

KOMPUTASI STATISTIK

LAPORAN UJIAN AKHIR SEMESTER



Dosen Pengampu:

Yuliagnis Transver Wijaya, S.S.T., M.Sc.

Disusun Oleh:

Nama : Raihan Taufiqurrahman Zaki

NIM : 222313332

Kelas : 2KS2

PROGRAM STUDI D-IV KOMPUTASI STATISTIK

POLITEKNIK STATISTIKA STIS

2025

BAB I

PENDAHULUAN

A. Latar Belakang

Analisis terhadap kerentanan sosial memiliki peran penting dalam proses perencanaan pembangunan dan penetapan kebijakan di Indonesia. Kerentanan sosial menggambarkan kapasitas masyarakat dalam merespons berbagai tantangan dan tekanan, baik yang bersifat sosial, ekonomi, maupun lingkungan. Untuk mendukung analisis yang menyeluruh dan mudah diakses oleh berbagai pihak, dibutuhkan sebuah platform yang mampu menggabungkan data multidimensi dengan tampilan antarmuka yang ramah pengguna.

Dashboard CLEAR-IN (Comprehensive Local Exploratory Analysis and Reporting Interface for Indonesia) dikembangkan sebagai jawaban atas kebutuhan tersebut. Platform ini menyuguhkan antarmuka web interaktif yang memungkinkan pengguna melakukan eksplorasi dan analisis statistik secara mendalam terhadap data kerentanan sosial di tingkat kabupaten dan kota di seluruh wilayah Indonesia.

B. Tujuan

Pengembangan dashboard ini bertujuan untuk:

1. Menyelesaikan Ujian Akhir Semester tidak terstruktur untuk mata kuliah Komputasi Statistik
2. Menyediakan platform analisis statistik yang terintegrasi untuk data kerentanan sosial Indonesia
3. Memfasilitasi eksplorasi data interaktif melalui visualisasi geospasial dan grafik statistik
4. Menyediakan tools untuk pengujian asumsi statistik dan analisis inferensial
5. Memberikan kemampuan analisis prediktif melalui model regresi linear berganda
6. Mendukung pengambilan keputusan berbasis data bagi akademisi dan pengambil kebijakan

C. Ruang Lingkup

Dashboard ini mencakup analisis data kerentanan sosial di seluruh kabupaten/kota di Indonesia dengan fokus pada:

1. Eksplorasi dan visualisasi data multidimensional

2. Pengujian asumsi statistik
3. Analisis inferensial dan metode resampling
4. Analisis regresi untuk pemodelan prediktif
5. Visualisasi geospasial interaktif

BAB II

TINJAUAN PUSTAKA DAN METODOLOGI

A. Kerangka Teoritis

1. Social Vulnerability Index (SOVI)

Indeks Kerentanan Sosial (Social Vulnerability Index/SOVI) adalah suatu indikator gabungan yang digunakan untuk menilai tingkat kerentanan sosial di suatu wilayah berdasarkan berbagai karakteristik demografis, sosial, dan ekonomi penduduknya. SOVI berperan penting dalam mengidentifikasi wilayah-wilayah yang memiliki tingkat kerentanan tinggi terhadap bencana maupun dinamika sosial-ekonomi.

2. Multidimensional Scaling (Dibantu Geospasial)

Multidimensional Scaling (MDS) merupakan metode analisis multivariat yang bertujuan untuk memetakan data berdimensi banyak ke dalam ruang berdimensi lebih rendah. Teknik ini mempermudah visualisasi dan interpretasi pola kemiripan atau perbedaan antar objek yang diteliti.

B. Sumber Data

Dashboard ini mengintegrasikan beberapa dataset utama:

1. Data SOVI (Social Vulnerability Index)
 - a. Tingkat agregasi: Kabupaten/Kota
 - b. Cakupan: Seluruh Indonesia
 - c. Berisi indikator sosial, ekonomi, dan demografi
2. Data Hasil MDS (Multidimensional Scaling)
 - a. Dimensi representatif dari indikator kerentanan
 - b. Digunakan untuk visualisasi pola kemiripan antar wilayah
3. Data Geospasial
 - a. File: GeoJSON 38 Provinsi Indonesia - Kabupaten.json
 - b. Format: Poligon batas administratif kabupaten/kota

C. Sumber Data

1. Data Preprocessing

Proses pengolahan data meliputi:

- a. Data Loading: Menggunakan `read_csv()` untuk data numerik dan `st_read()` untuk data spasial
- b. Data Cleaning: Standardisasi nama variabel dan penanganan missing values
- c. Data Integration: Penggabungan data spasial dengan data numerik berdasarkan kode wilayah
- d. Data Normalization: Transformasi skala untuk keperluan visualisasi dan analisis

BAB III

DESAIN DAN IMPLEMENTASI SISTEM

A. Arsitektur Dashboard

Dashboard CLEAR-IN dirancang dengan arsitektur modular yang terdiri dari beberapa komponen utama:

1. Struktur Navigasi

Dashboard menggunakan struktur tab-based navigation dengan 7 modul utama:

a. Tab Beranda

- 1) Pengantar aplikasi
- 2) Instruksi penggunaan dan panduan navigasi
- 3) Metadata dan informasi sumber data

Dokumentasi:

Sumber: SUSENAS 2017 - Badan Pusat Statistik Indonesia
Publikasi: Revisiting social vulnerability analysis in Indonesia data (2021)
DOI: 523523489218181818

Metadata Variabel Dataset		
Metadata Referensi		
DISTRICTCODE	CHILDREN	FEMALE
Varibel: Kode wilayah/distrik Penjelasan: Kode wilayah/distrik	Varibel: Anak-anak Penjelasan: Persentase populasi berusia di bawah lima tahun	Varibel: Perempuan Penjelasan: Persentase populasi perempuan
ELDERLY	FHEAD	FAMILYSIZE
Varibel: Lansia	Varibel: Kepala Rumah Tangga Perempuan	Varibel: Anggota Rumah Tangga

Metadata Struktural		
DISTRICTCODE	CHILDREN	FEMALE
Varibel: Kode wilayah/distrik Penjelasan: Kode wilayah/distrik	Varibel: Anak-anak Penjelasan: Persentase populasi berusia di bawah lima tahun	Varibel: Perempuan Penjelasan: Persentase populasi perempuan
ELDERLY	FHEAD	FAMILYSIZE
Varibel: Lansia	Varibel: Kepala Rumah Tangga Perempuan	Varibel: Anggota Rumah Tangga

The screenshot shows the 'Manajemen Data' tab of a software interface. On the left, a sidebar lists navigation options: Beranda, Manajemen Data, Eksplorasi Data, Uji Asumsi Data, Resampling, Statistik Inferensi, and Regresi Linear Berganda. Below these are two buttons: 'Cetak Semua Tab' and 'Cetak Tab Aktif'. The main content area displays five data variables in boxes:

- DPRONE**: Variabel: Rawan Bencana; Penjelasan: Persentase rumah tangga yang tinggal di daerah rawan bencana.
- RENTED**: Variabel: Kepemilikan Rumah; Penjelasan: Persentase rumah tangga yang menyewa rumah.
- NOSEWER**: Variabel: Drainase; Penjelasan: Persentase rumah tangga yang tidak memiliki sistem drainase.
- TAPWATER**: Variabel: Sumber Air; Penjelasan: Persentase rumah tangga yang menggunakan air ledeng.
- POPULATION**: Variabel: Populasi; Penjelasan: Jumlah Populasi.

Below the variables is a section titled 'Fitur' containing the following items:

- Manajemen Data**: Upload file CSV/Excel, transformasi variabel, dan kategorisasi data dengan antarmuka yang intuitif.
- Uji Eksplorasi Data**: Statistik deskriptif komprehensif, visualisasi interaktif, dan pemetaan data geografis.
- Uji Asumsi Data**: Pengujian normalitas dan homogenitas dengan interpretasi otomatis hasil uji.
- Metode Resampling**: Jackknife, Bootstrap, dan Permutasi untuk estimasi parameter dan pengujian hipotesis yang robust.
- Analisis Lanjutan**: Uji statistik inferensi dan regresi linier berganda dengan diagnostik lengkap.

To the right, there is a 'Panduan Memulai' section with five numbered steps:

- Upload Dataset**: Masukkan data Anda melalui menu 'Manajemen Data' dengan format CSV dan json.
- Preprocessing Data**: Lakukan transformasi variabel atau kategorisasi data sesuai kebutuhan analisis.
- Eksplorasi Awal**: Jelajahi karakteristik data dengan statistik deskriptif dan visualisasi.
- Uji Asumsi**: Periksa normalitas dan homogenitas data sebelum melakukan analisis inferensi.
- Analisis Statistik**: Terapkan metode resampling dan uji statistik sesuai tujuan penelitian Anda.

b. Tab Manajemen Data

- Preview dataset dalam format tabel interaktif
- Ringkasan statistik deskriptif otomatis
- Fitur validasi dan seleksi data
- Transformasi data kontinyu menjadi kategorik

Dokumentasi:

The screenshot shows the 'Manajemen Data' feature with three main panels:

- Data Primer**: A panel for uploading CSV files. It shows a file named 'sov_data.csv' has been uploaded. Options include selecting the first header row and choosing column separators (Koma, Titik Koma, Tab).
- Data MDS (Optional)**: A panel for Matrix distance. It shows a message 'Belum ada file dipilih' (No file selected). Options include selecting the first header row and choosing column separators (Koma, Titik Koma, Tab).
- Data Geografis**: A panel for geographical data. It shows a file named '38 Provinsi Indonesia - Kabupaten.json' has been uploaded. It also includes a note about supported file formats: JSON, GeoJSON.

Below these panels are two more sections:

- Transformasi Variabel**: A list of variables for selection: DISTRICTCODE, CHILDREN, FEMALE, ELDERLY, FHEAD, FAMILYSIZE, NOELECTRIC, LOWEDU, GROWTH, POVERTY, ILLITERATE, NOTRADING, and nspouse.
- Kategorisasi Data**: A section for categorizing numerical variables. It shows 'CHILDREN' selected. Options include specifying the number of categories (3) and using quantiles for categorization.

The screenshot displays the Data Processing Application interface with three main tabs:

- Transformasi Variabel**: Shows a list of variables for selection, with "CHILDREN" chosen for transformation. The transformation type is set to "Logaritma Natural (ln)". The result shows the variable has been transformed into "CHILDREN_log".
- Kategorisasi Data**: Allows for categorization of numerical variables. "CHILDREN" is selected as the variable, and "3" is specified as the number of categories. The method is set to "Quantile (berdasarkan persentil)". The output shows the transformation of "FEMALE" into "FEMALE_kategori1" based on quantiles.
- Data Hasil Pemrosesan**: Displays the processed data for "CHILDREN". The table shows the original "DISTRICTCODE" values (1101, 1102, 1103, 1104, 1105) and the corresponding transformed values ("CHILDREN_log").

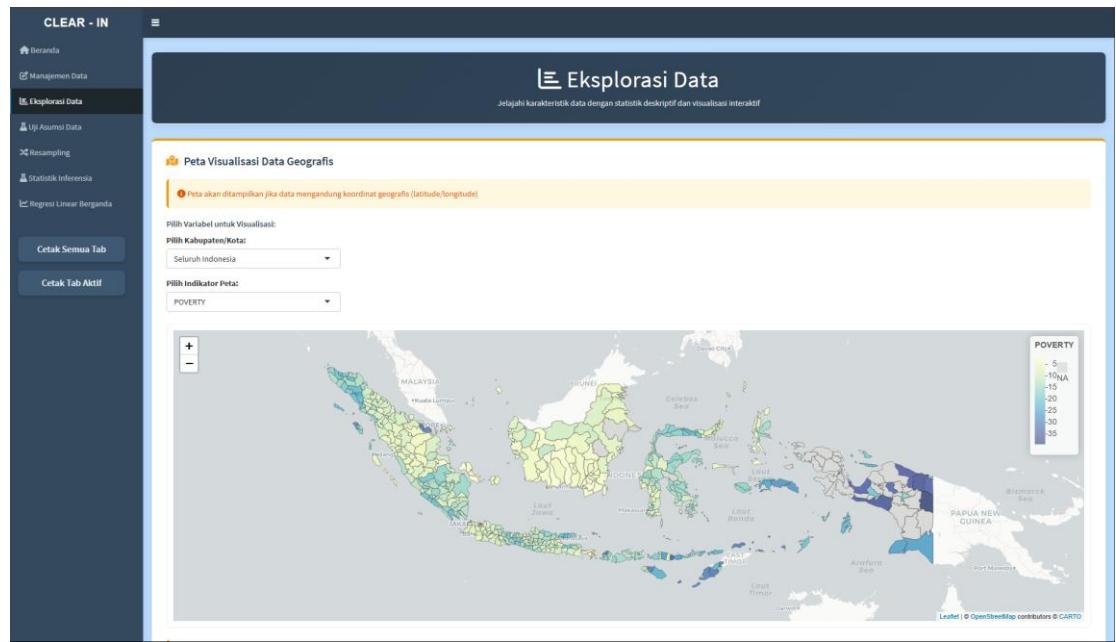
c. Tab Eksplorasi Data

1) Visualisasi geospasial menggunakan Leaflet

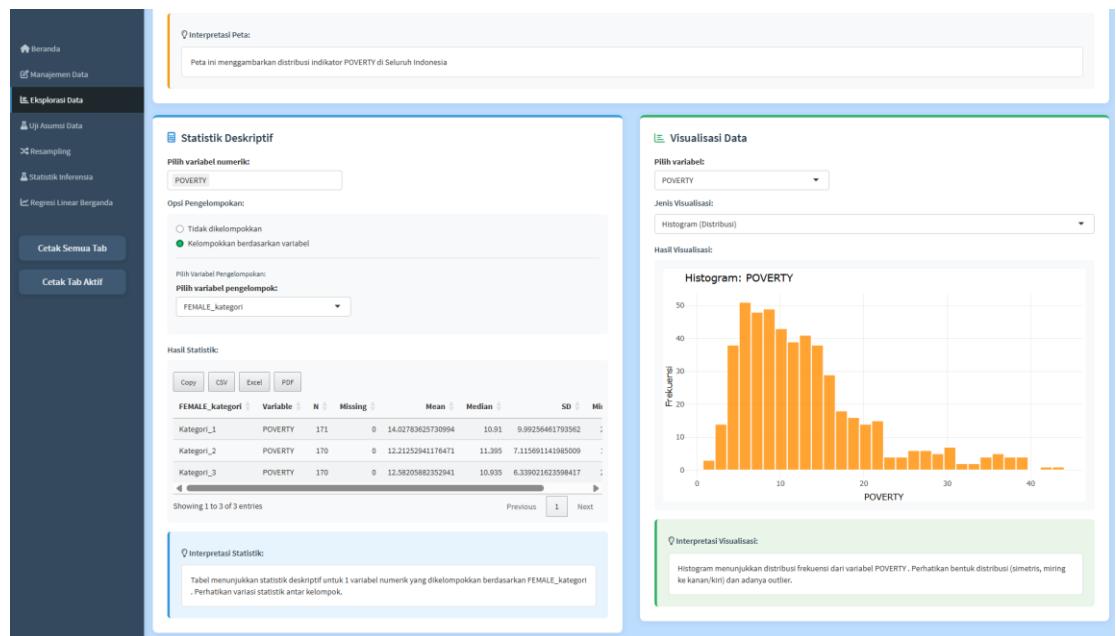
2) Scatter plot interaktif untuk analisis

Dokumentasi:

1) Peta Visualisasi



2) Statistik Deskriptif dan Visualisasi Data



d. Tab Uji Asumsi Data

- 1) Pengujian normalitas (Shapiro-Wilk, Lilliefors, KS, Jarque-Bera, Anderson-Darling)
- 2) Pengujian homogenitas (Levene, Bartlett)
- 3) Visualisasi diagnostik (histogram, Q-Q plot, boxplot)

Dokumentasi:

CLEAR - IN

Uji Asumsi Data

Uji Normalitas

Pilih variabel: POVERTY

Hipotesis Uji Normalitas:

- H₀ Data mengikuti distribusi normal
- H₁ Data tidak mengikuti distribusi normal

Kriteria Keputusan:

- Jika p-value >= 0,05 → Terima H₀ (data normal)
- Jika p-value < 0,05 → Tolak H₀ (data tidak normal)

Pilih Uji Normalitas:

- Shapiro-Wilk (selected)
- Kolmogorov-Smirnov
- Anderson-Darling
- Jarque-Bera

Uji Normalitas

Jumlah observasi: 511
Mean: 12.9429 | SD: 7.9999
Skewness: 1.3807 | Kurtosis: 4.8359

Shapiro-Wilk normality test

```
data: data
W = 0.88166, p-value < 2.2e-16
```

Uji Homogenitas

Pilih variabel numerik (dependent): POVERTY

Kelompokkan berdasarkan: FEMALE_kategori

Hipotesis Uji Homogenitas:

- H₀ Varians antar grup adalah homogen ($\sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$)
- H₁ Minimal ada satu varians grup yang berbeda

Kriteria Keputusan:

- Jika p-value >= 0,05 → Terima H₀ (varians homogen)
- Jika p-value < 0,05 → Tolak H₀ (varians tidak homogen)

Pilih Uji Homogenitas:

- Levene (selected)
- Bartlett

Uji Homogenitas

Ukuran sampel per grup:

Kategori_1	Kategori_2	Kategori_3
171	170	170

Deskripsi statistik per grup:

Grup Kategori_1	N = 171 Mean = 14.0278 SD = 9.9926 Var = 99.8513
Grup Kategori_2	N = 170 Mean = 12.2125 SD = 7.1157 Var = 50.6351
Grup Kategori_3	N = 170 Mean = 12.5821 SD = 6.339 Var = 40.1832

Ukuran keseluruhan: 63.9937

*** HASIL UJI HOMOGENITAS ***

Levene's Test for Homogeneity of Variance (center = median)

DF	F value	Pri<f>
group	2	9.1750 0.000122 ***
---	500	

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

Visualisasi:

Boxplot: POVERTY per FEMALE_kategori

Varian per Grup

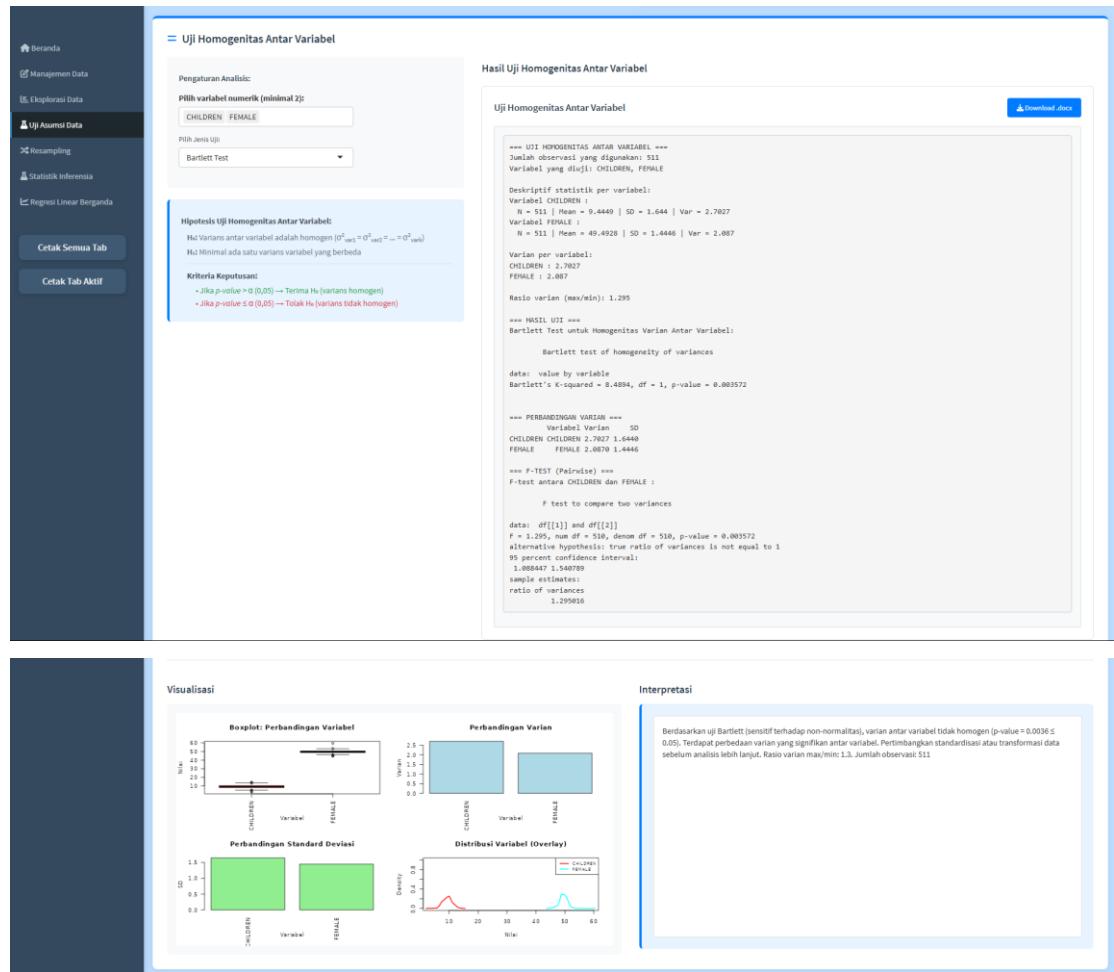
Residuals vs Fitted

Mean ± SD per Grup

Interpretasi:

Berdasarkan uji Shapiro-Wilk, data tidak terdistribusi normal ($p\text{-value} = 0.000122$). Pertimbangkan transformasi data (log, sqrt, Box-Cox) atau gunakan uji non-parametrik.

Berdasarkan uji Levene (robust terhadap non-normalitas), varian antar grup tidak homogen ($p\text{-value} = 0.000122$). Pertimbangkan transformasi data, gunakan uji Welch ANOVA, atau uji non-parametrik seperti Kruskal-Wallis. Ratio varian max/min: 2.48



e. Tab Resampling

- 1) Metode Bootstrap untuk estimasi parameter
- 2) Teknik Jackknife untuk biased correction
- 3) Metode Permutation untuk pengujian hipotesis

Dokumentasi:

CLEAR - IN

Beranda

Management Data

Eksplorasi Data

Uji Asumsi Data

Resampling

Statistik Inferensia

Regress Linear Berganda

Cetak Semua Tab

Cetak Tab Aktif

Metode Resampling

Teknik inferensi statistik non-parametrik dengan bootstrap, jackknife, dan permutasi

Jackknife Resampling

Pilih variabel: POVERTY

Pilih Statistik: Mean

Eksekusi: Jalankan Jackknife

Hasil Jackknife Resampling

```
--- JACKKNIFE RESAMPLING ---
Variabel: POVERTY
Statistik: mean
Jumlah observasi: 511
Observasi valid: 511

Estimasi original: 12.942935
Rata-rata jackknife: 12.942935
Bias: 0.000000
Estimasi terkoreksi: 12.942935
Standard error: 0.0353982
95% CI: [ 12.2477 , 13.6382 ]
Efisiensi relatif: 1.0000
```

Visualisasi

Interpretasi:

Jackknife memberikan estimasi bias sebesar 0.0000 (0.00 % dari nilai original) dan standard error 0.03539 . Estimasi terkoreksi adalah 12.9429 . Bias sangat kecil, estimator cukup unbiased.

Bootstrap Resampling

Pilih variabel: POVERTY

Pilih Statistik: Median

Replikasi: 1000

Confidence interval: 95%

Eksekusi: Jalankan Bootstrap

Hasil Bootstrap Resampling

```
--- BOOTSTRAP RESAMPLING ---
Variabel: POVERTY
Statistik: median
Jumlah replikasi: 1000
Replikasi valid: 1000
Tingkat Kepercayaan: 95 %

Estimasi original: 11.100000
Rata-rata bootstrap: 11.099160
Standard error: 0.322994
Bias: -0.000840
95% BCa CI: [ 10.2800 , 11.6300 ]
95% Percentil CI: [ 10.3436 , 11.8000 ]

--- DIAGNOSTICS ---
Skewness bootstrap dist: 0.067
Kurtosis bootstrap dist: 4.814
% extreme values: 1.4 %
```

Visualisasi

Interpretasi:

Bootstrap dengan 1000 replikasi menunjukkan distribusi sampling statistik. Standard error bootstrap adalah 0.32280 (CV = 2.95 %) dan bias estimasi adalah -0.0008 . Interval kepercayaan BCa 95 % adalah [10.2800 , 11.6300].

Permutation Test

Pilih variabel: POVERTY

Kelompok berdasarkan: FEMALE_kategori

Jumlah Permutasi: 1400

Hipotesis Alternatif: Two-sided

Eksekusi: Jalankan Permutasi

Hasil Uji Permutasi

```
--- PERMUTATION TEST ---
Variabel: POVERTY
Grup: FEMALE_kategori
Jumlah grup: 2
Jumlah observasi: 511
Jumlah permutasi: 1400

--- GROUP STATISTICS ---
Grup Kategori_1: n=171, mean=14.038, sd=0.993
Grup Kategori_2: n=170, mean=12.211, sd=7.116
Grup Kategori_3: n=170, mean=12.581, sd=6.339

--- TEST STATISTICS ---
F-statistic original: 2.4483
Max mean differences: 1.8153
Effect size ( $\eta^2$ ): 0.0096

--- P-VALUES ---
P-value (F-test): 0.0091
P-value (Max diff): 0.0071

--- CONCLUSION ---
F-test: Tidak signifikan pada  $\alpha = 0.05$ 
Max diff test: Tidak signifikan pada  $\alpha = 0.05$ 
Effect size: Kecil (< 1% variance explained)
```

Visualisasi

Interpretasi:

Permutation test menunjukkan perbedaan yang marginally significant antara grup ($p = 0.009$). Effect size kecil ($\eta^2 = 0.0096$), perbedaan praktis minimal.

f. Tab Statistik Inferensia

- 1) Uji t (satu sampel, dua sampel berpasangan/tidak berpasangan)
- 2) Uji proporsi (satu dan dua populasi)
- 3) Uji ragam (F-test)
- 4) Analysis of Variance (ANOVA) satu dan dua arah

Dokumentasi:

- 1) Uji beda rata-rata 1 populasi

The screenshot shows the 'Inferensi Statistik' application interface. On the left, a sidebar lists various statistical methods: 'CLEAR - IN', 'Beranda', 'Manajemen Data', 'Explorasi Data', 'Uji Asumsi Data', 'Resampling', 'Statistik Inferensia' (selected), and 'Regressi Linear Berganda'. Below these are 'Cetak Semua Tab' and 'Cetak Tab Aktif' buttons.

The main area has two panels: 'Pengaturan Uji Statistik' on the left and 'Hasil Uji Statistik' on the right.

Pengaturan Uji Statistik:

- Jenis Uji Statistik: Uji Rata-rata 1 Populasi
- Pilih Variabel Numerik: CHILDREN
- Nilai Hipotesis (μ_0): 0
- Hipotesis Uji Rata-rata 1 Populasi:
 - $H_0: \mu = \mu_0$ (rata-rata populasi sama dengan nilai yang dihipotesikan)
 - $H_1: \mu \neq \mu_0$ (rata-rata populasi tidak sama dengan nilai yang dihipotesikan)
- Kriteria Keputusan:
 - Jika $p\text{-value} > 0,05 \rightarrow$ Terima H_0 ($\mu = \mu_0$)
 - Jika $p\text{-value} \leq 0,05 \rightarrow$ Tolak H_0 ($\mu \neq \mu_0$)
- Uji apakah rata-rata populasi sama dengan nilai tertentu

Hasil Uji Statistik:

Hasil Uji Statistik

```
One Sample t-test
data: x
t = 129.87, df = 538, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 9.360221 9.587788
sample estimates:
mean of x
 9.444991

*** STATISTIK DESKRIPTIF ***
sample_mean: 9.444991
sample_sd: 1.644
sample_size: 533
```

Interpretasi:

Kesimpulan: H_0 DITOLAK ($p\text{-value} = 0 < \alpha = 0,05$)
Kesimpulan: Terdapat bukti statistik yang signifikan untuk mendukung hipotesis alternatif.

- 2) Uji beda rata-rata 2 populasi (independen)

The screenshot shows the 'Inferensi Statistik' application interface, similar to the previous one but for a Welch Two Sample t-test.

The sidebar and 'Pengaturan Uji Statistik' panel are identical to the previous screenshot, showing 'Uji Rata-rata 2 Populasi (Independen)' selected.

Hasil Uji Statistik:

Hasil Uji Statistik

```
Welch Two Sample t-test
data: x1 and x2
t = -9.6824, df = 953, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -4.440001 12.342992
sample estimates:
mean of x1 mean of x2
 9.444991 12.342995

*** STATISTIK DESKRIPTIF ***
sample_mean: 9.444991
group_sd: 1.644, 12.34299
group_size: 511, 511
group_names: CHILDREN, POVERTY
```

Interpretasi:

Kesimpulan: H_0 DITOLAK ($p\text{-value} = 0 < \alpha = 0,05$)
Kesimpulan: Terdapat bukti statistik yang signifikan untuk mendukung hipotesis alternatif.

3) Uji beda rata-rata 2 populasi (Berpasangan)

The screenshot shows the 'Inferensi Statistik' application interface. On the left, a sidebar lists various statistical tests: Beranda, Manajemen Data, Eksplorasi Data, Uji Asumsi Data, Resampling, Statistik Inferensia (which is selected and highlighted in dark blue), and Regresi Linear Berganda. Below these are two buttons: 'Cetak Semua Tab' and 'Cetak Tab Aktif'. The main content area has two tabs: 'Pengaturan Uji Statistik' and 'Hasil Uji Statistik'. In 'Pengaturan Uji Statistik', the 'Jenis Uji Statistik' dropdown is set to 'Uji Rata-rata 2 Populasi (Berpasangan)'. Under 'Variabel 1' and 'Variabel 2', both are set to 'CHILDREN'. The 'Hipotesis Uji Rata-rata 2 Populasi (Berpasangan)' section contains two hypotheses: $H_0: \mu_1 = 0$ atau $\mu_1 > \mu_2$ (tidak ada perbedaan rata-rata) and $H_1: \mu_1 \neq 0$ atau $\mu_1 < \mu_2$ (ada perbedaan rata-rata). The 'Kriteria Keputusan' section includes a note about p-value > 0.05 indicating acceptance of H_0 and p-value ≤ 0.05 indicating rejection of H_0 . The 'Pengaturan Uji:' section shows 'Dua sis' and 'Taraf Signifikansi (α): 0.05'. A large blue button at the bottom right says 'Lakukan Uji Statistik'. To the right, the 'Hasil Uji Statistik' tab displays the results of a Welch Two Sample t-test. The output shows:

```
Welch Two Sample t-test
data: v1 and v2
t = -0.6824, df = 553, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.6858 1.786395
sample estimates:
 mean of x mean of y
 9.44981 12.542935
```

 and a summary table:

STATISTIK DESKRIPТИF
group_mean: 9.44981, 12.5429
group_sd: 1.644, 7.9996
group_size: 511, 511
group_name: CHILDREN, POVERTY

. Below this is an 'Interpretasi' section with the text: 'Keputusan: H_0 DITOLAK ($p\text{-value} = 0 < \alpha = 0.05$)' and 'Kesimpulan: Terdapat bukti statistik yang signifikan untuk mendukung hipotesis alternatif.'

4) Uji proporsi 1 populasi

The screenshot shows the 'Inferensi Statistik' application interface. The sidebar and layout are identical to the previous screenshot, with 'Statistik Inferensia' selected. In 'Pengaturan Uji Statistik', the 'Jenis Uji Statistik' dropdown is set to 'Uji Proporsi 1 Populasi'. Under 'Jumlah Total Pengamatan (n):', the value is '511'. Under 'Jumlah Sukses/Berhasil:', the value is '275'. Under 'Proporsi Hipotesis (p₀):', the value is '0.5'. The 'Hipotesis Uji Proporsi 1 Populasi:' section contains two hypotheses: $H_0: p = p_0$ (proporsi populasi sama dengan nilai yang dihipotiskan) and $H_1: p \neq p_0$ (proporsi populasi tidak sama dengan nilai yang dihipotiskan). The 'Kriteria Keputusan' section includes a note about p-value > 0.05 indicating acceptance of H_0 and p-value ≤ 0.05 indicating rejection of H_0 . The 'Pengaturan Uji:' section shows 'Dua sis' and 'Taraf Signifikansi (α): 0.05'. A large blue button at the bottom right says 'Lakukan Uji Statistik'. To the right, the 'Hasil Uji Statistik' tab displays the results of a 1-sample proportions test without continuity correction. The output shows:

```
1-sample proportions test without continuity correction
data: n_success out of n_total, null probability input$p0
X-squared = 2.9795, df = 1, p-value = 0.08848
alternative hypothesis: true p is not equal to 0.5
95 percent confidence interval:
 0.494810 0.505188
sample estimates:
        p 
0.5301605
```

 and a summary table:

STATISTIK DESKRIPТИF
sample_prop: 0.5302
n_success: 275
n_total: 511
n_failure: 236

. Below this is an 'Interpretasi' section with the text: 'Keputusan: H_0 GAGAL DITOLAK ($p\text{-value} = 0.088481 \geq \alpha = 0.05$)' and 'Kesimpulan: Tidak terdapat bukti statistik yang cukup untuk mendukung hipotesis alternatif.'

5) Uji proporsi 2 populasi

The screenshot shows the software's main menu on the left and two main panels on the right.

Pengaturan Uji Statistik:

- Jenis Uji Statistik: Uji Proporsi 2 Populasi
- Kelompok 1: Jumlah Total Pengamatan Kelompok 1: 100, Jumlah Sukses Kelompok 1: 45
- Kelompok 2: Jumlah Total Pengamatan Kelompok 2: 100, Jumlah Sukses Kelompok 2: 55
- Hipotesis Uji Proporsi 2 Populasi:
 - $H_0: p_1 = p_2$ (proporsi kedua populasi sama)
 - $H_1: p_1 \neq p_2$ (proporsi kedua populasi berbeda)
- Kriteria Kepatuhan:
 - Jika $p\text{-value} > \alpha (0,05)$ → Terima $H_0 (p_1 = p_2)$
 - Jika $p\text{-value} \leq \alpha (0,05)$ → Tolak $H_0 (p_1 \neq p_2)$
- Pengaturan Uji:
 - Hipotesis Alternatif: Dua sisi
 - Taraf Signifikansi (α): 0,05

Hasil Uji Statistik:

```

2-sample test for equality of proportions without continuity correction
data: c(x1, x2) out of c(n1, n2)
X-squared = 2, df = 1, p-value = 0.1573
alternative hypothesis: two.sided
95 percent confidence interval:
-0.2779959 0.4559959
sample estimates:
prop 1 prop 2
0.45 0.55
  
```

Interpretasi:

Kesimpulan: H_0 GAGAL DITOLAK ($p\text{-value} = 0.157299 > \alpha = 0.05$)
Kesimpulan: Tidak terdapat bukti statistik yang cukup untuk mendukung hipotesis alternatif.

6) Uji ragam 1 populasi

The screenshot shows the software's main menu on the left and two main panels on the right.

Pengaturan Uji Statistik:

- Jenis Uji Statistik: Uji Ragam 1 Populasi
- Pilih Variabel Numerik: LOWDU
- Variansi Hipotesis (σ^2_h): 1
- Hipotesis Uji Ragam 1 Populasi:
 - $H_0: \sigma^2 = \sigma^2_h$ (ragam populasi sama dengan nilai yang dihipotesiskan)
 - $H_1: \sigma^2 \neq \sigma^2_h$ (ragam populasi tidak sama dengan nilai yang dihipotesiskan)
- Kriteria Kepatuhan:
 - Jika $p\text{-value} > \alpha (0,05)$ → Terima $H_0 (\text{ragam} = \sigma^2_h)$
 - Jika $p\text{-value} \leq \alpha (0,05)$ → Tolak $H_0 (\text{ragam} \neq \sigma^2_h)$
- Pengaturan Uji:
 - Hipotesis Alternatif: Dua sisi
 - Taraf Signifikansi (α): 0,05

Hasil Uji Statistik:

```

$statistic
X-squared
4337.19
$p.value
[1] 0
$parameter
df
510
$method
[1] "chi-square test for Variance"
$std.error
[1] "sqrt(x)"
$alternative
[1] "two.sided"
  
```

```

  
```

Interpretasi:

Kesimpulan: H_0 DITOLAK ($p\text{-value} = 0 < \alpha = 0.05$)
Kesimpulan: Terdapat bukti statistik yang signifikan bahwa varian populasi tidak sama dengan nilai yang dihipotesiskan.

7) Uji ragam 2 populasi

8) Uji ANOVA 1 populasi

9) Uji ANOVA 2 populasi

g. Tab Regresi Linear Berganda

- 1) Pemodelan prediktif dengan Regresi Linear Berganda
 - 2) Diagnostik model
 - 3) Interpretasi hasil

Dokumentasi:

1) Pemilihan variabel

The screenshot shows the 'CLEAR - IN' software interface. On the left, a sidebar menu includes 'Beranda', 'Management Data', 'Explorasi Data', 'Uji Asumsi Data', 'Resampling', 'Statistik Inferensia', and 'Regresi Linear Berganda'. Below these are two buttons: 'Cetak Semua Tab' and 'Cetak Tab Aktif'. The main panel title is 'Regresi Linear Berganda' with the subtitle 'Analisis hubungan linear antara variabel dependen dan independen dengan uji asumsi lingkup'. The 'Pengaturan Regresi Linear Berganda' section has a dropdown menu 'Variabel Dependens (Y):' set to 'POVERTY'. To the right is a list of 'Variabel Independens (X):' with several items checked: CHILDREN, FEMALE, ELDERLY, FHEAD, FAMILYSIZE, NOELECTRIC, LOWEDU, GROWTH, POVERTY, LLITERATE, NOTRANNING, OPIONE, RENTED, NOSEWER, TAPWATER, POPULATION, and CHILDREN_llg. A green button at the bottom right says 'Jalankan Regresi'.

2) Ringkasan Model

The screenshot shows the 'CLEAR - IN' software interface. The sidebar menu is identical to the previous screenshot. The main panel title is 'Hasil Analisis' with tabs for 'Ringkasan Model', 'Uji Asumsi', 'Plot Diagnostik', and 'Prediksi'. The 'Ringkasan Model MLR' tab is active. It displays the R code for the regression model, the output results, and a summary table. The R code includes the formula 'POVERTY ~ LOWEDU + LLITERATE', the call to the formula function, residual statistics, coefficients, and other model details. The output results show the model summary, including R-squared (0.3939), adjusted R-squared (0.3895), and F-statistic (163.7). The summary table includes columns for Estimate, Std. Error, t value, and Pr(>|t|). The interpretation section notes that the model explains 39.19% of the variance in the dependent variable and that the intercept is significant (p < 0.001).

3) Uji Asumsi

Hasil Analisis

1. Uji Normalitas Residual

```
UJI NORMALITAS RESIDUAL

Shapiro-Wilk normality test
data: residuals
W = 0.92089, p-value = 2.192e-15

INTERPRETASI (shapiro-wilk):
P-value (0) ≤ 0.05
KESIMPULAN: Residual TIDAK berdistribusi normal (asumsi normalitas TIDAK TERPENUHI)
Saran: Pertimbangkan transformasi data atau gunakan metode non-parametrik

Asympotic one-sample Kolmogorov-Smirnov test

data: residuals
D = 0.10655, p-value = 2.18de-05
alternative hypothesis: two-sided

INTERPRETASI (kolmogorov-smirnov):
P-value (0) ≤ 0.05
KESIMPULAN: Residual TIDAK berdistribusi normal (asumsi normalitas TIDAK TERPENUHI)
Saran: Pertimbangkan transformasi data atau gunakan metode non-parametrik
```

2. Uji Multikolinearitas

```
UJI MULTIKOLINEARITAS (VIF)

Nilai VIF untuk setiap variabel:
LOMBA_ILLITERATE
1.0338 1.0338

INTERPRETASI:
VIF > 5: Tidak ada multikolinearitas
5 < VIF < 10: Multikolinearitas sedang
VIF > 10: Multikolinearitas tinggi (bermasalah)

- LOBANG : VIF = 1.0338 - BAIK (tidak ada multikolinearitas)
- ILLITERATE : VIF = 1.0338 - BAIK (tidak ada multikolinearitas)
```

3. Uji Homoskedastisitas

```
UJI HOMOSKEDASTISITAS (BREUSCH-PAGAN)

studentized Breusch-Pagan test

data: mlr_model()
BP = 5.5347, df = 3, p-value = 0.06222

INTERPRETASI:
P-value (0.0622) > 0.05
KESIMPULAN: Varians residual homogen (asumsi homoskedastisitas TERPENUHI)
```

4. Uji Autokorelasi

```
UJI AUTOKORELASI (DURBIN-WATSON)

Durbin-Watson test

data: mlr_model()
DW = 0.83879, p-value < 2.2e-16
alternative hypothesis: true autocorrelation is greater than 8

INTERPRETASI:
Nilai Durbin-Watson: 0.8168
P-value: 0

INTERPRETASI NILAI DURBIN-WATSON:
X Terdapat autokorelasi positif (asumsi independensi TIDAK TERPENUHI)

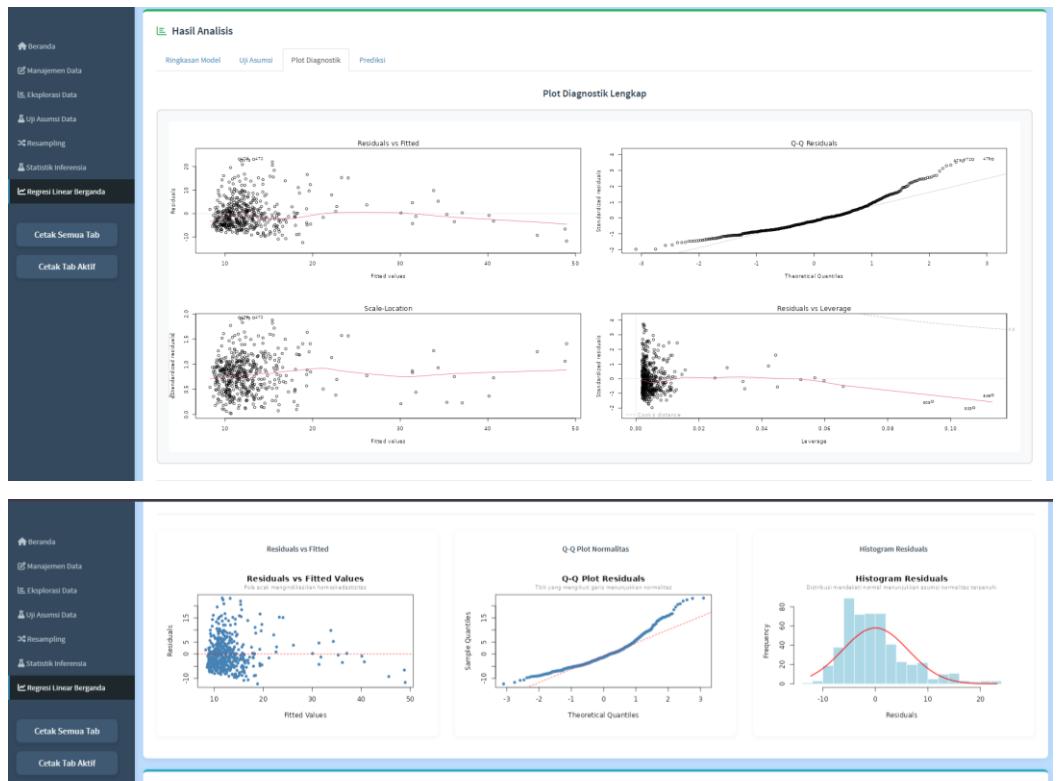
INTERPRETASI P-VALUE:
X P-value (0) ≤ 0.05: Ada bukti autokorelasi yang signifikan

KESIMPULAN KESILIRUMAN:
A: ASUMSI INDEPENDENSI TIDAK TERPENUHI
(A: bukti autokorelasi yang signifikan secara statistik)

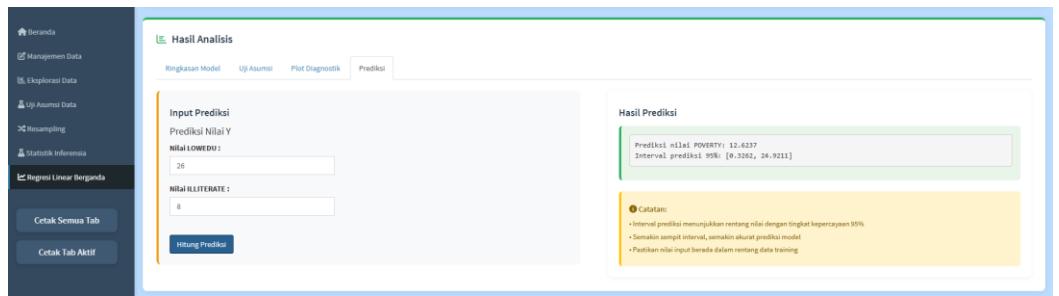
SARAN PENGAMANAN:
Autokorelasi positif terdeteksi:
- Periksa urutan data (pastikan data time series diurutkan dengan benar)
- Pertimbangkan penggunaan lag variabel
- Pertimbangkan model ARIMA atau teknik lainnya
- Transformasi data atau penambahan variabel kontrol
- Pertimbangan menggunakan Generalized Least Squares (GLS)

CATATAN:
• Nilai Durbin-Watson berkisar 0-4
• Nilai <2 menunjukkan tidak ada autokorelasi
• Nilai >2 menunjukkan autokorelasi positif
• Nilai <2 menunjukkan autokorelasi negatif
```

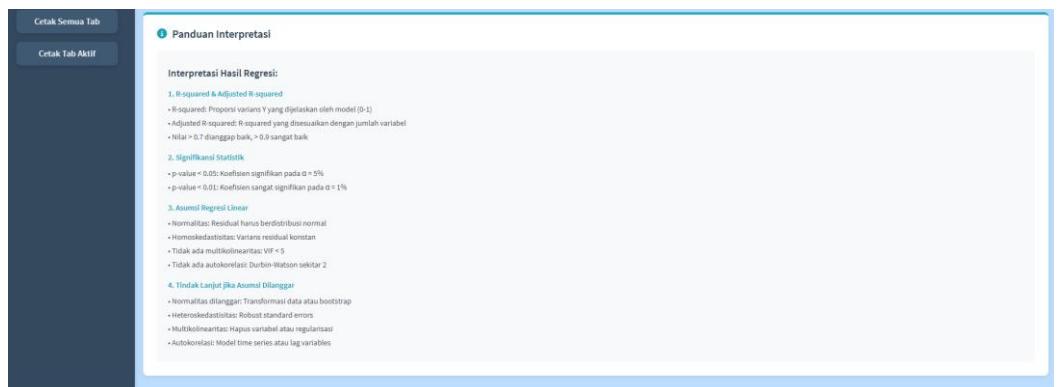
4) Plot Diagnostik



5) Prediksi



6) Panduan



B. Arsitektur Dashboard

1. Visualisasi Interaktif

- a. Peta : Menampilkan distribusi spasial indeks kerentanan sosial
- b. Scatter Plot Dinamis : Representasi hasil MDS dalam ruang 2D

2. Analisis Statistik Terintegrasi

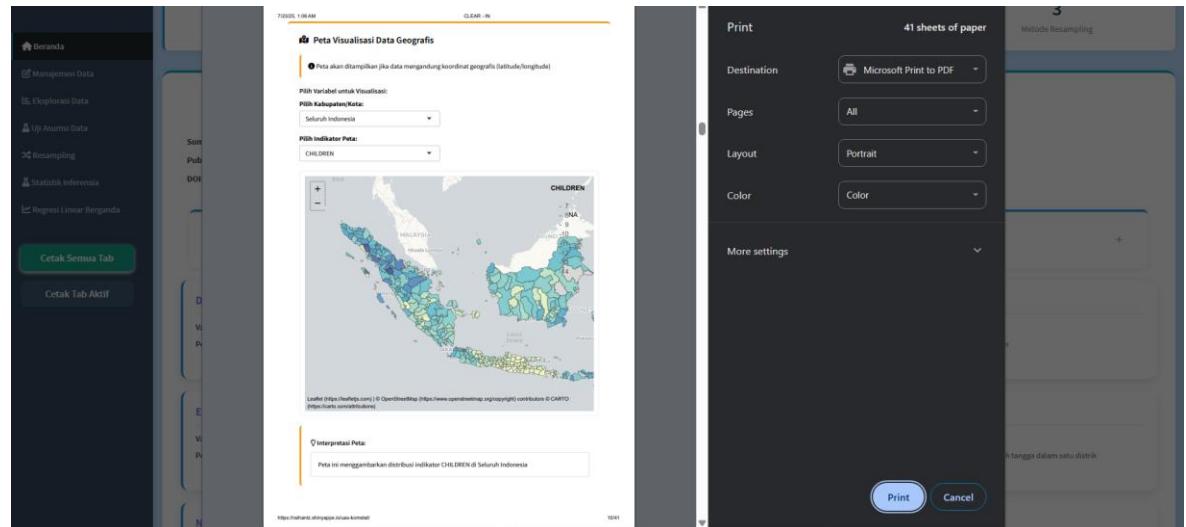
- a. Pengujian Asumsi : Automated testing dengan interpretasi hasil
- b. Metode Resampling : Bootstrap, Jackknife, dan Permutation testing
- c. Regression Analysis : Pemeriksaan penuh untuk asumsi

3. Sistem Reporting

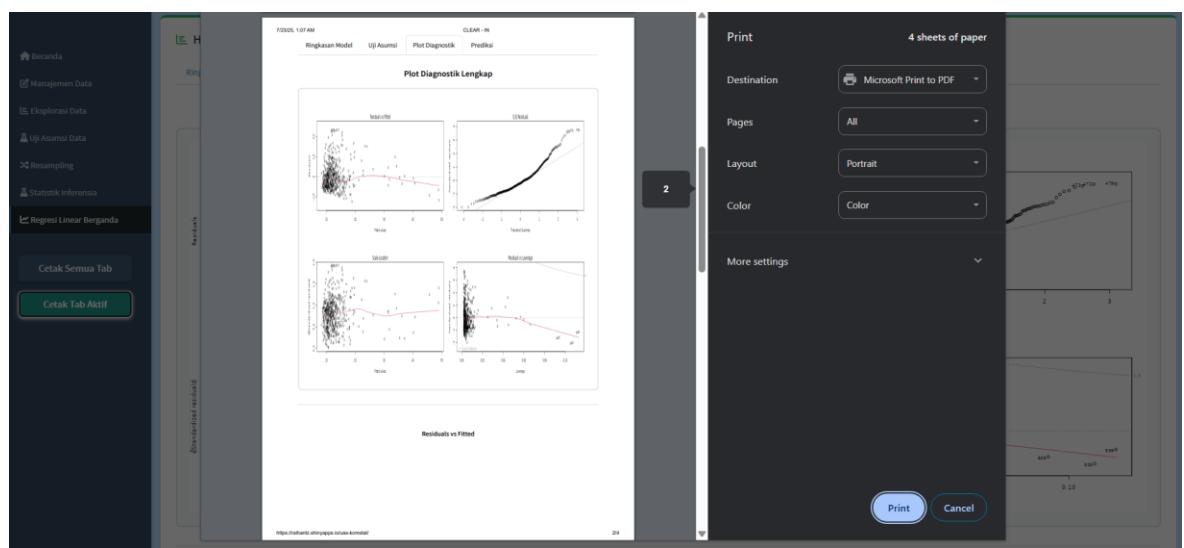
- a. Export Functionality: Tombol "Cetak Semua Tab" dan "Cetak Tab Aktif"
- b. PDF Generation: Automatic report generation dalam format PDF
- c. Reproducible Results: Konsistensi hasil untuk replikasi analisis

Dokumentasi:

- a. Cetak Semua Tab



b. Cetak Tab Aktif



BAB IV

STUDI KASUS

A. Pendahuluan

1. Latar Belakang

Kemiskinan merupakan indikator kunci dalam menilai tingkat kesejahteraan masyarakat di suatu daerah. Di Indonesia, permasalahan kemiskinan tidak hanya berkaitan dengan aspek ekonomi, tetapi juga dipengaruhi oleh berbagai faktor sosial, terutama tingkat pendidikan dan kemampuan literasi masyarakat. Tersedianya data sosial dan demografis yang lengkap dari berbagai wilayah membuka peluang untuk melakukan analisis kuantitatif untuk mengidentifikasi faktor-faktor yang berperan dalam menentukan tingkat kemiskinan.

Regresi linier berganda menjadi salah satu metode statistik yang sesuai untuk menganalisis hubungan antara variabel-variabel tersebut. Dalam penelitian ini, digunakan pendekatan ekonometrika untuk mengkaji sejauh mana pengaruh tingkat pendidikan rendah (LOWEDU) dan tingkat buta huruf (ILLITERATE) terhadap tingkat kemiskinan (POVERTY) di suatu wilayah.

2. Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah:

- a. Apakah tingkat pendidikan rendah (LOWEDU) berpengaruh signifikan terhadap tingkat kemiskinan?
- b. Apakah tingkat buta huruf (ILLITERATE) berpengaruh signifikan terhadap tingkat kemiskinan?
- c. Seberapa besar kontribusi kedua variabel tersebut dalam menjelaskan variasi tingkat kemiskinan?

3. Tujuan Penelitian

Tujuan penelitian ini adalah:

- a. Menganalisis hubungan antara tingkat pendidikan rendah dan kemiskinan.
- b. Menganalisis hubungan antara tingkat buta huruf dan kemiskinan.
- c. Menyusun model regresi linier untuk memprediksi tingkat kemiskinan berdasarkan kedua variabel tersebut.

B. Metodologi Penelitian

1. Data dan Variabel

Data yang digunakan dalam penelitian ini mencakup beberapa variabel sosial dalam bentuk persentase (%), dengan fokus pada tiga variabel utama:

- a. POVERTY : Persentase tingkat kemiskinan (variabel dependen)
- b. LOWEDU : Persentase tingkat pendidikan rendah (variabel independen)
- c. ILLITERATE : Persentase tingkat buta huruf (variabel independen)

2. Metode Analisis

Metode yang digunakan adalah regresi linier berganda dengan model sebagai berikut:

$$POVERTY_i = \beta_0 + \beta_1 \cdot LOWEDU_i + \beta_2 \cdot ILLITERATE_i + \varepsilon_i$$

Dimana:

$POVERTY_i$ = tingkat kemiskinan daerah ke-i

$LOWEDU_i$ = tingkat pendidikan rendah daerah ke-i

$ILLITERATE_i$ = tingkat buta huruf daerah ke-i

β_0 = konstanta

β_1, β_2 = koefisien regresi

ε_i = error term

Asumsi klasik regresi linier yang diuji meliputi normalitas residual, homoskedastisitas, non-autokorelasi, dan non-multikolinearitas.

C. Hasil dan Pembahasan (Tingkat Signifikansi 95%)

Uji dan hasil analisis dilakukan dengan memanfaatkan dashboard CLEAR-IN yang sudah dikembangkan sebelumnya.

1. Pengujian Asumsi Klasik - Tahap 1 (Variabel Awal)

a) Normalitas

Hipotesis:

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

```
1) === UJI NORMALITAS RESIDUAL ===
2)
3)      Shapiro-Wilk normality test
4)
5) data: residuals
6) W = 0.92409, p-value = 2.192e-15
7)
8) INTERPRETASI (shapiro-wilk):
9) P-value ( 0 ) ≤ 0.05
10) KESIMPULAN: Residual TIDAK berdistribusi normal (asumsi normalitas TIDAK TERPENUHI)
11) SARAN: Pertimbangkan transformasi data atau gunakan metode non-parametrik
12)
13)      Asymptotic one-sample Kolmogorov-Smirnov test
14)
15) data: residuals
16) D = 0.10455, p-value = 2.816e-05
17) alternative hypothesis: two-sided
18)
19) INTERPRETASI (kolmogorov-smirnov):
20) P-value ( 0 ) ≤ 0.05
21) KESIMPULAN: Residual TIDAK berdistribusi normal (asumsi normalitas TIDAK TERPENUHI)
22) SARAN: Pertimbangkan transformasi data atau gunakan metode non-parametrik
```

Berdasarkan hasil uji normalitas, hipotesis nol ditolak (H_0 ditolak), yang menunjukkan bahwa asumsi normalitas residual tidak terpenuhi. Pelanggaran asumsi ini dapat mempengaruhi validitas inferensi statistik, sehingga diperlukan transformasi data untuk memperbaiki distribusi residual.

b) Homoskedastisitas

Hipotesis:

H_0 : Varians residual konstan (homoskedastisitas)

H_1 : Varians residual tidak konstan (heteroskedastisitas)

- 1) === UJI MULTIKOLINEARITAS (VIF - Variance Inflation Factor) ==
=
- 2)
- 3) Nilai VIF untuk setiap variabel:
- 4) LOWEDU ILLITERATE
- 5) 1.0338 1.0338
- 6)
- 7) INTERPRETASI:
- 8) $VIF < 5$: Tidak ada multikolinearitas
- 9) $5 \leq VIF < 10$: Multikolinearitas sedang
- 10) $VIF \geq 10$: Multikolinearitas tinggi (bermasalah)
- 11)
- 12) - LOWEDU : $VIF = 1.0338$ - BAIK (**tidak ada multikolinearitas**)
- 13) - ILLITERATE : $VIF = 1.0338$ - BAIK (**tidak ada multikolinearitas**)

Hasil uji menunjukkan bahwa hipotesis nol tidak dapat ditolak (H_0 diterima), yang berarti varians residual bersifat konstan. Terpenuhinya asumsi ini mengindikasikan bahwa model tidak mengalami masalah heteroskedastisitas, sehingga estimator yang dihasilkan tetap efisien.

c) Nonautokorelasi

Hipotesis:

H_0 : Tidak ada autokorelasi antar residual

H_1 : Terdapat autokorelasi antar residual

1) === UJI AUTOKORELASI (DURBIN-WATSON) ===

2) Durbin-Watson test

3)

4) data: mlr_model()

5) DW = 0.81679, p-value < 2.2e-16

6) alternative hypothesis: true autocorrelation is greater than 0

7)

8) === INTERPRETASI ===

9) Nilai Durbin-Watson: 0.8168

10) P-value: 0

11)

12) INTERPRETASI NILAI DURBIN-WATSON:

13) X Terdapat autokorelasi positif (asumsi independensi TIDAK TERPENUHI)

14)

15) INTERPRETASI P-VALUE:

16) X P-value (0) ≤ 0.05: Ada bukti autokorelasi yang signifikan

17)

18) === KESIMPULAN KESELURUHAN ===

19) Δ ASUMSI INDEPENDENSI TIDAK TERPENUHI

20) (Ada bukti autokorelasi yang signifikan secara statistik)

21)

22) SARAN PERBAIKAN:

23) • Autokorelasi positif terdeteksi:

24) - Periksa urutan data (pastikan data time series diurutkan dengan benar)

- 25) - Pertimbangkan menambahkan lag variables
 26) - Gunakan model ARIMA atau time series lainnya
 27) - Transformasi data atau penambahan variabel kontrol
 28) - Pertimbangkan menggunakan Generalized Least Squares (GLS)
 29)
 30) CATATAN:
 31) • Nilai Durbin-Watson berkisar 0-4
 32) • Nilai ~2 menunjukkan tidak ada autokorelasi
 33) • Nilai <2 menunjukkan autokorelasi positif
 34) • Nilai >2 menunjukkan autokorelasi negatif

Hasil pengujian menunjukkan bahwa hipotesis nol ditolak (H_0 ditolak), mengindikasikan adanya autokorelasi positif antar residual. Pelanggaran asumsi ini dapat menyebabkan estimator menjadi tidak efisien dan standar error yang bias, sehingga diperlukan perlakuan khusus seperti metode lagging.

d) Nonmultikolinieritas

- ```

1) === UJI HOMOSKEDASTISITAS (BREUSCH-PAGAN) ===

2)

3) studentized Breusch-Pagan test

4)

5) data: mlr_model()

6) BP = 5.5541, df = 2, p-value = 0.06222

7)

8) INTERPRETASI:

9) P-value (0.0622) > 0.05

```

10) KESIMPULAN: Varians residual homogen (**asumsi homoskedastisitas TERPENUHI**)

Hasil uji menunjukkan bahwa tidak terdapat multikolinearitas yang serius antar variabel independen. Hal ini mengindikasikan bahwa variabel LOWEDU dan ILLITERATE tidak memiliki korelasi linier yang sempurna, sehingga kedua variabel dapat digunakan secara bersamaan dalam model regresi

## 2. Transformasi Data dan Pengujian Asumsi - Tahap 2

Untuk mengatasi pelanggaran asumsi normalitas dan autokorelasi, dilakukan transformasi data menggunakan logaritma natural dan metode lagging melalui dashboard CLEAR-IN.

### a) Transformasi Logaritma Natural

Variabel-variabel ditransformasi menjadi:

- 1) POVERTY → POVERTY\_log
- 2) LOWEDU → LOWEDU\_log
- 3) ILLITERATE → ILLITERATE\_log

Variabel untuk Transformasi:

Pilih variabel numerik untuk transformasi:

POVERTY

Jenis Transformasi:

Logaritma Natural (ln)

Terapkan Transformasi

Hasil Transformasi:

```
== HASIL TRANSFORMASI ==
Variabel asli: POVERTY
Transformasi: log
Variabel baru: POVERTY_log
```

Ringkasan data sebelum transformasi:  
Min. 1st Qu. Median Mean 3rd Qu. Max.  
1.760 7.305 11.100 12.943 15.980 43.630

Ringkasan data setelah transformasi:  
Min. 1st Qu. Median Mean 3rd Qu. Max.  
0.5653 1.9885 2.4069 2.3842 2.7713 3.7757

Variabel untuk Transformasi:

Pilih variabel numerik untuk transformasi:

LOWEDU

Jenis Transformasi:

Logaritma Natural (ln)

Terapkan Transformasi

Hasil Transformasi:

```
== HASIL TRANSFORMASI ==
Variabel asli: LOWEDU
Transformasi: log
Variabel baru: LOWEDU_log
```

```
Ringkasan data sebelum transformasi:
Min. 1st Qu. Median Mean 3rd Qu. Max.
9.298 25.921 33.296 32.448 39.669 55.416
```

```
Ringkasan data setelah transformasi:
Min. 1st Qu. Median Mean 3rd Qu. Max.
2.230 3.255 3.505 3.432 3.681 4.015
```

Variabel untuk Transformasi:

Pilih variabel numerik untuk transformasi:

ILLITERATE

Jenis Transformasi:

Logaritma Natural (ln)

Terapkan Transformasi

Hasil Transformasi:

```
== HASIL TRANSFORMASI ==
Variabel asli: ILLITERATE
Transformasi: log
Variabel baru: ILLITERATE_log
```

```
Ringkasan data sebelum transformasi:
Min. 1st Qu. Median Mean 3rd Qu. Max.
1.541 4.382 6.366 8.323 9.114 66.099
```

```
Ringkasan data setelah transformasi:
Min. 1st Qu. Median Mean 3rd Qu. Max.
0.4327 1.4775 1.8510 1.8863 2.2098 4.1912
```

Variabel bebas dan terikat telah ditransformasi menjadi variabel baru sebagai berikut:

Data Hasil Pemrosesan Unduh Data Terproses

| POVERTY_log       | LOWEDU_log        | ILLITERATE_log    |
|-------------------|-------------------|-------------------|
| 3.005682604407159 | 3.244624920354664 | 1.613237322725055 |
| 3.096029994869355 | 3.357757898533258 | 2.396318595040439 |
| 2.644044871126298 | 3.393818262808489 | 2.0435175164844   |
| 2.698673039289613 | 2.820877695357768 | 1.89813997232179  |
| 2.724579503053421 | 3.491640478531834 | 1.894350012903252 |

Showing 1 to 5 of 511 entries

Previous 1 2 3 4 5 ... 103 Next

Hasil Uji Normalitas setelah Transformasi:

```

1) === UJI NORMALITAS RESIDUAL ===

2)

3)

4) Shapiro-Wilk normality test

5)

6) data: residuals

7) W = 0.99612, p-value = 0.2455

8)

9)

10) INTERPRETASI (shapiro-wilk):
11) P-value (0.2455) > 0.05

12) KESIMPULAN: Residual berdistribusi normal (asumsi normalitas TERPENUHI)

13)

14) Asymptotic one-sample Kolmogorov-Smirnov test

15)

```

```

16) data: residuals

17) D = 0.026526, p-value = 0.8648

18) alternative hypothesis: two-sided

19) INTERPRETASI (kolmogorov-smirnov):

20) P-value (0.8648) > 0.05

21) KESIMPULAN: Residual berdistribusi normal (asumsi normalitas T
ERPENUHI)

```

Transformasi logaritma natural berhasil memperbaiki distribusi data, sehingga asumsi normalitas residual kini terpenuhi. Hal ini memungkinkan penggunaan inferensi statistik yang valid dalam model regresi.

### b) Transformasi Lagging

Untuk mengatasi masalah autokorelasi, dilakukan transformasi lagging pada variabel yang telah di-log, menghasilkan:

- 1) POVERTY\_log → POVERTY\_log\_lag
- 2) LOWEDU\_log → LOWEDU\_log\_lag
- 3) ILLITERATE\_log → ILLITERATE\_log\_lag

Variabel untuk Transformasi:

Pilih variabel numerik untuk transformasi:

POVERTY\_log

Jenis Transformasi:

lag (nonautokorelasi)

Terapkan Transformasi

Hasil Transformasi:

```

==== HASIL TRANSFORMASI ====
Variabel asli: POVERTY_log
Transformasi: lag
Variabel baru: POVERTY_log_lag

Ringkasan data sebelum transformasi:
Min. 1st Qu. Median Mean 3rd Qu. Max.
0.5653 1.9885 2.4069 2.3842 2.7713 3.7757

Ringkasan data setelah transformasi:
Min. 1st Qu. Median Mean 3rd Qu. Max.
-2.251809 -0.266418 -0.015297 -0.001111 0.271553 1

```

Variabel untuk Transformasi:

Pilih variabel numerik untuk transformasi:

Jenis Transformasi:

Terapkan Transformasi

Hasil Transformasi:

```
== HASIL TRANSFORMASI ==
Variabel asli: LOWEDU_log
Transformasi: lag
Variabel baru: LOHEDU_log_lag

Ringkasan data sebelum transformasi:
 Min. 1st Qu. Median Mean 3rd Qu. Max.
 2.230 3.255 3.505 3.432 3.681 4.015

Ringkasan data setelah transformasi:
 Min. 1st Qu. Median Mean 3rd Q
-1.0618289 -0.1479233 0.0074293 0.0001593 0.16549
```

Variabel untuk Transformasi:

Pilih variabel numerik untuk transformasi:

Jenis Transformasi:

Terapkan Transformasi

Hasil Transformasi:

```
== HASIL TRANSFORMASI ==
Variabel asli: ILLITERATE_log
Transformasi: lag
Variabel baru: ILLITERATE_log_lag

Ringkasan data sebelum transformasi:
 Min. 1st Qu. Median Mean 3rd Qu. Max.
 0.4327 1.4775 1.8510 1.8863 2.2098 4.1912

Ringkasan data setelah transformasi:
 Min. 1st Qu. Median Mean 3rd Q
-2.2310128 -0.2469977 -0.0028210 -0.0002036 0.28377
```

Variabel bebas dan terikat telah ditransformasi menjadi variabel baru sebagai berikut:

**Data Hasil Pemrosesan**

[Unduh Data Terproses](#)

| POVERTY_log_lag     | LOWEDU_log_lag      | ILLITERATE_log_lag    |
|---------------------|---------------------|-----------------------|
| 0.09034739046219586 | 0.1131329781785944  | 0.7830812723153835    |
| -0.4519851237430572 | 0.03606036427523041 | -0.3528010785560385   |
| 0.05462816816331539 | -0.5729405674507211 | -0.1453775441626102   |
| 0.02590646376380734 | 0.6707627831740659  | -0.003789959418537991 |

Showing 1 to 5 of 511 entries

Previous [1](#) [2](#) [3](#) [4](#) [5](#) ... [103](#) Next

### Hasil Uji Durbin-Watson::

```

1) === UJI AUTOKORELASI (DURBIN-WATSON) ===

2)

3) Durbin-Watson test

4)

5) data: mlr_model()

6) DW = 2.7908, p-value = 1

7) alternative hypothesis: true autocorrelation is greater than 0

8)

9) === INTERPRETASI ===

10) Nilai Durbin-Watson: 2.7908

11) P-value: 1

12)

13) INTERPRETASI NILAI DURBIN-WATSON:

14) X Terdapat autokorelasi negatif (asumsi independensi TIDAK TERPENUHI)

15)

16) INTERPRETASI P-VALUE:

```

- 17) ✓ P-value ( 1 ) > 0.05: Tidak ada bukti autokorelasi yang signifikan
- 18)
- 19) === KESIMPULAN KESELURUHAN ===
- 20)  ASUMSI INDEPENDENSI TERPENUHI - Model dapat dilanjutkan
- 21) (Tidak ada bukti autokorelasi yang signifikan secara statistik)
- 22)
- 23) CATATAN:
- 24) • Nilai Durbin-Watson berkisar 0-4
- 25) • Nilai ~2 menunjukkan tidak ada autokorelasi
- 26) • Nilai <2 menunjukkan autokorelasi positif
- 27) • Nilai >2 menunjukkan autokorelasi negatif
- 28)

Nilai Durbin-Watson sebesar 2.7908 dengan p-value sebesar 1 menunjukkan bahwa tidak terdapat bukti autokorelasi yang signifikan secara statistik. Meskipun nilai DW > 2 mengindikasikan kemungkinan autokorelasi negatif berdasarkan nilai statistik, namun p-value yang sangat tinggi ( $1 > 0.05$ ) menunjukkan bahwa autokorelasi tersebut tidak signifikan secara statistik. Dengan demikian, asumsi independensi residual telah terpenuhi dan model dapat dilanjutkan untuk analisis.

### 3. Hasil Estimasi Model Regresi

- 1) === RINGKASAN MODEL REGRESI LINEAR BERGANDA ===
- 2)
- 3) Formula: POVERTY\_log\_lag ~ LOWEDU\_log\_lag + ILLITERATE\_log\_lag
- 4)
- 5) Call:

```

6) lm(formula = formula_obj, data = values$processedData)

7)

8) Residuals:

9) Min 1Q Median 3Q Max
10) -1.65102 -0.22883 -0.00302 0.24228 1.74743

11)

12) Coefficients:

13) Estimate Std. Error t value Pr(>|t|)

14) (Intercept) -0.001085 0.019036 -0.057 0.955
15) LOWEDU_log_lag 0.299245 0.068625 4.361 1.57e-05 ***
16) ILLITERATE_log_lag 0.363037 0.043361 8.372 5.55e-16 ***
17) ---
18) Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
19)

20) Residual standard error: 0.4299 on 507 degrees of freedom
21) (1 observation deleted due to missingness)
22) Multiple R-squared: 0.2548, Adjusted R-squared: 0.2519
23) F-statistic: 86.69 on 2 and 507 DF, p-value: < 2.2e-16
24)
25)
26) === INTERPRETASI ===
27) R-squared: 0.2548 - Model menjelaskan 25.48 % variasi dalam va
 riable dependen
28) Adjusted R-squared: 0.2519
29) F-statistic: 86.6945 dengan p-value: < 0.001
30)
31) INTERPRETASI KOEFISIEN:
32) - Intercept: -0.0011 (tidak signifikan)

```

33) - LOWEDU\_log\_lag : Setiap peningkatan 1 unit akan meningkatkan Y sebesar 0.2992 (signifikan)

34) - ILLITERATE\_log\_lag : Setiap peningkatan 1 unit akan meningkatkan Y sebesar 0.363 (signifikan)

a) Persamaan Regresi Final

Berdasarkan analisis regresi linier berganda dengan variabel yang telah ditransformasi, diperoleh persamaan:

$$\widehat{POVERTY} = -0.001085 + 0.299245 LOWEDU + 0.363037 ILLITERATE$$

b) Evaluasi Kesesuaian Model

- 1) R-squared: 0.2548 (25.48%)
- 2) Adjusted R-squared: 0.2519 (25.19%)
- 3) F-statistic: 86.69 dengan p-value < 2.2e-16
- 4) Residual standard error: 0.4299

Nilai R-squared sebesar 0.2548 menunjukkan bahwa model mampu menjelaskan 25.48% variasi dalam tingkat kemiskinan, sedangkan 74.52% sisanya dijelaskan oleh faktor lain yang tidak dimasukkan dalam model. Meskipun nilai R-squared relatif rendah, hal ini umum terjadi dalam data cross-sectional sosial ekonomi yang memiliki variabilitas tinggi. Nilai F-statistic yang signifikan (p-value < 0.001) mengindikasikan bahwa model secara keseluruhan signifikan secara statistik

c) Interpretasi Koefisien Regresi

- 1) Konstanta (Intercept): -0.001085

Nilai: -0.001085 (tidak signifikan, p-value = 0.955)

Ketika kedua variabel independen bernilai nol, tingkat kemiskinan diperkirakan sebesar -0.001085%. Namun, koefisien ini tidak signifikan secara statistik.

2) Koefisien LOWEDU\_Ln\_lag: 0.299245

Nilai: 0.299245 (signifikan, p-value = 1.57e-05)

Setiap peningkatan 1% dalam tingkat pendidikan rendah akan meningkatkan tingkat kemiskinan sebesar 0.299%, dengan asumsi variabel lain konstan. Koefisien ini signifikan pada tingkat kepercayaan 95%.

3) Koefisien ILLITERATE\_Ln\_lag: 0.363037

Nilai: 0.363037 (signifikan, p-value = 5.55e-16)

Setiap peningkatan 1% dalam tingkat buta huruf akan meningkatkan tingkat kemiskinan sebesar 0.363%, dengan asumsi variabel lain konstan. Koefisien ini signifikan pada tingkat kepercayaan 95%.

#### D. Jawaban Rumusan Masalah

Berdasarkan hasil analisis regresi linier berganda yang telah dilakukan, berikut adalah jawaban untuk setiap rumusan masalah:

1. Apakah tingkat pendidikan rendah (LOWEDU) berpengaruh signifikan terhadap tingkat kemiskinan?

Jawab:

Tingkat pendidikan rendah (LOWEDU) berpengaruh positif dan signifikan terhadap tingkat kemiskinan. Hal ini dibuktikan dengan:

- a. Koefisien regresi: 0.299245
- b. P-value: 1.57e-05 (< 0.05)
- c. Tingkat signifikansi: 95% ( $\alpha = 0.05$ )

Artinya, setiap peningkatan 1% tingkat pendidikan rendah akan meningkatkan tingkat kemiskinan sebesar 0.299%.

2. Apakah tingkat buta huruf (ILLITERATE) berpengaruh signifikan terhadap tingkat kemiskinan?

Jawab:

Tingkat buta huruf (ILLITERATE) berpengaruh positif dan signifikan terhadap tingkat kemiskinan. Hal ini dibuktikan dengan:

- a. Koefisien regresi: 0.363037
- b. P-value: 5.55e-16 (< 0.05)
- c. Tingkat signifikansi: 95% ( $\alpha = 0.05$ )

Artinya, setiap peningkatan 1% tingkat buta huruf akan meningkatkan tingkat kemiskinan sebesar 0.363%.

3. Seberapa besar kontribusi kedua variabel tersebut dalam menjelaskan variasi tingkat kemiskinan?

Jawab:

Berdasarkan nilai R-squared, kedua variabel (tingkat pendidikan rendah dan tingkat buta huruf) secara bersama-sama mampu menjelaskan 25.48% variasi tingkat kemiskinan di wilayah penelitian. Hal ini menunjukkan bahwa:

- a. 25.48% variasi kemiskinan dijelaskan oleh model
- b. 74.52% variasi kemiskinan dijelaskan oleh faktor lain di luar model
- c. Model secara keseluruhan signifikan (F-statistic: 86.69, p-value < 2.2e-16)

Tingkat buta huruf memiliki pengaruh yang lebih besar dibandingkan tingkat pendidikan rendah ( $0.363 > 0.299$ ).

## E. Kesimpulan

Berdasarkan hasil analisis regresi linier berganda, dapat disimpulkan bahwa:

1. Tingkat pendidikan rendah berpengaruh positif dan signifikan terhadap tingkat kemiskinan dengan koefisien 0.299 dan tingkat signifikansi 99%.
2. Tingkat buta huruf berpengaruh positif dan signifikan terhadap tingkat kemiskinan dengan koefisien 0.363 dan tingkat signifikansi 99%.
3. Variabel tingkat buta huruf memiliki pengaruh yang lebih besar terhadap kemiskinan dibandingkan dengan tingkat pendidikan rendah ( $0.363 > 0.299$ ).
4. Model regresi yang dihasilkan telah memenuhi asumsi klasik setelah dilakukan transformasi logaritma dan lagging, sehingga hasil estimasi dapat dipercaya.

5. Model mampu menjelaskan 25.48% variasi tingkat kemiskinan, yang menunjukkan bahwa masih terdapat faktor-faktor lain yang mempengaruhi kemiskinan selain tingkat pendidikan dan literasi.

Hasil penelitian ini mengonfirmasi pentingnya investasi dalam bidang pendidikan dan program pemberantasan buta huruf sebagai strategi pengentasan kemiskinan yang efektif.

## **BAB V**

### **PENUTUP**

#### A. Kesimpulan

Dashboard CLEAR-IN (Comprehensive Local Exploratory Analysis and Reporting Interface for Indonesia) telah berhasil dikembangkan sebagai bagian dari laporan Ujian Akhir Semester mata kuliah Komputasi Statistik. Selain itu, dashboard ini juga berfungsi sebagai platform analisis statistik terintegrasi yang dirancang untuk mengolah data kerentanan sosial di tingkat kabupaten/kota di seluruh Indonesia.

Berdasarkan hasil implementasi sistem, dapat disimpulkan beberapa hal berikut:

##### 1. Integrasi Data Multidimensional

Dashboard mampu menggabungkan berbagai jenis data, seperti Social Vulnerability Index (SOVI) dan data spasial (peta), ke dalam satu platform yang menyatu dan mudah digunakan.

##### 2. Kemampuan Analisis Lengkap

Platform menyediakan berbagai alat analisis, mulai dari eksplorasi data dasar hingga analisis statistik lanjutan seperti regresi linier berganda, sehingga pengguna bisa melakukan analisis yang komprehensif.

##### 3. Tampilan Ramah Pengguna

Struktur sistem menggunakan navigasi tab yang memudahkan pengguna untuk berpindah antar menu dan melakukan analisis secara runtut, serta dilengkapi dengan visualisasi data yang interaktif dan responsif.

##### 4. Mendukung Pengambilan Keputusan

Fitur-fitur seperti pengujian otomatis, interpretasi hasil, dan sistem pelaporan mendukung pengambilan keputusan berbasis data, baik untuk peneliti, akademisi, maupun pembuat kebijakan.

##### 5. Mendukung Penelitian yang Dapat Direplikasi

Hasil analisis bisa disimpan dan diekspor, sehingga pengguna bisa mengulang analisis yang sama di waktu berbeda, atau membagikan hasilnya untuk keperluan ilmiah secara bertanggung jawab.

## B. Saran

### 1. Pengembangan Lanjutan

Menambahkan fitur-fitur baru seperti analisis spasial lanjutan, clustering, atau integrasi data real-time.

### 2. Optimalisasi Sistem

Meningkatkan efisiensi loading data, performa tampilan, dan kompatibilitas lintas perangkat.

### 3. Kolaborasi dan Diseminasi

Mendorong kerja sama dengan akademisi, pemerintah daerah, dan organisasi masyarakat sipil untuk memanfaatkan platform ini secara luas.

### 4. Keberlanjutan Platform

Menjaga agar platform ini terus diperbarui, didukung oleh komunitas, dan digunakan secara aktif dalam pengambilan keputusan pembangunan.