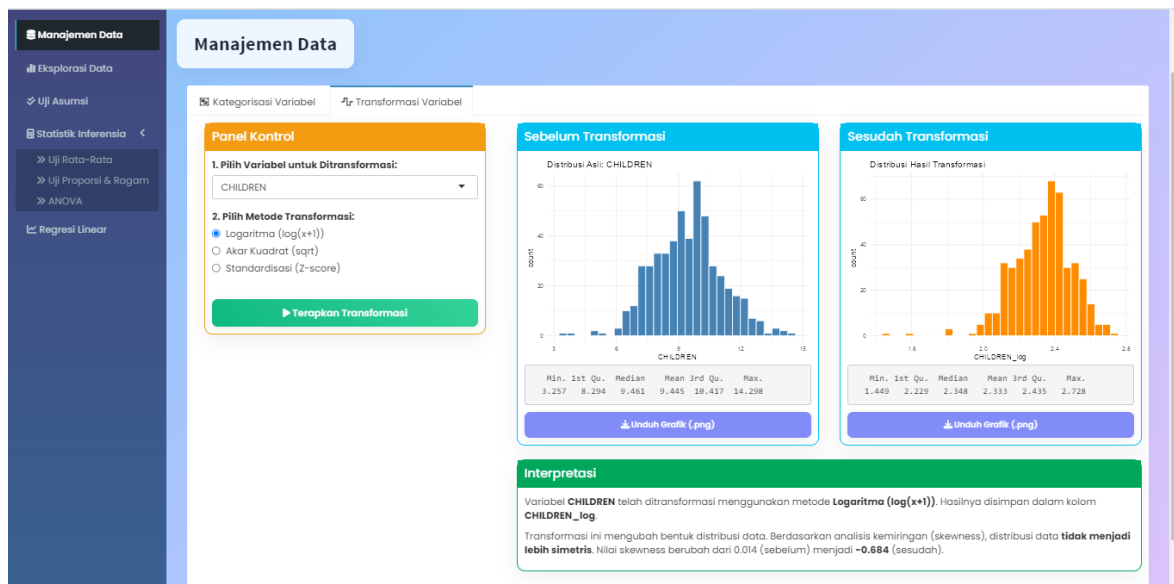
2. Manajemen Data

a) Kategorisasi Variabel



Di menu ini, bisa memilih variabel yang ingin dikategorisasi dengan metode kuartil dan interval lalu bisa memilih jumlah kategori. Setelah dipilih, bisa melihat grafik histogram variabel sebelum kategorisasi, dan bar chart sesudah kategorisasi. Variabel yang sudah dikategorisasi, akan diinterpretasikan dan tersimpan di data, bisa dilihat di Lihat Data, dan diunduh filenya.

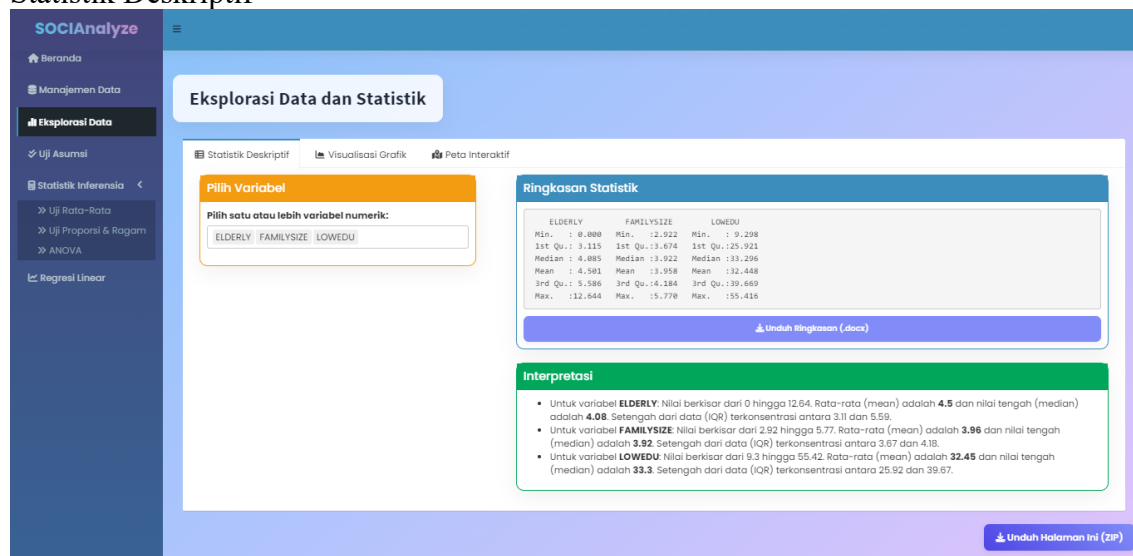
b) Transformasi Variabel



Di menu ini, bisa memilih variabel yang ingin ditransformasi dengan metodenya ada 3 yakni Logaritma, Akar Kuadrat, dan Standardisasi. Setelah dipilih, pengguna bisa melihat grafik histogram variabel sebelum transformasi, dan bar chart sesudah transformasi. Variabel yang sudah ditransformasi, akan diinterpretasikan dan tersimpan di data, bisa dilihat di Lihat Data, dan diunduh filenya.

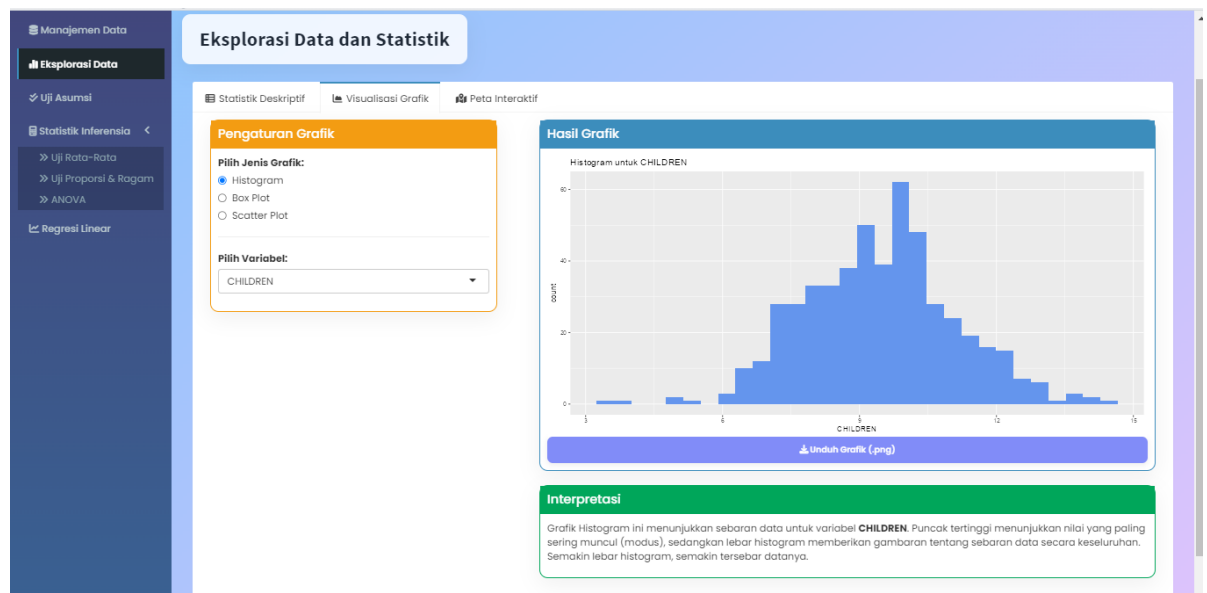
3. Eksplorasi Data

a) Statistik Deskriptif



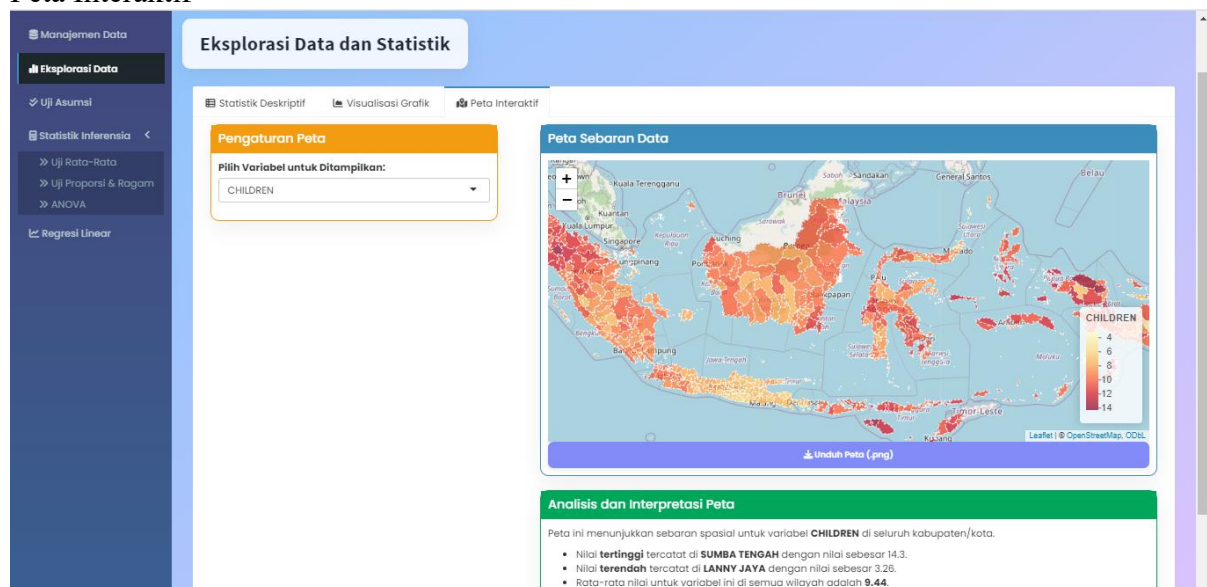
Di menu ini, pengguna bisa memilih lebih dari satu variabel numerik yang ingin dianalisis, dengan ringkasan statistik dan interpretasi yang dihasilkan setelah pengguna memilih. Hasil analisis bisa diunduh dalam bentuk word.

b) Visualisasi Grafik



Di menu ini, pengguna bisa memilih jenis grafik yang disediakan yakni histogram, box plot, dan scatter plot beserta variabel yang ingin divisualisasikan. Setelah memilih, pengguna bisa melihat grafik dan interpretasi yang dihasilkan.

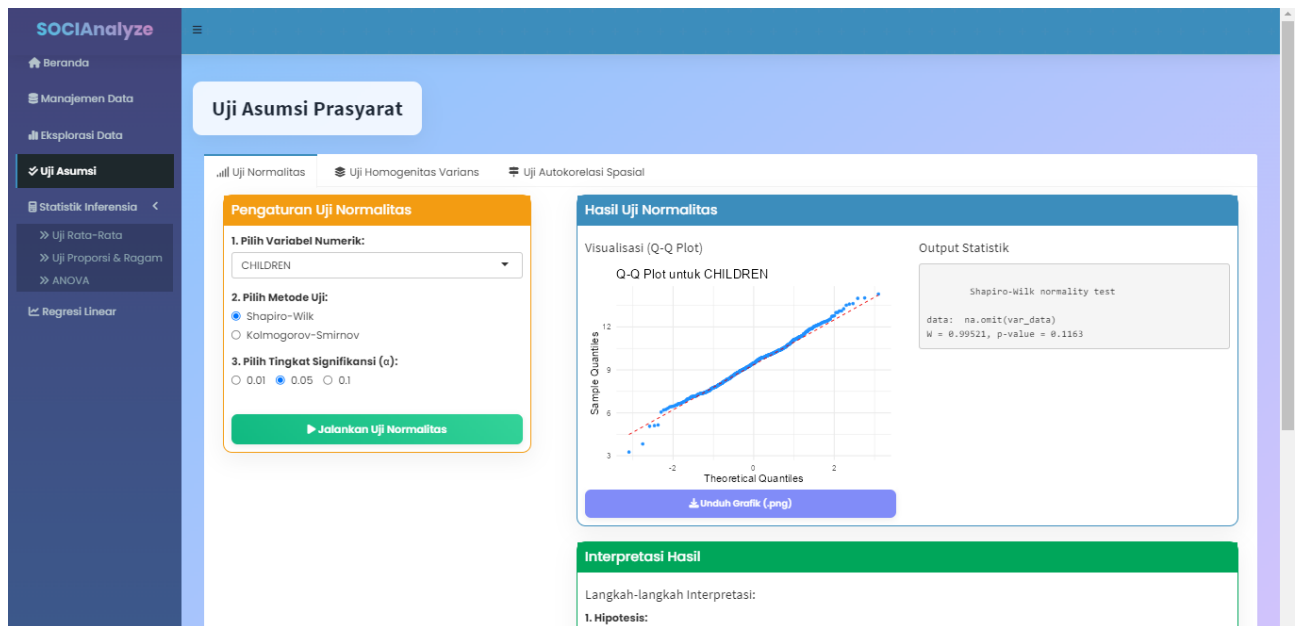
c) Peta Interaktif



Di menu ini, pengguna bisa memilih variabel yang ingin dianalisis dalam bentuk peta. Peta yang dihasilkan akan diinterpretasikan dan bisa diunduh outputnya.

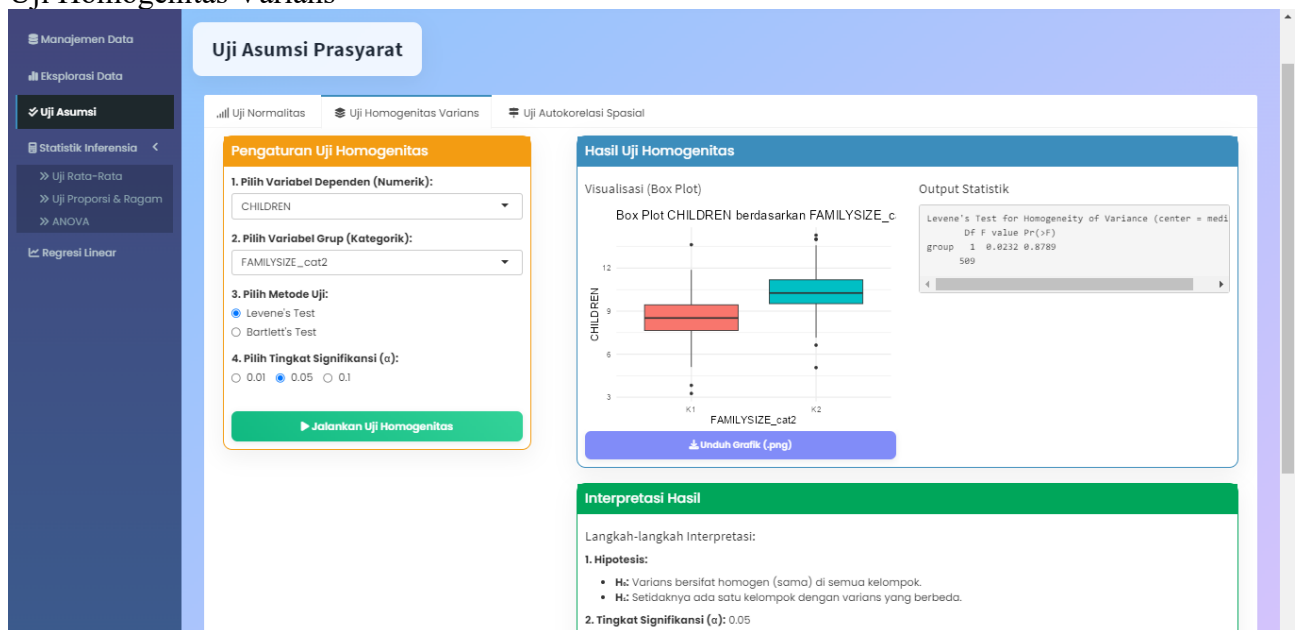
4. Uji Asumsi

a) Uji Normalitas



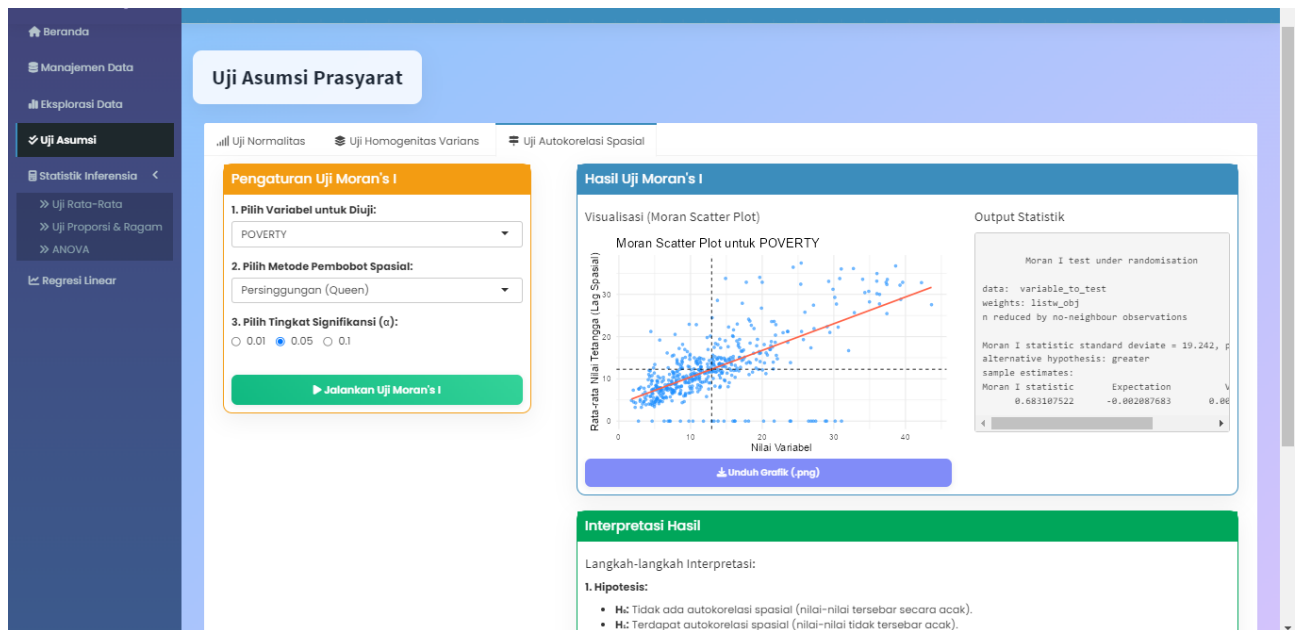
Di menu ini, pengguna bisa memilih variabel, metode uji (shapiro-wilk & kolmogorov-smirnov) serta tingkat signifikansi yang diinginkan. Setelah dipilih, output yang dihasilkan berupa Q-Q Plot, hasil uji, dan interpretasi hasil yang bisa diunduh.

b) Uji Homogenitas Varians



Di menu ini, pengguna bisa memilih variabel, metode uji (Levene's & Bartlett's Test) serta tingkat signifikansi yang diinginkan. Setelah dipilih, output yang dihasilkan berupa Box Plot, hasil uji, dan interpretasi hasil yang bisa diunduh.

c) Uji Autokorelasi Spasial

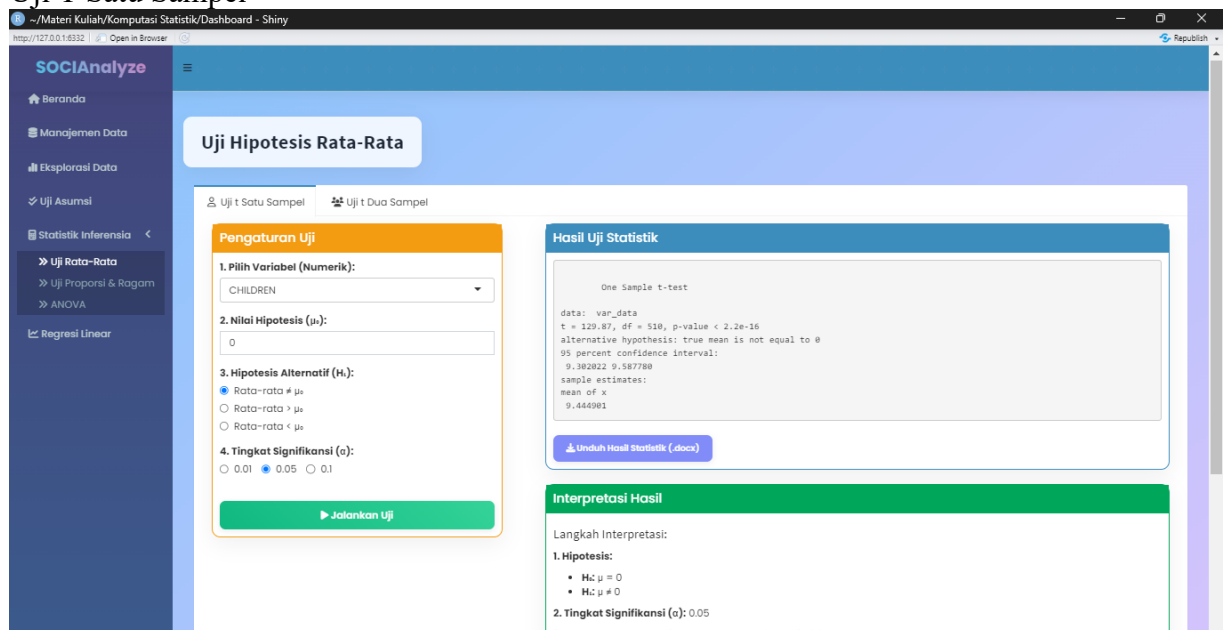


Di menu ini, pengguna bisa memilih variabel, metode uji (Queen & Jarak Inversi) serta tingkat signifikansi yang diinginkan. Setelah dipilih, output yang dihasilkan berupa Scatter Plot, hasil uji, dan interpretasi hasil yang bisa diunduh.

5. Statistik Inferensia

a) Uji Rata-Rata

a. Uji T Satu Sampel



Di menu ini, pengguna bisa memilih variabel, nilai hipotesis, hipotesis alternatif, serta tingkat signifikansi yang diinginkan. Setelah dipilih, output yang dihasilkan berupa hasil uji, dan interpretasi hasil yang bisa diunduh.

b. Uji T Dua Sampel

SOCIAnalyze

Uji Hipotesis Rata-Rata

Uji t Satu Sampel | Uji t Dua Sampel

Pengaturan Uji

1. Pilih Variabel Dependen (Numerik):
CHILDREN

2. Pilih Variabel Grup (2 Kategori):
FAMILYSIZE_cat2

3. Hipotesis Alternatif (H_1):
☒ Rata-rata Grup 1 \neq Grup 2
☐ Rata-rata Grup 1 $>$ Grup 2
☐ Rata-rata Grup 1 $<$ Grup 2

4. Tingkat Signifikansi (α):
☐ 0.01 ☒ 0.05 ☐ 0.1

[▶ Jalankan Uji](#)

Hasil Uji Statistik

Welch Two Sample t-test

```
data: num_var by cat_var
t = -14.654, df = 587.78, p-value < 2.2e-16
alternative hypothesis: true difference in means between group K1 and group K2 is not equal to 0
95 percent confidence interval:
 -2.829268 -1.549476
sample estimates:
mean in group K1 mean in group K2
 8.551966      10.341338
```

[Unduh Hasil Statistik \(.docx\)](#)

Interpretasi Hasil

Langkah Interpretasi:

1. Hipotesis:

- $H_0: \mu_1 = \mu_2$
- $H_1: \mu_1 \neq \mu_2$

2. Tingkat Signifikansi (α): 0.05

3. Aturan Keputusan: Jika p-value < α , maka H_0 ditolak.

Di menu ini, pengguna bisa memilih variabel dependen dan variabel grup khusus 2 kategori, hipotesis alternatif, serta tingkat signifikansi yang diinginkan. Setelah dipilih, output yang dihasilkan berupa hasil uji, dan interpretasi hasil yang bisa diunduh.

b) Uji Ragam & Proporsi

a. Uji Ragam (1 Populasi & 2 Populasi)

SOCIAnalyze

Uji Hipotesis Proporsi dan Ragam

Uji Proporsi | Uji Ragam

Uji Proporsi Satu Sampel

1. Pilih Variabel (Kategorik):
FAMILYSIZE_cat2

2. Pilih Level 'Sukses':
K1

3. Nilai Hipotesis (p_0):
0,5

4. Hipotesis Alternatif (H_1):
☒ Proporsi $\neq p_0$
☐ Proporsi $> p_0$
☐ Proporsi $< p_0$

5. Tingkat Signifikansi (α):
☐ 0.01 ☒ 0.05 ☐ 0.1

[▶ Jalankan Uji 1 Sampel](#)

[Unduh Hasil \(.docx\)](#)

Uji Proporsi Dua Sampel

Hasil Uji Proporsi

1-sample proportions test with continuity correction

```
data: sukses out of total_n, null probability input$prop_one_p0
X-squared = 0, df = 1, p-value = 1
alternative hypothesis: true p is not equal to 0.5
95 percent confidence interval:
 0.4568162 0.5431838
sample estimates:
p
0.5089785
```

Interpretasi Hasil

Uji Proporsi Satu Sampel

Hipotesis: $H_0: p = 0.5$ vs. $H_1: p \neq 0.5$

Aturan Keputusan: Tolak H_0 jika p-value < α (0.05).

Proporsi sampel (\hat{p}) = 0.501.

Statistik $X^2 = 0$ dengan p-value = 1.

Karena p-value $\geq \alpha$, keputusannya **GAGAL TOLAK H_0** .

Kesimpulan:

Pada $\alpha = 5\%$, tidak ada cukup bukti bahwa proporsi populasi berbeda dari 0.5.

Di menu ini, pengguna bisa memilih variabel kategorik, level 'Sukses', nilai hipotesis, hipotesis alternatif, serta tingkat signifikansi yang diinginkan. Setelah dipilih, output yang dihasilkan berupa hasil uji, dan interpretasi hasil yang bisa diunduh.

☐ Ragam > σ^2
☐ Ragam < σ^2
☒ Ragam = σ^2

4. Tingkat Signifikansi (α):
☐ 0.01 ☒ 0.05 ☐ 0.1

Jalankan Uji 1 Sampel

Unduh Hasil (.docx)

Uji Ragam Dua Sampel (F-Test)

1. Pilih Variabel Dependen (Numerik):
CHILDREN

2. Pilih Variabel Grup (2 Kategori):
FAMILYSIZE_cat2

3. Hipotesis Alternatif (H_a):
☒ Rasio Ragam $\neq 1$
☐ Rasio Ragam > 1
☐ Rasio Ragam < 1

4. Tingkat Signifikansi (α):
☐ 0.01 ☒ 0.05 ☐ 0.1

Jalankan Uji 2 Sampel

Unduh Hasil (.docx)

0.5989785

Interpretasi Hasil

Uji Proporsi Satu Sampel
Hipotesis: $H_0: p = 0.5$ vs. $H_a: p \neq 0.5$
Aturan Keputusan: Tolak H_0 jika $p\text{-value} < \alpha$ (0.05).

Proporsi sampel (\hat{p}) = **0.501**
Statistik $X^2 = 0$ dengan $p\text{-value} = 1$.
Karena $p\text{-value} \geq \alpha$, keputusannya **GAGAL TOLAK H_0** .
Kesimpulan:
Pada $\alpha = 5\%$, tidak ada cukup bukti bahwa proporsi populasi berbeda dari 0.5.

Unduh Halaman Ini (ZIP)

Di menu ini, pengguna bisa memilih variabel dependen (numerik), variabel grup 2 kategori, hipotesis alternatif, serta tingkat signifikansi yang diinginkan. Setelah dipilih, output yang dihasilkan berupa hasil uji, dan interpretasi hasil yang bisa diunduh.

c) ANOVA

a. Hasil Utama & Plot

SOCIAnalyze

Beranda

Manajemen Data

Eksplorasi Data

Uji Asumsi

Statistik Inferensia

Uji Rata-Rata

Uji Proporsi & Ragam

ANOVA

Regresi Linear

Uji Beda Rata-rata > 2 Kelompok (ANOVA)

Pengaturan ANOVA

ANOVA digunakan untuk menguji apakah terdapat perbedaan rata-rata yang signifikan antara tiga atau lebih kelompok independen.

1. Pilih Variabel Dependen (Numerik):
CHILDREN

2. Pilih Variabel Grup (Kategori > 2 Level):
LOWEDU_cat4

3. Pilih Tingkat Signifikansi (α):
☐ 0.01 ☒ 0.05 ☐ 0.1

Jalankan Analisis ANOVA

Hasil Utama & Plot

Uji Lanjutan (Post-Hoc)

Interpretasi

Tabel Ringkasan ANOVA

Visualisasi Box Plot

Box Plot untuk CHILDREN berdasarkan LOWEDU

b. Uji Lanjutan (Post Hoc)

SOCIAnalyze

Uji Beda Rata-rata > 2 Kelompok (ANOVA)

Pengaturan ANOVA

ANOVA digunakan untuk menguji apakah terdapat perbedaan rata-rata yang signifikan antara tiga atau lebih kelompok independen.

1. Pilih Variabel Dependen (Numerik):
CHILDREN

2. Pilih Variabel Grup (Kategorik > 2 Level):
LOWEDU_cat4

3. Pilih Tingkat Signifikansi (α):
☐ 0.01 ☒ 0.05 ☐ 0.1

Jalankan Analisis ANOVA

Hasil Utama & Plot | Uji Lanjutan (Post-Hoc) | Interpretasi

Tukey Honest Significant Differences (HSD)

Tabel ini menunjukkan perbandingan rata-rata antar setiap pasangan kelompok. Perhatikan kolom 'p adj' (p-value yang disesuaikan). Jika 'p adj' < α , maka terdapat perbedaan yang signifikan antara kedua kelompok tersebut.

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

Fit: aov(formula = formula, data = df_clean)

Scat_var	diff	lwr	upr	p adj
K2-K1	0.55527859	0.0372607	1.07328048	0.0301581
K3-K1	0.07572295	-0.4433856	0.59475154	0.9818592
K4-K1	-0.47166381	-0.9896737	0.04634608	0.0889977
K3-K2	-0.47954764	-0.9995762	0.03948095	0.0815406
K4-K2	-1.02693440	-1.5449443	-0.50892453	0.0009027
K4-K3	-0.54738676	-1.0664154	-0.02835817	0.0341707

Unduh Hasil Uji Lanjutan (.docx)

Unduh Halaman Ini (.zip)

c. Interpretasi

SOCIAnalyze

Uji Beda Rata-rata > 2 Kelompok (ANOVA)

Pengaturan ANOVA

ANOVA digunakan untuk menguji apakah terdapat perbedaan rata-rata yang signifikan antara tiga atau lebih kelompok independen.

1. Pilih Variabel Dependen (Numerik):
CHILDREN

2. Pilih Variabel Grup (Kategorik > 2 Level):
LOWEDU_cat4

3. Pilih Tingkat Signifikansi (α):
☐ 0.01 ☒ 0.05 ☐ 0.1

Jalankan Analisis ANOVA

Hasil Utama & Plot | Uji Lanjutan (Post-Hoc) | Interpretasi

Langkah-langkah Interpretasi ANOVA:

1. **Hipotesis:**

- H_0 (Hipotesis Nol): Rata-rata semua kelompok adalah sama ($\mu_1 = \mu_2 = \dots = \mu_k$).
- H_1 (Hipotesis Alternatif): Setidaknya ada satu rata-rata kelompok yang berbeda.

2. **Tingkat Signifikansi (α):** 0.05

3. **Aturan Keputusan:** Jika p-value < α , maka H_0 ditolak.

Hasil Analisis:

Berdasarkan analisis varians, diperoleh statistik uji F sebesar **8.751** dengan nilai p-value sebesar **1.145e-05**. Karena p-value ($1.145e-05$) < α (0.05), maka keputusannya adalah **TOLAK H_0** .

Kesimpulan:

Pada tingkat signifikansi 5%, terdapat cukup bukti untuk menyatakan bahwa **setidaknya ada satu kelompok yang memiliki rata-rata yang berbeda secara signifikan** dari kelompok lainnya.

Interpretasi Uji Lanjutan (Tukey HSD)

Karena hasil ANOVA signifikan (H_0 ditolak), kita melihat uji Tukey HSD untuk mengetahui pasangan kelompok mana yang berbeda secara signifikan. Pada tabel di tab 'Uji Lanjutan (Post-Hoc)', lihatlah pasangan dengan nilai **p adj < 0.05**. Pasangan-pasangan inilah yang rata-ratanya berbeda secara nyata.

Unduh Laporan Interpretasi (.docx)

Unduh Halaman Ini (.zip)

Di menu ini, pengguna bisa memilih variabel dependen (numerik), variabel grup kategorik > 2 level, serta tingkat signifikansi yang diinginkan. Setelah dipilih, output yang dihasilkan berupa hasil uji utama & plot, uji lanjutan (Post-Hoc), dan interpretasi hasil yang bisa diunduh.

6. Regresi Linier

a) Ringkasan Model

SOCIAnalyze

- Beranda
- Manajemen Data
- Eksplorasi Data
- Uji Asumsi
- Statistik Inferensia
 - Uji Rata-Rata
 - Uji Proporsi & Ragam
 - ANOVA
 - Regresi Linear

Analisis Regresi Linear Berganda

Pengaturan Model Regresi

Bangun model untuk memprediksi satu variabel dependen (Y) dari satu atau lebih variabel independen (X).

1. Pilih Variabel Dependen (Y):
CHILDREN

2. Pilih Variabel Independen (X):
FEMALE FHEAD FAMILYSIZE

3. Pilih Tingkat Signifikansi (α):
☐ 0.01 ☒ 0.05 ☐ 0.1

Jalankan Analisis Regresi

Ringkasan Model
Uji Asumsi Klasik
Visualisasi Residual
Interpretasi Lengkap

Output Model Regresi

Call:
lm(formula = formula, data = rv\$data)

Residuals:
Min 1Q Median 3Q Max
-5.9691 -0.7860 -0.0209 0.6853 5.8131

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 10.02319 2.78697 3.596 0.000354 ***
FEMALE -0.22679 0.05924 -3.828 0.000145 ***
FHEAD 0.04299 0.01743 2.467 0.013965 *
FAMILYSIZE 2.53518 0.13701 18.503 < 2e-16 ***

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

Residual standard error: 1.264 on 507 degrees of freedom
Multiple R-squared: 0.4124, Adjusted R-squared: 0.4089
F-statistic: 118.6 on 3 and 507 Df, p-value: < 2.2e-16

Unduh Ringkasan Model (.docx)

Unduh Halaman Ini (ZIP)

b) Uji Asumsi Klasik

SOCIAnalyze

- Beranda
- Manajemen Data
- Eksplorasi Data
- Uji Asumsi
- Statistik Inferensia
 - Uji Rata-Rata
 - Uji Proporsi & Ragam
 - ANOVA
 - Regresi Linear

Analisis Regresi Linear Berganda

Pengaturan Model Regresi

Bangun model untuk memprediksi satu variabel dependen (Y) dari satu atau lebih variabel independen (X).

1. Pilih Variabel Dependen (Y):
CHILDREN

2. Pilih Variabel Independen (X):
FEMALE FHEAD FAMILYSIZE

3. Pilih Tingkat Signifikansi (α):
☐ 0.01 ☒ 0.05 ☐ 0.1

Jalankan Analisis Regresi

Ringkasan Model
Uji Asumsi Klasik
Visualisasi Residual
Interpretasi Lengkap

1. Uji Normalitas Residual (Shapiro-Wilk)

Shapiro-Wilk normality test
data: residuals(lm_model)
W = 0.97193, p-value = 2.522e-88

2. Uji Homoskedastisitas (Breusch-Pagan)

studentized Breusch-Pagan test
data: lm_model
BP = 9.1028, df = 3, p-value = 0.02796

3. Uji Autokorelasi (Durbin-Watson)

Durbin-Watson test
data: lm_model
DW = 1.4387, p-value = 6.338e-11
alternative hypothesis: true autocorrelation is greater than 0

4. Uji Multikolinearitas (Variance Inflation Factor - VIF)

c) Visualisasi Residual

SOCIAnalyze

- Beranda
- Manajemen Data
- Eksplorasi Data
- Uji Asumsi
- Statistik Inferensia
 - Uji Rata-Rata
 - Uji Proporsi & Ragam
 - ANOVA
 - Regresi Linear

Analisis Regresi Linear Berganda

Pengaturan Model Regresi

Bangun model untuk memprediksi satu variabel dependen (Y) dari satu atau lebih variabel independen (X).

1. Pilih Variabel Dependen (Y):
CHILDREN

2. Pilih Variabel Independen (X):
FEMALE FHEAD FAMILYSIZE

3. Pilih Tingkat Signifikansi (α):
☐ 0.01 ☒ 0.05 ☐ 0.1

Jalankan Analisis Regresi

Ringkasan Model
Uji Asumsi Klasik
Visualisasi Residual
Interpretasi Lengkap

Residuals vs Fitted Plot

Normal Q-Q Plot

Unduh Visualisasi Residual (.png)

Unduh Halaman Ini (ZIP)

d) Interpretasi Lengkap

SOCIAnalyze

Beranda

Manajemen Data

Eksplorasi Data

Uji Asumsi

Statistik Inferensia

Uji Rata-Rata

Uji Proporsi & Ragam

ANOVA

Regresi Linear

Analisis Regresi Linear Berganda

Pengaturan Model Regresi

Bangun model untuk memprediksi satu variabel dependen (Y) dari satu atau lebih variabel independen (X).

1. Pilih Variabel Dependen (Y):

CHILDREN

2. Pilih Variabel Independen (X):

FEMALE

FHEAD

FAMILYSIZE

3. Pilih Tingkat Signifikansi (α):

☐ 0.01

☒ 0.05

☐ 0.1

Jalankan Analisis Regresi

Ringkasan Model

Uji Asumsi Klasik

Visualisasi Residual

Interpretasi Lengkap

1. Evaluasi Model

a. Kelayakan Model (Uji F):

Nilai p-value dari F-statistik ($\ll 2.2e-16$) lebih kecil dari α (0.05), maka model secara keseluruhan SIGNIFIKAN dalam menjelaskan variabel dependen.

b. Koefisien Determinasi (Kemampuan Menjelaskan):

Nilai Adjusted R-squared sebesar 0.4089. Ini berarti sekitar 40.89% variasi pada variabel dependen 'CHILDREN' dapat dijelaskan oleh variabel-variabel independen dalam model.

c. Pengaruh Variabel Individual (Uji t):

(Intercept):

p-value = 4e-04. Variabel ini berpengaruh signifikan terhadap 'CHILDREN'.

FEMALE: p-value = 1e-04. Variabel ini berpengaruh signifikan terhadap 'CHILDREN'.

FHEAD: p-value = 0.014. Variabel ini berpengaruh signifikan terhadap 'CHILDREN'.

FAMILYSIZE: p-value = 0. Variabel ini berpengaruh signifikan terhadap 'CHILDREN'.

2. Evaluasi Asumsi Klasik

a. Normalitas Residual:

p-value = 0. Kesimpulan: Tidak Terpenuhi (residual tidak berdistribusi normal).

b. Homoskedastisitas (Varians Residual Konstan):

p-value = 0.028. Kesimpulan: Tidak Terpenuhi (terdapat masalah heteroskedastisitas).

c. Autokorelasi (Independensi Residual):

p-value = 0. Kesimpulan: Tidak Terpenuhi (terdapat masalah autokorelasi).

d. Multikolinearitas (Hubungan Antar Variabel X):

FEMALE: VIF = 2.339 (Aman)

FHEAD: VIF = 2.345 (Aman)

FAMILYSIZE: VIF = 1.005 (Aman)

Kesimpulan Akhir

Di menu ini, pengguna bisa memilih variabel dependen (numerik), variabel independen yang bisa lebih dari satu, serta tingkat signifikansi yang diinginkan. Setelah dipilih, output yang dihasilkan berupa hasil ringkasan, uji asumsi klasik, visualisasi residual, dan interpretasi hasil yang bisa diunduh.

BAB IV

KESIMPULAN PENGEMBANGAN

A. Hasil Akhir Proyek

Proyek pengembangan dashboard R Shiny SOCIAalyze: Analisis Kerentanan Sosial di Indonesia telah berhasil diwujudkan sebagai alat analitik interaktif yang menyeluruh. Dashboard ini dirancang untuk memfasilitasi eksplorasi berbagai indikator sosial-ekonomi yang memengaruhi tingkat kerentanan sosial (SoVI) di wilayah Indonesia. Indikator yang dianalisis mencakup kemiskinan, tingkat buta huruf, akses listrik, status hunian, serta kepemimpinan keluarga perempuan.

Fitur-fitur utama dari SOCIAalyze meliputi:

- **Antarmuka Pengguna yang Interaktif:** Menggunakan desain bs4Dash yang intuitif dan ramah pengguna, dashboard memungkinkan interaksi yang lancar bagi pengguna dari berbagai latar belakang untuk menjelajahi data secara bertahap.
- **Manajemen Data yang Adaptif:** Pengguna dapat melakukan diskretisasi otomatis pada variabel numerik hingga lima kategori dan menyaring data berdasarkan provinsi untuk mendukung analisis regional.
- **Eksplorasi Data yang Kaya:** Menyediakan statistik deskriptif, visualisasi grafik interaktif seperti histogram, boxplot, dan peta choropleth yang dinamis untuk mengenali pola spasial dan wilayah rentan. Seluruh grafik disertai interpretasi otomatis guna meningkatkan pemahaman data.
- **Uji Asumsi Statistik:** Dashboard menyediakan alat untuk menguji asumsi dasar analisis statistik seperti normalitas (melalui histogram, QQ Plot, dan uji Shapiro-Wilk) serta homogenitas (menggunakan Levene's Test).
- **Statistik Inferensia yang Lengkap:** Menyediakan berbagai uji statistik, termasuk uji beda rata-rata (untuk satu dan dua kelompok), uji proporsi dan variansi, serta ANOVA satu dan dua arah, dilengkapi dengan alternatif non-parametrik seperti uji Kruskal-Wallis.
- **Pemodelan Regresi Linear:** Menawarkan fitur analisis regresi linear berganda untuk mengevaluasi pengaruh berbagai faktor terhadap tingkat kerentanan sosial, lengkap dengan pengujian asumsi klasik dan interpretasi hasil.
- **Pembuatan Laporan Otomatis:** Hasil analisis dan visualisasi dapat diunduh dalam format DOCX maupun gambar (JPG/PNG), memudahkan dokumentasi dan penyebaran hasil.

Secara keseluruhan, SOCIAalyze menjadi alat yang bermanfaat bagi peneliti, akademisi, maupun pembuat kebijakan dalam memahami dinamika kerentanan sosial antar wilayah, sehingga mendukung pengambilan keputusan berbasis data yang lebih tepat sasaran.

B. Potensi Pengembangan Lanjutan

Meskipun dashboard SOCIAalyze telah menghadirkan berbagai fitur analitik yang kuat, masih terdapat ruang pengembangan yang berpotensi meningkatkan kapabilitas, fleksibilitas, dan relevansi dashboard ini ke depannya. Beberapa arah pengembangan yang direkomendasikan antara lain:

1. **Integrasi Data Multi-Tahun dan Terbaru**
Saat ini, dashboard masih menggunakan data SoVI dari tahun 2017. Menambahkan data dari tahun-tahun yang lebih baru serta fitur perbandingan antar tahun akan memungkinkan analisis tren perubahan kerentanan sosial secara temporal yang lebih dinamis.
2. **Penambahan Indikator Sosial-Ekonomi Baru**

Memperluas cakupan indikator dengan memasukkan variabel lain yang relevan seperti akses layanan kesehatan, kualitas sanitasi, atau partisipasi angkatan kerja dapat memberikan gambaran yang lebih menyeluruh terhadap kondisi sosial suatu wilayah.

3. Analisis Spasial yang Lebih Mendalam

Selain peta choropleth, penerapan metode analisis spasial tingkat lanjut seperti *Spatial Lag Model*, *Spatial Error Model*, atau *Geographically Weighted Regression (GWR)* akan meningkatkan pemahaman terhadap variasi spasial antar wilayah. Visualisasi autokorelasi spasial seperti *Moran's I scatter plot* secara interaktif juga dapat menambah kedalaman analisis.

4. Fitur Simulasi dan Prediksi Skenario

Menyediakan fitur simulasi untuk mengevaluasi dampak perubahan suatu indikator (contoh: simulasi penurunan angka kemiskinan) terhadap indeks kerentanan dapat menjadi alat bantu strategis dalam perencanaan kebijakan berbasis data.

5. Peningkatan Penanganan Asumsi Model Regresi

Selain mengidentifikasi asumsi klasik yang tidak terpenuhi, dashboard dapat dilengkapi dengan mekanisme perbaikan otomatis seperti transformasi data (Box-Cox), penggunaan *robust standard errors*, atau saran pengelolaan outlier secara adaptif.

6. Validasi Model yang Lebih Kuat

Penambahan fitur validasi model seperti *k-fold cross validation* dan visualisasi diagnostik tambahan akan membantu meningkatkan keandalan hasil model regresi yang dibangun.

7. Peningkatan Interaktivitas Laporan

Selain menyajikan hasil dalam format dokumen atau gambar, laporan dapat dibuat lebih interaktif—misalnya, laporan Word atau HTML yang memungkinkan pengguna memperbesar grafik atau menavigasi ke bagian tertentu secara dinamis.

8. Dukungan Bahasa Ganda

Untuk memperluas jangkauan pengguna, terutama dari komunitas global, penambahan opsi bahasa (seperti Bahasa Inggris) akan menjadikan SOCIAalyze lebih inklusif dan dapat diakses oleh audiens lintas negara.

Dengan pengembangan lanjutan ini, SOCIAalyze berpotensi menjadi platform analisis kerentanan sosial yang lebih mendalam, adaptif, dan mendukung pengambilan kebijakan yang berbasis pada data spasial dan temporal yang komprehensif.

DAFTAR PUSTAKA

- Hida, A., Sukmono, A. B., & Firdaus, F. R. (2019). *Model Geographically Weighted Regression untuk Menganalisis Faktor-Faktor yang Mempengaruhi Kerentanan Sosial di Kota Semarang*. Jurnal Gaussian, 8(4), 785–794. <https://doi.org/10.14710/gaussian.v8i4.785-794>
- Jaya, H. R., & Andriyana, A. (2020). *Analisis Regresi Spasial pada Indeks Pembangunan Manusia (IPM) di Provinsi Sumatera Utara*. Jurnal Statistika, 10(1), 23–32. <https://doi.org/10.24198/jst.v10i1.23943>
- Badan Pusat Statistik. (2018). *Statistik Kesejahteraan Rakyat Indonesia 2017*. Jakarta: BPS RI. <https://www.bps.go.id/publication>