

# **LAPORAN TUGAS AKHIR REGRESI TERAPAN**

**Analisis Regresi Tingkat Kriminal : Pengaruh Variabel Sosial dan Ekonomi Terhadap  
Kejadian Kriminal**



## **Anggota Kelompok :**

|                           |           |
|---------------------------|-----------|
| Wanda Desi Rahmawati      | 164221026 |
| Fauziah Hamidah Al Hanief | 164221055 |
| Arkan Syafiq Attaqy       | 164221062 |
| Riko Handono              | 164221066 |

**PROGRAM STUDI S1 TEKNOLOGI SAINS DATA  
FAKULTAS TEKNOLOGI MAJU DAN MULTIDISIPLIN  
UNIVERSITAS AIRLANGGA  
SURABAYA  
2023**

## DAFTAR ISI

|  |           |
|--|-----------|
| <b>DAFTAR ISI.....</b>                     | <b>2</b>  |
| <b>DAFTAR TABEL .....</b>                  | <b>3</b>  |
| <b>DAFTAR GAMBAR.....</b>                  | <b>4</b>  |
| <b>BAB I PENDAHULUAN.....</b>              | <b>4</b>  |
| 1.1. Latar Belakang .....                  | 4         |
| 1.2. Rumusan Masalah.....                  | 5         |
| 1.3. Tujuan Penelitian .....               | 5         |
| <b>BAB II TINJAUAN PUSTAKA .....</b>       | <b>6</b>  |
| 2.1. Tinjauan Pustaka Statistika .....     | 6         |
| 2.2. Tinjauan Pustaka Non-Statistika.....  | 7         |
| <b>BAB III METODOLOGI.....</b>             | <b>9</b>  |
| 3.1. Sumber Data.....                      | 9         |
| 3.2. Variabel Penelitian.....              | 9         |
| 3.3. Metode Penelitian .....               | 10        |
| <b>BAB IV ANALISIS DAN PEMBAHASAN.....</b> | <b>12</b> |
| 4.1. Exploratory Data Analysis .....       | 12        |
| 4.2. Analisis Regresi .....                | 16        |
| 4.3. Pengujian Hipotesis .....             | 18        |
| 4.4. Pemilihan Model Terbaik .....         | 23        |
| <b>BAB V KESIMPULAN DAN SARAN.....</b>     | <b>24</b> |
| 5.1. Kesimpulan .....                      | 24        |
| 5.2. Saran .....                           | 24        |
| <b>DAFTAR PUSTAKA.....</b>                 | <b>25</b> |
| <b>LAMPIRAN.....</b>                       | <b>26</b> |

## **DAFTAR TABEL**

|  |    |
|--|----|
| Tabel 1: Variabel Penelitian.....              | 9  |
| Tabel 2 : Analisis Deskriptif .....            | 16 |
| Tabel 3 : Output OLS Regression.....           | 16 |
| Tabel 4 : Output Robust Regression .....       | 17 |
| Tabel 5 : Output Uji Serentak .....            | 18 |
| Tabel 6 : Output Uji Parsial .....             | 18 |
| Tabel 7 : Daerah Kritis Uji Autokorelasi ..... | 20 |
| Tabel 8 : Output Uji Autokorelasi .....        | 20 |
| Tabel 9 : Output Uji Heteroskedastisitas ..... | 21 |
| Tabel 10 : Output Uji Multikolinearitas.....   | 22 |
| Tabel 11 : Perbandingan Model.....             | 23 |

## DAFTAR GAMBAR

|   |    |
|---|----|
| Gambar 1 : Histogram.....                       | 12 |
| Gambar 2 : Scatter Plot .....                   | 13 |
| Gambar 3 : Heatmap .....                        | 13 |
| Gambar 4 : Density Plot.....                    | 14 |
| Gambar 5 : Boxplot.....                         | 15 |
| Gambar 6 : Grafik Normalitas .....              | 19 |
| Gambar 7 : Grafik Uji Heteroskedastisitas ..... | 21 |

## **BAB I**

### **PENDAHULUAN**

#### **1.1. Latar Belakang**

Kejahatan, dalam segala bentuknya, merugikan pihak lain dan melanggar norma serta Undang-Undang. Perspektif kejahatan tidak hanya ditinjau dari aspek kesusilaan, kesopanan, dan ketertiban masyarakat, tetapi juga melibatkan elemen-elemen sosial dan ekonomi. Kejahatan bisa dilakukan oleh siapa saja, termasuk mereka yang menghadapi kesulitan hidup atau bahkan individu dengan profesi tertentu. Kejahatan sendiri bersifat oportunistik, dan pelaku seringkali tidak mempertimbangkan konsekuensi panjang dari perbuatannya. Jenis kejahatan meliputi berbagai bidang, mulai dari kejahatan ekonomi hingga kejahatan sosial.

Manusia, sebagai makhluk sosial, dihadapkan pada tuntutan untuk berinteraksi dengan sesama, yang memiliki latar belakang beragam seperti suku, ras, agama, budaya, dan kebiasaan. Perbedaan-perbedaan ini seringkali menjadi pemicu kejahatan (Anjari, 2014). Tak hanya faktor sosial, tetapi juga faktor ekonomi seperti kemiskinan, pengangguran, pertumbuhan ekonomi, dan kepadatan penduduk dapat berperan dalam peningkatan tingkat kriminalitas (Purwanti, 2019).

Pada penelitian ini, peneliti ingin melakukan pemodelan analisis regresi guna mengidentifikasi faktor-faktor yang berkontribusi terhadap tingginya angka kejahatan kekerasan. Variabel independen yang digunakan dalam penelitian ini mencakup persentase penduduk yang tinggal di wilayah metropolitan, persentase penduduk yang berpendidikan Sekolah Menengah Atas (SMA), persentase penduduk yang hidup di bawah garis kemiskinan, dan persentase penduduk yang merupakan orang tua tunggal atau *single parent*.

Analisis regresi yang dilakukan akan memberikan wawasan mendalam tentang bagaimana variabel sosial dan ekonomi dapat menjadi prediktor kejahatan kekerasan. Hasil penelitian diharapkan dapat memberikan kontribusi pada pemahaman lebih lanjut terkait faktor-faktor yang perlu diperhatikan dalam upaya pencegahan dan pengendalian kejahatan. Selain itu, penelitian ini dapat menjadi dasar bagi kebijakan publik yang lebih efektif dalam mengatasi permasalahan keamanan masyarakat.

## **1.2. Rumusan Masalah**

Berdasarkan latar belakang masalah, maka dapat dirumuskan masalah sebagai berikut

1. Bagaimana pengaruh variabel sosial dan ekonomi berdasarkan faktor-faktor metro, sma, *single*, dan kemiskinan terhadap tindak kriminalitas?
2. Apa pemodelan regresi terbaik untuk menentukan analisis regresi faktor-faktor metro, sma, *single*, dan kemiskinan terhadap tindak kriminalitas?

## **1.3. Tujuan Penelitian**

Berdasarkan rumusan masalah diatas, tujuan yang ingin dicapai dalam penulisan laporan ini adalah :

1. Untuk mengetahui pengaruh variabel sosial dan ekonomi berdasarkan faktor-faktor metro, sma, *single*, dan kemiskinan terhadap tindak kriminalitas.
2. Untuk mengetahui pemodelan regresi terbaik untuk menentukan analisis regresi faktor-faktor metro, sma, *single*, dan kemiskinan terhadap tindak kriminalitas.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1. Tinjauan Pustaka Statistika**

##### **2.1.1. Statistika**

Menurut Budiyono (2017), statistika mencakup pengetahuan yang terkait dengan proses pengumpulan, penyusunan, penyajian, dan pengambilan kesimpulan (generalisasi) terhadap populasi dengan menggunakan data sampel yang diperoleh. Dengan demikian, secara umum, statistika dapat diklasifikasikan menjadi dua kategori utama, yaitu statistika deskriptif dan statistika inferensial

##### **2.1.2. Uji Asumsi Klasik**

Menurut Ghozali (2018), uji asumsi klasik merupakan langkah awal yang diterapkan sebelum melakukan analisis regresi linier berganda. Tujuan dari uji ini adalah untuk memastikan bahwa koefisien regresi tidak terpengaruh oleh bias, konsisten, dan memiliki tingkat ketepatan yang baik dalam proses estimasi. Pengujian asumsi klasik dilakukan untuk menunjukkan bahwa data telah memenuhi syarat normalitas, tidak mengalami multikolinieritas, autokorelasi, dan heteroskedastisitas.

##### **2.1.3. Pencilan**

Montgomery dan Peck (1992) menjelaskan bahwa pencilan merujuk pada pengamatan yang menonjol dari sebagian besar data. Pencilan terjadi ketika nilai residual secara signifikan lebih besar daripada nilai residual yang lain, seringkali terletak pada tiga atau empat deviasi standar dari rata-rata. Pencilan dapat diidentifikasi sebagai titik-titik data yang tidak sejajar dengan pola mayoritas titik data lainnya. Pada dasarnya, pencilan mencirikan data yang berbeda secara signifikan dari sebagian besar pengamatan lain dalam dataset. Analisis hasil menunjukkan bahwa pencilan dan pengamatan berpengaruh secara signifikan terhadap perubahan nilai intercept, slope,  $R^2$ , dan  $s^2$ .

#### 2.1.4. Analisis Regresi

##### 2.1.4.1. *Ordinary Least Squares*

Metode OLS diterapkan pada dataset kejahatan untuk mengidentifikasi hubungan linier antara variabel independen (metro, SMA, miskin, *single*) dan variabel dependen (kriminal) dengan meminimalkan selisih kuadrat antara nilai yang diobservasi dan diprediksi. OLS menyediakan estimasi parameter regresi yang optimal, tetapi keandalannya dapat dipengaruhi oleh ketidakpastian dan kompleksitas struktur dalam data kejahatan. Namun, OLS tetap relevan sebagai pendekatan awal untuk melihat hubungan linier antar variabel tersebut. Berikut adalah model regresi linier:

Model Regresi Linier

$$y = X\beta + \varepsilon, \quad \text{dengan } \varepsilon \sim N(\mathbf{0}, I\sigma^2)$$

##### 2.1.4.2. *Robust Regression*

*Robust Regression* adalah metode yang dirancang untuk meningkatkan ketahanan terhadap *outlier* atau pencilan dalam data. Outlier dapat memiliki pengaruh yang signifikan pada hasil regresi, terutama jika menggunakan metode OLS. *Robust Regression* menggunakan fungsi kerugian (*loss function*) yang lebih tahan terhadap nilai-nilai ekstrim, sehingga memberikan bobot yang lebih rendah pada outlier.

## 2.2. Tinjauan Pustaka Non-Statistika

### 2.2.1. Teori Kriminologi dan Faktor-Faktor Kejahatan

Kejahatan sebagai fenomena kompleks telah menarik perhatian banyak peneliti di bidang kriminologi. Teori-teori kriminologi seperti teori kesempatan, teori kontrol sosial, dan teori ketidaksetaraan sosial telah digunakan untuk memahami penyebab dan faktor-faktor yang mempengaruhi tingkat kejahatan. Studi-studi terdahulu menunjukkan bahwa faktor ekonomi, pendidikan, wilayah, dan kondisi sosial memainkan peran penting dalam memicu kejahatan.

### 2.2.2. Hubungan antara Faktor Wilayah dan Kejahatan



Studi kriminologi telah mengidentifikasi hubungan antara wilayah metropolitan dan tingkat kejahatan. Wilayah metropolitan sering kali menjadi fokus penelitian karena kepadatan penduduk dan kompleksitas struktur sosialnya mempengaruhi dinamika kejahatan. Teori kriminologi perkotaan menyoroti ketidaksetaraan, peluang kejahatan, dan tantangan keamanan di wilayah metropolitan. Penelitian terbaru, menunjukkan bahwa wilayah metropolitan dengan persentase penduduk yang tinggi cenderung memiliki tingkat kejahatan yang lebih tinggi.

#### **2.2.3. Hubungan antara Faktor Pendidikan dan Kejahatan**

Tingkat pendidikan masyarakat juga memiliki korelasi dengan tingkat kejahatan. Penelitian menunjukkan bahwa masyarakat yang lebih teredukasi cenderung memiliki tingkat kejahatan yang lebih rendah. Oleh karena itu, analisis regresi yang melibatkan variabel pendidikan seperti persentase penduduk SMA dapat memberikan wawasan tentang bagaimana pendidikan dapat berperan dalam mengurangi tingkat kejahatan. Penelitian terbaru menunjukkan bahwa tingkat pendidikan yang lebih tinggi dapat membantu mengurangi tingkat kejahatan.

#### **2.2.4. Hubungan antara Faktor Orang Tua Tunggal dan Kejahatan**

Peran orang tua tunggal dalam suatu masyarakat memainkan peran kunci dalam dinamika kejahatan. Studi terbaru menunjukkan bahwa kehadiran orang tua tunggal dapat memberikan dampak signifikan pada tingkat kejahatan, terutama pada rentang usia anak-anak. Anak-anak dari keluarga dengan orang tua tunggal cenderung lebih rentan terhadap risiko kejahatan karena faktor-faktor seperti kurangnya pengawasan dan dukungan emosional yang optimal.

### BAB III

#### METODOLOGI

##### 3.1. Sumber Data

*Project* ini menggunakan data sekunder yang berasal dari dosen pembimbing Indah Fahmiah S.Si., M.Stat. Data yang digunakan berisi 51 sampel yang memuat informasi tentang variabel kriminal, metro, SMA, miskin, dan *single*. Data dapat diakses pada link berikut: <https://bit.ly/DataKejahatanKelompok3>. Perlu diakui bahwa penggunaan data sekunder memiliki keterbatasan tertentu. Beberapa batasan mungkin seperti minimnya informasi data, kurangnya kontrol dalam proses pengumpulan data, dan lain-lain.

##### 3.2. Variabel Penelitian

Variabel yang digunakan pada *project* ini terdiri dari variabel dependen atau variabel respons (Y) dan variabel independen atau variabel prediktor (X). Variabel dependen (Y) yaitu variabel yang nilainya tergantung variabel lain sedangkan variabel independen (X) yaitu variabel yang diasumsikan memiliki pengaruh terhadap nilai variabel respons. Berikut variabel yang digunakan:

| Variabel | Nama Variabel | Deskripsi   | Tipe    |
|----------|---------------|---|---------|
| Y        | Kriminal      | Jumlah kejahatan kekerasan per 10000 orang  | Numerik |
| X1       | Metro         | Persentase penduduk yang tinggal di wilayah metropolitan                            | Numerik |
| X2       | SMA           | Persentase penduduk yang berada pada jenjang pendidikan sekolah menengah atas (SMA) | Numerik |
| X3       | Miskin        | Persentase penduduk yang hidup di bawah garis kemiskinan                            | Numerik |
| X4       | <i>Single</i> | Persentase penduduk yang menjadi orang tua <i>single</i> atau <i>single parent</i>  | Numerik |

(Tabel 1: Variabel Penelitian)

### 3.3. Metode Penelitian

#### 3.3.1. Teknik Analisis

##### 3.3.1.1. *Ordinary Least Squares Regression*

Penelitian menggunakan metode *OLS Regression* untuk memperoleh model regresi berganda dari hubungan variabel dependen (kriminal) dan variabel independen (metro, SMA, miskin, dan *single*). Model regresi yang ingin diperoleh yaitu:

$$Y = \beta + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

Metode *OLS Regression* digunakan karena metode tersebut bersifat linier, tak bias, dan varians minimum atau *BLUE (Best Linear Unbiased Estimator)* serta mudah diinterpretasi.

##### 3.3.1.2. Uji Hipotesis

Uji hipotesis meliputi uji signifikansi koefisien regresi dan uji asumsi. Uji signifikansi koefisien regresi yang dilakukan adalah uji serentak menggunakan uji-F dan uji parsial menggunakan uji-t. Uji signifikansi koefisien regresi dilakukan untuk evaluasi pengaruh variabel independen dan validasi pengaruh secara statistik.

Uji asumsi meliputi uji normalitas, autokorelasi, heteroskedastisitas, dan multikolinearitas. Uji normalitas dilakukan dengan *Kolmogorov-Smirnov* dan *QQ-plot* untuk memastikan bahwa residual berdistribusi normal. Uji autokorelasi dilakukan dengan *Durbin Watson Test* dan *plot residual vs time order* untuk menguji apakah dalam model ada korelasi antara error pada periode  $t$  dengan error pada periode  $t-1$ . Model yang baik harus bebas autokorelasi. Uji heteroskedastisitas dilakukan dengan *Glejser Test* dan plot residual dengan nilai prediksi untuk memastikan model homoskedastisitas. Uji multikolinearitas dilakukan dengan perhitungan nilai VIF (Variance Inflation Factor) untuk memastikan bahwa tiap variabel independen tidak memiliki ketergantungan.

##### 3.3.1.3. *Robust Regression*

Penelitian menggunakan metode *Robust Regression* untuk mengatasi nilai *outlier* pada data dan mendapat model terbaik karena *OLS* sangat sensitif terhadap *outlier*.

#### 3.3.2. Prosedur Analisis

Analisis menggunakan bantuan tools SPSS dan python. Langkah analisis meliputi:

- a. Melakukan *exploratory data* untuk mengetahui karakteristik data
- b. Mencari model regresi menggunakan *OLS Regression*
- c. Menguji asumsi model
- d. Mencari model regresi menggunakan *Robust Regression*
- e. Menguji signifikansi koefisien regresi
- f. Membandingkan model dan mencari model terbaik

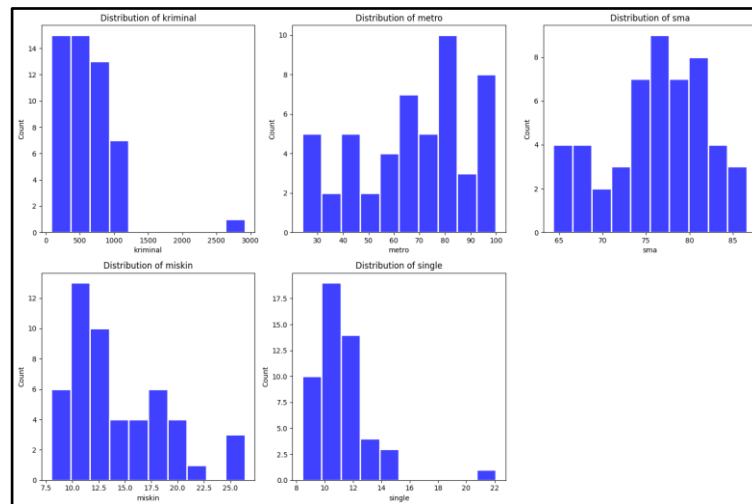
## BAB IV

### ANALISIS DAN PEMBAHASAN

#### 4.1. Exploratory Data Analysis

##### 4.1.1. Visualisasi Data

##### 4.1.1.1. Histogram

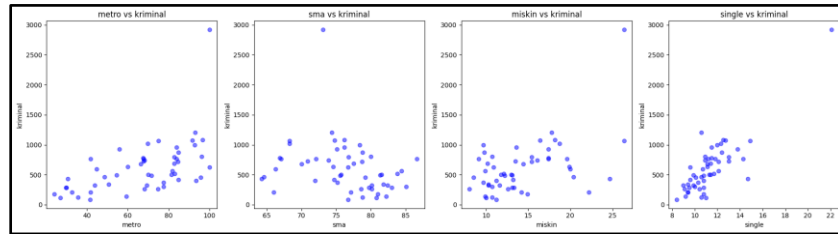


(Gambar 1 : Histogram)

#### Interpretasi :

Histogram di atas menjelaskan frekuensi dan distribusi data. Variabel kriminal memiliki rentang nilai kurang dari 3000 dan terdapat *outlier*. Variabel metro memiliki rentang nilai 0-100. Variabel SMA memiliki rentang nilai kurang dari 90. Variabel miskin memiliki rentang nilai kurang dari 30. Variabel *single* memiliki rentang nilai kurang dari 23 dan terdapat *outlier*. Variabel metro dan SMA cenderung miring ke kiri yang menandakan  $\text{mean} < \text{median} < \text{modus}$ . Variabel kriminal, single, dan miskin cenderung miring ke kanan yang menandakan  $\text{modus} < \text{median} < \text{mean}$ .

##### 4.1.1.2. Scatter Plot

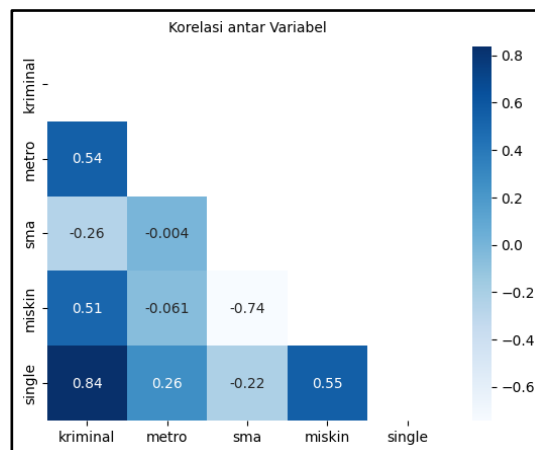


(Gambar 2 : Scatter Plot)

#### Interpretasi :

Scatter plot di atas menjelaskan hubungan antara variabel Y (kriminal) dan variabel X (metro, SMA, miskin, dan *single*). Variabel metro dan *single* membentuk titik ke atas sehingga dapat diidentifikasi memiliki hubungan positif. Variabel miskin dan SMA membentuk titik yang menyebar sehingga hubungan yang terbentuk dengan variabel kriminal tidak dapat diketahui dengan jelas secara visual.

#### 4.1.1.3. Heatmap

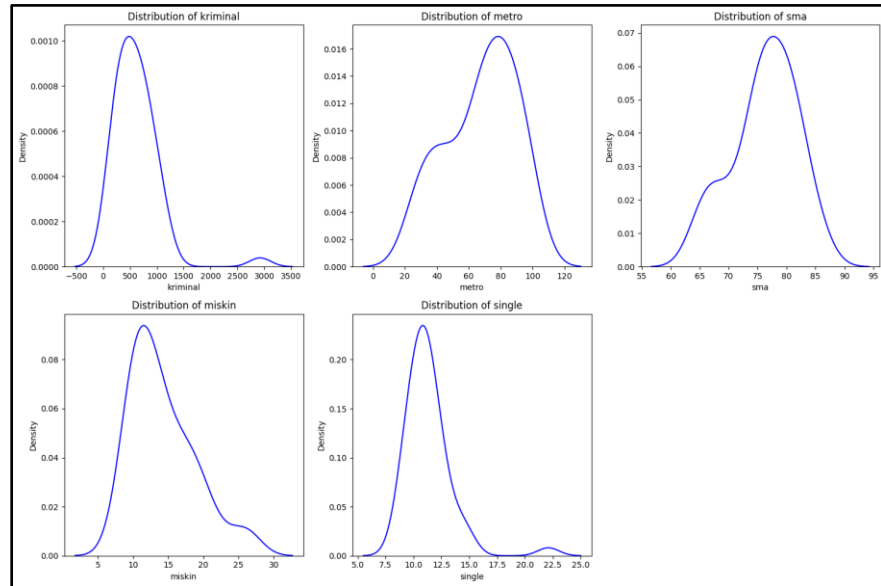


(Gambar 3 : Heatmap)

#### Interpretasi :

Heatmap di atas menunjukkan korelasi antara setiap variabel dengan variabel lainnya yang ditandai dengan warna gelap terang, dimana warna gelap menunjukkan korelasi positif yang sangat kuat antar variabel dan warna terang menunjukkan korelasi negatif. Dalam hal ini, korelasi positif terkuat dimiliki oleh variabel kriminal dan *single*, sedangkan korelasi negatif terkuat dimiliki oleh variabel SMA dan miskin.

#### 4.1.1.4. Density Plot

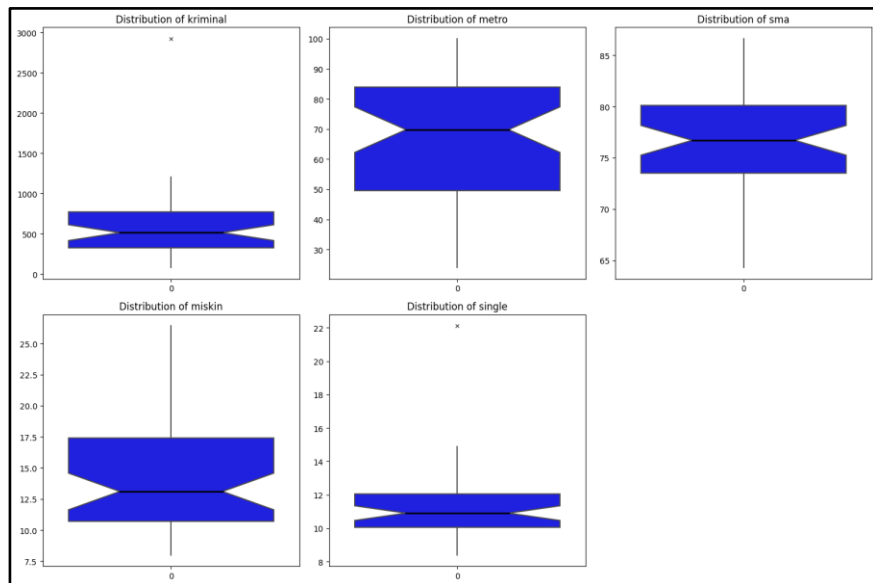


(Gambar 4 : Density Plot)

### Interpretasi :

Distribusi frekuensi kriminal menunjukkan mayoritas data berkumpul di sekitar nilai 500, dengan penurunan saat nilai kriminal meningkat. Grafik metro menunjukkan bahwa area metro memiliki frekuensi tertinggi di sekitar nilai 60. Distribusi sma memiliki dua puncak di sekitar 65 dan 80, menandakan adanya dua kelompok utama dalam data ini. Grafik miskin memiliki puncak tajam di sekitar nilai 5, menunjukkan konsentrasi tinggi pada tingkat kemiskinan tersebut. Distribusi *single* memiliki puncak di antara 15-17.5, menunjukkan kebanyakan data berkumpul pada rentang nilai tersebut.

#### 4.1.1.5. Boxplot



(Gambar 5 : Boxplot)

#### Interpretasi :

Dari analisis visual menggunakan boxplot, tampak bahwa variabel kriminal dan *single* menunjukkan keberadaan masing-masing satu pencilan (outlier), sementara ketiga variabel lainnya tidak menunjukkan adanya pencilan. Pencilan dalam variabel kriminal dan *single* dapat menjadi titik data yang signifikan atau ekstrem jika dibandingkan dengan sebagian besar data lainnya. Hal ini menyoroti kemungkinan adanya kejadian luar biasa atau tidak biasa yang perlu mendapatkan perhatian khusus dalam interpretasi data.

#### 4.1.2. Analisis Deskriptif

Analisis statistika deskriptif bertujuan untuk memberikan pemahaman dan gambaran awal tentang karakteristik data yang diteliti. Beberapa metode yang diterapkan dalam analisis ini diantaranya adalah perhitungan rata-rata (mean), median, Varians, standar deviasi, rentang (range), nilai minimal, dan nilai maksimal dari setiap dataset. Berikut adalah hasil dari analisis statistika deskriptif dari dataset yang diteliti :



| Variabel | Mean   | Median | Variance | SD     | Range | Sum    | Min  | Max   |
|----------|--------|--------|----------|--------|-------|--------|------|-------|
| Kriminal | 612.84 | 515    | 190754.4 | 436.75 | 2840  | 31255  | 82.0 | 2922  |
| Metro    | 67.39  | 69.8   | 472.66   | 21.74  | 76.0  | 3436.9 | 24.0 | 100.0 |
| SMA      | 76.22  | 76.7   | 30.65    | 5.53   | 22.3  | 3887.4 | 64.3 | 86.6  |
| Miskin   | 14.25  | 13.1   | 20.60    | 4.53   | 18.4  | 727.2  | 8.0  | 26.4  |
| Single   | 11.32  | 10.9   | 4.41     | 2.10   | 13.7  | 577.6  | 8.4  | 22.1  |

(Tabel 2 : Analisis Deskriptif)

## 4.2. Analisis Regresi

### 4.2.1. Ordinary Least Squares Regression

| OLS Regression Results  |                  |                     |          |       |           |          |
|---|------------------|---------------------|----------|-------|-----------|----------|
| =====   |                  |                     |          |       |           |          |
| Dep. Variable:  | kriminal         | R-squared:          | 0.841    |       |           |          |
| Model:  | OLS              | Adj. R-squared:     | 0.827    |       |           |          |
| Method:   | Least Squares    | F-statistic:        | 60.90    |       |           |          |
| Date:   | Sat, 23 Dec 2023 | Prob (F-statistic): | 8.53e-18 |       |           |          |
| Time:   | 09:58:46         | Log-Likelihood:     | -335.50  |       |           |          |
| No. Observations:   | 51               | AIC:                | 681.0    |       |           |          |
| Df Residuals:   | 46               | BIC:                | 690.7    |       |           |          |
| Df Model:   | 4                |                     |          |       |           |          |
| Covariance Type:  | nonrobust        |                     |          |       |           |          |
| =====   |                  |                     |          |       |           |          |
|   | coef             | std err             | t        | P> t  | [0.025    | 0.975]   |
| -----   |                  |                     |          |       |           |          |
| const   | -2060.9260       | 659.487             | -3.125   | 0.003 | -3388.404 | -733.447 |
| metro   | 8.0019           | 1.294               | 6.183    | 0.000 | 5.397     | 10.607   |
| sma   | 4.6349           | 7.548               | 0.614    | 0.542 | -10.559   | 19.829   |
| miskin  | 22.9626          | 11.083              | 2.072    | 0.044 | 0.654     | 45.272   |
| single  | 128.3664         | 16.938              | 7.578    | 0.000 | 94.271    | 162.461  |
| =====   |                  |                     |          |       |           |          |
| Omnibus:  | 1.246            | Durbin-Watson:      | 1.652    |       |           |          |
| Prob(Omnibus):  | 0.536            | Jarque-Bera (JB):   | 0.524    |       |           |          |
| Skew:   | -0.109           | Prob(JB):           | 0.769    |       |           |          |
| Kurtosis:   | 3.446            | Cond. No.           | 2.68e+03 |       |           |          |
| =====   |                  |                     |          |       |           |          |
| Notes:  |                  |                     |          |       |           |          |
| [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.   |                  |                     |          |       |           |          |
| [2] The condition number is large, 2.68e+03. This might indicate that there are strong multicollinearity or other numerical problems. |                  |                     |          |       |           |          |
| Root Mean Squared Error (RMSE): 174.07058039152062  |                  |                     |          |       |           |          |

(Tabel 3 : Output OLS Regression)

Dari hasil *output python*, didapatkan persamaan *OLS Regression* dari output tersebut adalah :

- Kriminal = -2060.9260 + 8.0019 (Metro) + 4.6349 (SMA) + 22.9626 (Miskin) + 128.3664 (Single).

- b. Model *OLS Regression* memiliki *R-square* sebesar 0,841 yang menunjukkan bahwa model mampu menjelaskan sekitar 84,1% dari variansi dalam variabel target (kriminal).
- c. Model *OLS Regression* memiliki *Adjusted R-Square* sebesar 0,827 yang menunjukkan bahwa model mampu menjelaskan sekitar 82,7% dari variansi dalam variabel setelah mempertimbangkan jumlah variabel independen yang digunakan dalam model.
- d. Model *OLS Regression* memiliki RMSE sekitar 174 menunjukkan bahwa rata-rata kesalahan prediksi model ini sekitar 174 unit dari nilai aktual.

#### 4.2.2. *Robust Regression*

Berdasarkan exploratory data yang telah dilakukan, data terindikasi terdapat outlier, untuk mengatasi outlier tersebut dilakukan robust regression sebagai berikut

```

Coefficients: [ 0.          5.98413587 -9.59551515 16.59584867 153.50134155]
Intercept: -984.5609151591948
R^2 Score: 0.7908497997373205
Adjusted R^2 Score: 0.7993986704354338
Root Mean Squared Error (RMSE): 199.74063776865464

```

(Tabel 4 : *Output Robust Regression*)

Dari hasil *output python*, didapatkan persamaan *Robust Regression* dari output tersebut adalah:

- a. Kriminal = -984.5609 + 5.9841 (Metro) - 9.5955 (SMA) + 16.5958 (Miskin) + 153.5013 (*Single*).
- b. Model *Robust Regression* memiliki *R-square* sebesar 0,791 yang menunjukkan bahwa model mampu menjelaskan sekitar 79,1% dari variansi dalam variabel target (kriminal).
- c. Model *Robust Regression* memiliki *Adjusted R-Square* sebesar 0,799 yang menunjukkan bahwa model mampu menjelaskan sekitar 79,9% dari variansi dalam variabel setelah mempertimbangkan jumlah variabel independen yang digunakan dalam model.
- d. Model *Robust Regression* memiliki RMSE sekitar 200 menunjukkan bahwa rata-rata kesalahan prediksi model ini sekitar 200 unit dari nilai aktual.

### 4.3. Pengujian Hipotesis

#### 4.3.1. Uji Serentak (Uji F)

Hipotesis :

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

$$H_1 : \text{setidaknya ada satu } \beta_i \neq 0, i = 1, 2, 3, 4$$

$$\alpha : 0.05$$

Daerah kritis :  $p\text{-value} < \alpha$

Uji Statistik :

|   | F-hitung  | p-value      | alpha | Kesimpulan |
|---|-----------|--------------|-------|------------|
| 0 | 60.897183 | 8.529875e-18 | 0.05  | Tolak H0   |

(Tabel 5 : Output Uji Serentak)

Kesimpulan :

Berdasarkan perhitungan diatas, dapat kita lihat bahwa nilai  $p\text{-value}$  (8.529875e-18) lebih kecil dari nilai  $\alpha$  (0.05) yang artinya kita dapat simpulkan **Tolak H<sub>0</sub>**. Berarti terdapat hubungan antara variabel independen dengan variabel dependen.

#### 4.3.2. Uji Parsial (Uji t)

Hipotesis :

$$H_0 : \beta_i = 0, i = 1, 2, 3, 4$$

$$H_1 : \beta_i \neq 0, i = 1, 2, 3, 4$$

$$\alpha : 0.05$$

Daerah kritis :  $p\text{-value} < \alpha$

Uji Statistik :

|   | Variabel | t-hitung  | p-value      | alpha | Kesimpulan     |
|---|----------|-----------|--------------|-------|----------------|
| 0 | const    | -3.125044 | 3.076077e-03 | 0.05  | Tolak H0       |
| 1 | metro    | 6.183279  | 1.537333e-07 | 0.05  | Tolak H0       |
| 2 | sma      | 0.614020  | 5.422272e-01 | 0.05  | Gagal Tolak H0 |
| 3 | miskin   | 2.071864  | 4.391182e-02 | 0.05  | Tolak H0       |
| 4 | single   | 7.578474  | 1.245130e-09 | 0.05  | Tolak H0       |

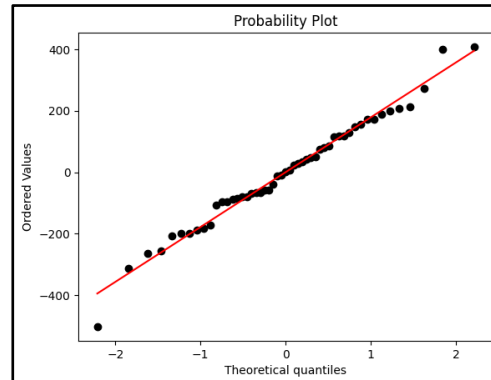
(Tabel 6 : Output Uji Parsial)

Kesimpulan :

- **Gagal Tolak  $H_0$**  : Untuk variabel SMA yang artinya variabel tersebut tidak berpengaruh terhadap tindak kriminal yang ada.
- **Tolak  $H_0$**  : Untuk variabel metro, miskin, dan *single* yang artinya variabel tersebut berpengaruh terhadap tindak kriminal yang ada.

#### 4.3.3. Uji Normalitas

a. Pengujian dengan grafik :



(Gambar 6 : Grafik Normalitas)

Sebaran data berada di sekitar garis diagonal dan mengikuti arah diagonal, maka dapat disimpulkan bahwa model regresi berdistribusi normal dan memenuhi asumsi normalitas.

b. Pengujian dengan statistik (Kolmogorov-Smirnov) :

Hipotesis :

$H_0$ : Error berdistribusi normal

$H_1$ : Error tidak berdistribusi normal

$\alpha$  : 0.05

Daerah kritis :  $p\text{-value} < \alpha$

```
Statistic KS: 0.07322782884154155
P-value: 0.928697393742446
Data Berdistribusi Normal (Gagal Tolak  $H_0$ )
```

(Tabel 7 : Output Uji Normalitas)

Dari *output python*, didapatkan p-value (0.928) > batas kritis penelitian (0.05), maka keputusan hipotesis **Gagal Tolak  $H_0$** , sehingga dapat dinyatakan bahwa error berdistribusi normal.

#### 4.3.4. Uji Autokorelasi

Hipotesis:

| Hipotesis Nol                               | Keputusan           | Jika                          |
|---|---------------------|-------------------------------|
| Tidak ada autokorelasi positif              | Tolak               | $0 < d < d_L$                 |
| Tidak ada autokorelasi positif              | Tidak ada keputusan | $d_L \leq d \leq d_U$         |
| Tidak ada autokorelasi negatif              | Tolak               | $4 - d_L < d < 4$             |
| Tidak ada autokorelasi negatif              | Tidak ada keputusan | $4 - d_U \leq d \leq 4 - d_L$ |
| Tidak ada autokorelasi positif atau negatif | Tidak ditolak       | $d_U < d < 4 - d_U$           |

(Tabel 7 : Daerah Kritis Uji Autokorelasi)

Statistik:

Berdasarkan tabel durbin watson(  $n = 51$ ,  $k = 4$ ):

$d_L = 1.3855$

$d_U = 1.7218$

Didapatkan output nilai  $d_W$  pada python:

```
Durbin-Watson statistic: 1.6516280405697554
Nilai dU dengan N = 51 dan K = 4 adalah 1.7218
Nilai dL dengan N = 51 dan K = 4 adalah 1.3855

Tidak ada keputusan. dL <= DW <= dU
(1.3855 <= 1.6516280405697554 <= 1.7218)
```

(Tabel 8 : Output Uji Autokorelasi)

Dari hasil statistik didapatkan  $d_L \leq d_W \leq d_U$  sehingga **tidak didapatkan keputusan** apakah terjadi korelasi positif atau tidak, tetapi dipastikan tidak terdapat autokorelasi negatif.

#### 4.3.5. Uji Heteroskedastisitas

|        | p-value  | Keputusan      |
|--------|----------|----------------|
| const  | 0.389627 | Gagal Tolak H0 |
| metro  | 0.584250 | Gagal Tolak H0 |
| sma    | 0.547906 | Gagal Tolak H0 |
| miskin | 0.150533 | Gagal Tolak H0 |
| single | 0.317166 | Gagal Tolak H0 |

(Tabel 9 : Output Uji Heteroskedastisitas)

Hipotesis :

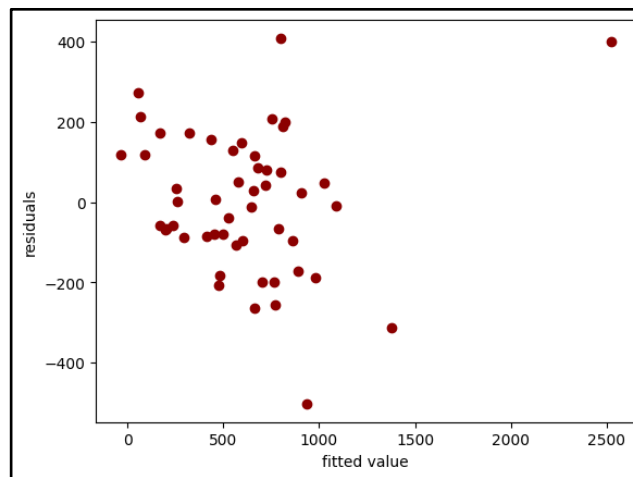
$H_0$  : Variansi Error Konstan (homoskedastisitas)

$H_1$  : Variansi Error Tidak Konstan (heteroskedastisitas)

$\alpha$  : 0.05

Daerah kritis :  $p\text{-value} < \alpha$

Berdasarkan output diatas, diketahui 4 variabel independen memiliki nilai p-value lebih besar dari nilai confidence level (0.05), maka dapat disimpulkan **Gagal Tolak H<sub>0</sub>**. Artinya, tidak terdapat masalah heteroskedastisitas.



(Gambar 7 :Grafik Uji Heteroskedastisitas)

Dengan melihat grafik scatterplot di atas, titik-titik yang terbentuk menyebar secara acak, baik di atas maupun di bawah angka 0 pada sumbu-Y. Maka dapat disimpulkan bahwa tidak terjadi heteroskedastisitas pada model regresi.

#### 4.3.6. Uji Multikolinearitas

|   | feature | VIF        | keputusan      |
|---|---------|------------|----------------|
| 0 | const   | 660.266956 | Tolak H0       |
| 1 | metro   | 1.201715   | Gagal Tolak H0 |
| 2 | sma     | 2.652013   | Gagal Tolak H0 |
| 3 | miskin  | 3.842047   | Gagal Tolak H0 |
| 4 | single  | 1.921894   | Gagal Tolak H0 |

(Tabel 10 : Output Uji Multikolinearitas)

Hipotesis :

$H_0$  : VIF < 10 (tidak terjadi multikolinearitas)

$H_1$  : VIF > 10 (terjadi multikolinearitas)

Sesuai dengan syarat Uji Multikolinearitas, jika nilai VIF kurang dari 10 maka dapat disimpulkan bahwa tidak terdapat masalah multikolinearitas. Nilai VIF yang rendah menunjukkan bahwa variabel-variabel independen dalam model regresi tidak memiliki korelasi yang tinggi satu sama lain, mengindikasikan ketidakberlebihan dalam pengaruh variabel-variabel tersebut terhadap variabel dependen (kriminal).

#### 4.4. Pemilihan Model Terbaik

Berdasarkan pembentukan model melalui metode *OLS regression* dan *robust regression* yang telah dilakukan, berikut tabel perbandingan antara kedua model :

| Pembanding                | <i>OLS Regression</i> | <i>Robust Regression</i> |
|---------------------------|-----------------------|--------------------------|
| Nilai RMSE                | 174                   | 200                      |
| <i>R-Squared</i>          | 0.841                 | 0.791                    |
| <i>Adjusted R-Squared</i> | 0.827                 | 0.791                    |

(Tabel 11 : Perbandingan Model)

Kriteria model terbaik yaitu:

1. Model memiliki nilai RMSE terkecil. Karena semakin kecil RMSE atau kesalahan prediksi maka semakin baik model.
2. Model memiliki *R-Squared* yang paling mendekati 1. Karena semakin *R-Squared* mendekati 1 maka semakin besar model dapat menjelaskan variansi dalam variabel dependen.
3. Model memiliki *Adjusted R-Squared* yang paling mendekati 1. Karena semakin *Adjusted R-Squared* mendekati 1 maka semakin besar model dapat menjelaskan variansi dalam variabel setelah mempertimbangkan jumlah variabel independen yang digunakan dalam model.

Berdasarkan ketiga kriteria tersebut, maka model terbaik untuk regresi tingkat kriminal dengan variabel independen metro, SMA, miskin, dan single adalah *OLS Regression* model dengan persamaan Kriminal =  $-2060.9260 + 8.0019 (\text{Metro}) + 4.6349 (\text{SMA}) + 22.9626 (\text{Miskin}) + 128.3664 (\text{Single})$ .



## BAB V

### KESIMPULAN DAN SARAN

#### 5.1. Kesimpulan

Pada uji asumsi klasik, dilakukan beberapa uji dan didapatkan hasil :

1. Uji Serentak, membuktikan bahwa terdapat hubungan antara variabel independen dengan variabel independen.
2. Uji Parsial, membuktikan bahwa hanya variabel sma yang tidak berpengaruh pada tindak kriminal, sedangkan variabel metro, miskin, dan *single* berpengaruh pada tindak kriminal.
3. Uji normalitas, membuktikan bahwa error berdistribusi normal.
4. Uji Autokorelasi, membuktikan bahwa dapat dipastikan tidak terdapat autokorelasi negatif tetapi tidak dapat diambil keputusan apakah terjadi autokorelasi positif.
5. Uji heteroskedastisitas, membuktikan bahwa tidak terdapat masalah heteroskedastisitas.
6. Uji multikolinearitas, membuktikan bahwa VIF kurang dari 10, sehingga tidak terdapat masalah multikolinearitas.

Selain *Ordinary Least Square Regression*, peneliti memilih untuk menganalisis data menggunakan *Robust Regression* karena berdasarkan exploratory data yang telah dilakukan, terindikasi bahwa data memiliki outlier.

Dari perbandingan RMSE, *R-Squared*, dan *Adjusted R-Squared* antara *OLS* dan *Robust Regression*, dapat disimpulkan bahwa model terbaik untuk regresi tingkat kriminal dengan variabel independen metro, SMA, miskin, dan single adalah *OLS Regression* model dengan persamaan Kriminal =  $-2060.9260 + 8.0019 (\text{Metro}) + 4.6349 (\text{SMA}) + 22.9626 (\text{Miskin}) + 128.3664 (\text{Single})$  dengan *R-Squared* sebesar 0.841

#### 5.2. Saran

Berdasarkan hasil penelitian, peneliti menyarankan :

1. Menambah data untuk validasi model
2. Evaluasi lebih lanjut terhadap outlier yang diidentifikasi dalam *Robust Regression* untuk memahami dampaknya

## DAFTAR PUSTAKA

- Anjari, Warih. (2014). Fenomena Kekerasan sebagai Bentuk Kekerasan (Violence). Jakarta. E-Journal WIDYA Yustisia. Vol I (1).
- Budiyono. (2017). Statistika untuk penelitian. UNS Press.
- Ghozali, Imam. 2018. Aplikasi Analisis Multivariat dengan Program SPSS. Semarang: Badan Penerbitan Universitas Diponegoro.
- Montgomery, D. C., Peck, E. A. 1992. Introduction to Linear Regression Analysis, 2nd edition. New York: John Wiley & Sons, Inc.
- Muthiah, Hanifa. (2016). Analisis Faktor yang Mempengaruhi Jumlah Kejahatan di Indonesia dengan Pendekatan Regresi Spasial. Yogyakarta.
- Purwanti, E. Y., & Widyaningsih, E. (2019). Analisis Faktor Ekonomi yang Mempengaruhi Kriminalitas di Jawa Timur. Jurnal Ekonomi-Qu, 9(2), 157. Semarang.
- Rawlings, J. O., Pantula, S. G., & Dickey, D. A. (1998). *Applied Regression Analysis: A Research Tool* (2nd ed.). Springer New York.

### *Lampiran*

Lampiran 1. Perhitungan menggunakan python: <https://bit.ly/PythonFinalProjectRegter>