

**APLIKASI REGRESI *ROBUST* DATA PANEL
DENGAN PENDUGA *LEAST TRIMMED SQUARE* (LTS)
PADA DATA GARIS KEMISKINAN DI PROVINSI LAMPUNG**

(Skripsi)

Oleh

WINDI LESTARI



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2025**

ABSTRACT

ROBUST PANEL DATA REGRESSION ANALYSIS USING THE LEAST TRIMMED SQUARE (LTS) ESTIMATOR ON POVERTY LINE DATA IN LAMPUNG PROVINCE

BY

WINDI LESTARI

Robust regression is an alternative method in regression analysis designed to produce stable parameter estimates, even when the data contain outliers or deviate from classical assumptions. One of its estimation techniques, the Least Trimmed Square (LTS), works by minimizing the smallest squared residuals, thereby assigning smaller weights to extreme data points. This method serves as a solution when classical approaches such as the Ordinary Least Squares (OLS) fail to meet the assumptions, especially in socio-economic data that are often complex and prone to outliers. This study applies robust regression with the LTS estimator on panel data to analyze the influence of population size (JP), population density (KP), and registered job vacancies (LK) on poverty lines in Lampung Province. The data cover 15 districts/cities during the period 2019–2023. The analysis results show that the model obtained has a coefficient of determination of $R^2 = 0.8909$, which means that 89.09% of the variation in the poverty line can be explained by the three predictor variables.

Keywords: Poverty Line, Robust Regression, Least Trimmed Square (LTS), Panel Data

ABSTRAK

APLIKASI REGRESI *ROBUST* DATA PANEL DENGAN PENDUGA *LEAST TRIMMED SQUARE* (LTS) PADA DATA GARIS KEMISKINAN DI PROVINSI LAMPUNG

OLEH

WINDI LESTARI

Regresi *robust* merupakan metode alternatif dalam analisis regresi yang dirancang untuk menghasilkan estimasi parameter yang tetap stabil meskipun data mengandung pencilan atau penyimpangan dari asumsi klasik. Salah satu metode estimasinya, yaitu *Least Trimmed Square* (LTS), bekerja dengan meminimalkan jumlah kuadrat galat terkecil, sehingga memberikan bobot yang lebih kecil pada data ekstrem. Metode ini menjadi solusi ketika pendekatan klasik seperti Metode Kuadrat Terkecil (MKT) tidak lagi memenuhi asumsi, terutama pada data sosio-ekonomi yang seringkali kompleks dan rentan terhadap pencilan. Penelitian ini mengaplikasikan regresi *robust* dengan penduga LTS pada data panel untuk menganalisis pengaruh jumlah penduduk (JP), kepadatan penduduk (KP), dan lowongan kerja terdaftar (LK) terhadap garis kemiskinan di Provinsi Lampung. Data yang digunakan mencakup 15 Kabupaten/kota selama periode 2019–2023. Hasil analisis menunjukkan bahwa model yang diperoleh memiliki nilai koefisien determinasi $R^2 = 0.8909$, yang berarti 89,09% variasi garis kemiskinan dapat dijelaskan oleh ketiga variabel prediktor tersebut.

Kata Kunci: Garis Kemiskinan, Regresi *Robust*, *Least Trimmed Square* (LTS), Data Panel

**APLIKASI REGRESI *ROBUST* DATA PANEL
DENGAN PENDUGA *LEAST TRIMMED SQUARE* (LTS)
PADA DATA GARIS KEMISKINAN DI PROVINSI LAMPUNG**

Oleh

Windi Lestari

Skripsi

**Sebagai Salah Satu Syarat Untuk Mencapai Gelar
SARJANA MATEMATIKA**

Pada

**Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2025**

Judul Skripsi : **APLIKASI REGRESI *ROBUST* DATA PANEL
DENGAN PENDUGA *LEAST TRIMMED*
SQUARE (LTS) PADA DATA GARIS
KEMISKINAN DI PROVINSI LAMPUNG**

Nama Mahasiswa : **Windi Lestari**

Nomor Pokok Mahasiswa : **2117031045**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI

1. Komisi Pembimbing

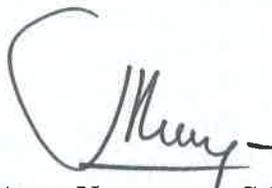


Widiarti, S.Si., M.Si.
NIP. 198005022005012003



Bernadhita Herindri S. U., M.Sc.
NIP. 199206302023212034

2. Ketua Jurusan Matematika

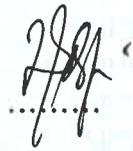


Dr. Aang Nuryaman, S.Si., M.Si.
NIP. 197403162005011001

MENGESAHKAN

1. Tim Penguji

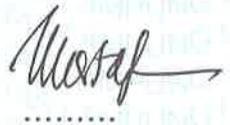
Ketua : Widiarti, S.Si., M.Si.



Sekretaris : Bernadhita Herindri S. U., M.Sc.



**Penguji
Bukan Pembimbing : Prof. Dr. Mustofa Usman, M.A., Ph.D.**



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Heri Satria, S.Si., M.Si.
NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi : 24 April 2025

PERNYATAAN SKRIPSI MAHASISWA

Saya yang bertanda tangan dibawah ini :

Nama Mahasiswa : Windi Lestari
Nomor Pokok Mahasiswa : 2117031045
Jurusan : Matematika
Fakultas : Matematika dan Ilmu Pengetahuan Alam
Judul Skripsi : Aplikasi Regresi *Robust* Data Panel Dengan
Penduga *Least Trimmed Square* (LTS) Pada Data
Garis Kemiskinan Di Provinsi Lampung

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri. Apabila dikemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, Mei 2025



Windi Lestari
NPM 2117031045

RIWAYAT HIDUP

Penulis memiliki nama lengkap Windi Lestari dan lahir di Sukaraja pada tanggal 03 Agustus 2003. Penulis merupakan putri tunggal dari pasangan Bapak Dakun dan Ibu Pujiyani.

Penulis mengawali pendidikan di Taman Kanak-Kanak (TK) Primayudha pada tahun 2007-2009 dan menempuh pendidikan dasar di SDN 2 Karangrejo pada tahun 2009-2015. Kemudian penulis melanjutkan pendidikan di SMPN 2 Semaka pada tahun 2015-2018 dan Sekolah Menengah Atas di SMAN 1 Semaka pada tahun 2018-2021. Setelah itu pada tahun 2021, penulis melanjutkan pendidikan Strata Satu di Jurusan Matematika FMIPA Universitas Lampung melalui jalur SBMPTN.

Selama menjadi mahasiswi, penulis aktif di beberapa kegiatan di antaranya aktif dalam kepengurusan organisasi ROIS FMIPA Unila sebagai anggota Bidang Akademik & Riset pada tahun 2022 dan 2023. Penulis juga menjadi salah satu anggota UKM ESo Unila di cabang Scrabble pada tahun 2022. Kemudian pada periode 26 Desember 2023 - 3 Februari 2024, penulis melakukan Praktik Kerja Lapangan (PKL) di Badan Pusat Statistik Kabupaten Tanggamus. Dan Pada periode 25 Juni – 3 Agustus 2024, penulis melakukan Kuliah Kerja Nyata (KKN) sebagai bentuk pengabdian kepada masyarakat dan juga berkontribusi dalam pembangunan Desa Penyangga Taman Nasional Way Kambas di Desa Braja Luhur, Kecamatan Braja Selehah, Kabupaten Lampung Timur, Provinsi Lampung.

KATA INSPIRASI

“Maka sesungguhnya bersama kesulitan pasti ada kemudahan. Maka apabila engkau telah selesai (dari suatu urusan), tetaplah bekerja keras (untuk urusan yang lain). Dan hanya kepada Tuhan m ulah engkau berharap”

(QS. Al-Insyirah : 6-8)

“Allah tidak akan membebani seseorang melainkan sesuai dengan kesanggupannya....”

(QS. Al-Baqarah : 286)

“Sesungguhnya Allah akan mengangkat derajat seseorang dengan ilmu, dan menjadikan ilmu itu sebagai cahaya bagi yang mencarinya.”

(Imam Syafi’i)

“Bisikkanlah Terima kasih pada diri sendiri, hebat dia terus menjaga mu dan sayangimu”

(Tulus - Diri)

“Long story short, i survived”

(Taylor Swift)

PERSEMBAHAN

Alhamdulillahirobbil' alamin, segala puji syukur kepada Allah SWT atas nikmat dan karunia-nya sehingga skripsi ini dapat terselesaikan dengan baik. Dengan kerendahan hati kupersembahkan karya kecilku untuk:

Kedua orang tuaku Bapak Dakun dan Ibu Pujiyani:

Bapak dan Ibu tersayang yang selalu mendoakan untuk kebaikan anaknya, selalu memberi kasih sayang, cinta, dan dukungan, serta motivasi. Terima kasih banyak atas segala pengorbanan dan doa tulus kasih yang di berikan. Mereka memang tidak sempat merasakan pendidikan di bangku perkuliahan, namun mereka mampu dan senantiasa memberikan yang terbaik untuk anaknya. Semoga Bapak dan Ibu sehat, panjang umur dan bahagia selalu.

Bapak Ibu Dosen Pembimbing dan Pembahas:

Terima kasih kepada Bapak/Ibu Dosen Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi selama proses penelitian ini. Terima kasih atas kesabaran, ilmu, dan waktu yang telah Bapak/Ibu berikan.

Keluarga Besar dan Sahabat-Sahabatku Tercinta:

Terima kasih banyak untuk keluarga besar dan sahabat tercinta, yang selalu memberikan dukungan, cinta, dan semangat tanpa henti. Kalian adalah sumber kekuatan saya dalam setiap langkah perjalanan ini.

Almamater Tercinta, Universitas Lampung

SANWACANA

Alhamdulillah, segala puji dan syukur penulis panjatkan ke hadirat Allah SWT karena atas rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “Aplikasi Regresi *Robust* Data Panel Dengan Penduga *Least Trimmed Square* (LTS) Pada Data Garis Kemiskinan di Provinsi Lampung”. Shalawat beserta salam senantiasa tercurah kepada baginda Nabi Muhammad SAW, suri tauladan untuk kita semua, semoga dikemudian hari mendapat syafaat dari beliau.

Dalam penyusunan skripsi ini penulis banyak mendapatkan bantuan dan dukungan dari berbagai pihak. Oleh karena itu penulis ingin mengungkapkan rasa Terima kasih kepada:

1. Ibu Widiarti, S.Si., M.Si., selaku Dosen Pembimbing I atas kesediaannya meluangkan waktu untuk membimbing, memberikan kritik, saran, serta arahan selama proses penyelesaian studi dan skripsi ini.
2. Ibu Dr. Bernadhita Herindri Samodera Utami, M.Sc. selaku Dosen Pembimbing II atas kesediaannya untuk membimbing serta memotivasi penulis dalam menyelesaikan skripsi ini.
3. Bapak Prof. Dr. Mustofa Usman, M.A., Ph.D. selaku Dosen Pembahas yang telah memberikan arahan serta kritik dan saran yang membantu dalam memperbaiki skripsi ini.
4. Bapak Dr. Aang Nuryaman, S.Si., M.Si. selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
5. Ibu Dra. Dorrah Azis, M.Si. selaku Dosen Pembimbing Akademik atas kesediaannya memberikan arahan serta dukungan selama penyelesaian studi ini.

6. Bapak Dr. Eng. Heri Satria, S.Si., M.Si. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Seluruh dosen dan staf Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Cinta pertama dan panutan penulis, Bapak Dakun dan pintu surga penulis, Ibu Pujiyani. yang sangat penulis cintai dan menjadi sumber semangat, doa, dan kasih sayang tanpa batas yang selalu mendukung dan menguatkan penulis.
9. Sahabat-sahabat penulis (Silvia, Anggun, Dzita, Leni, Arya, Cici, Fahmi) dan saudara penulis (Winda) terima kasih atas dukungan, semangat, dan pelajaran berharga yang telah diberikan.
10. Syarli Dita Anjani & Andi Wahyudiansyah, dua orang luar biasa yang selalu hadir dalam tiap fase perjuangan ini.
11. Rohana, Irma, Amanda, Buena, Damai, Sayida, Dede, Lusiana, Afika, Teman-teman KKN keluarga berencana (Margeliza, Khairi, Bila, dan Hafidz), Teman-teman seperbimbingan (Syarli, Irma, Buena, Dede, Siska, Shelvira, Rosa) yang telah memberi warna kehidupan dan pengalaman paling berharga, menemani dan kebersamai suka dan duka selama perkuliahan.
12. Seluruh pihak terkait yang telah membantu menyelesaikan skripsi ini yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa masih banyak terdapat kekurangan dalam penulisan skripsi ini. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun dari semua pihak.

Bandar Lampung, 24 April 2025
Penulis,

Windi Lestari
NPM. 2117031045

DAFTAR ISI

DAFTAR TABEL	ii
DAFTAR GAMBAR	v
I. PENDAHULUAN	1
1.1 Latar Belakang dan Masalah	1
1.2 Tujuan Penelitian	4
1.3 Manfaat Penelitian	4
II. TINJAUAN PUSTAKA	5
2.1 Analisis Regresi	5
2.2 Analisis Regresi Data Panel	6
2.3 Estimasi Model Regresi Data Panel	8
2.3.1 <i>Common Effect Model</i> (CEM)	8
2.3.2 <i>Fixed Effect Model</i> (FEM)	9
2.3.3 <i>Random Effect Model</i> (REM)	9
2.4 Pemilihan Model Regresi Data Panel	10
2.4.1 Uji <i>Chow</i>	11
2.4.2 Uji <i>Hausman</i>	12
2.5 Uji Asumsi Model	13
2.5.1 Uji Normalitas	13
2.5.2 Uji Multikolinearitas	14
2.5.3 Uji Heteroskedastisitas	15
2.5.4 Uji Autokorelasi	15
2.6 Pendeteksian Pencilan	16
2.6.1 Nilai Pengaruh (<i>Leverage Value</i>)	16
2.6.2 TRES (<i>Studentized Deleted Residual</i>)	17

2.7 Regresi <i>Robust</i>	18
2.8 Estimasi <i>Least Trimmed Square</i> (LTS).....	19
2.9 Pemilihan Model Terbaik Menggunakan Nilai R^2	21
2.10Garis Kemiskinan	21
III. METODE PENELITIAN.....	23
3.1 Waktu dan Tempat Penelitian	23
3.2 Data Penelitian.....	23
3.3 Metode Penelitian	24
IV. HASIL DAN PEMBAHASAN.....	27
4.1 Analisis Deskriptif.....	27
4.2 Uji Multikolinearitas	30
4.3 Pemodelan Regresi Data Panel.....	31
4.3.1 <i>Common Effect Model</i> (CEM)	31
4.3.2 <i>Fixed Effect Model</i> (FEM)	32
4.3.3 <i>Random Effect Model</i> (REM)	34
4.4 Pemilihan Model Regresi Data Panel.....	34
4.4.1 Uji <i>Chow</i>	35
4.4.2 Uji <i>Hausman</i>	36
4.5 Uji Asumsi Model.....	36
4.5.1 Uji Normalitas.....	37
4.5.2 Uji Heteroskedastisitas.....	38
4.5.3 Uji Autokorelasi	39
4.6 Pendeteksian Pencilan	39
4.7 Regresi <i>Robust Least Trimmed Square</i> (LTS)	42
V. KESIMPULAN.....	44
DAFTAR PUSTAKA	45

DAFTAR TABEL

Tabel 1. Analisis Deskriptif Tabel Penelitian	29
Tabel 2. <i>Output</i> Uji Multikolinearitas	30
Tabel 3. <i>Output</i> Estimasi CEM	31
Tabel 4. <i>Output</i> Estimasi FEM.....	32
Tabel 5. <i>Output</i> Efek Individu.....	33
Tabel 6. <i>Output</i> Estimasi REM	34
Tabel 7. <i>Output</i> Uji <i>Chow</i>	35
Tabel 8. <i>Output</i> Uji <i>Hausman</i>	36
Tabel 9. <i>Output</i> Uji Normalitas	37
Tabel 10. <i>Output</i> Uji Heteroskedastisitas	38
Tabel 11. <i>Output</i> Uji Autokorelasi	39
Tabel 12. Hasil Identifikasi Pengamatan Berpengaruh Dengan <i>Leverage Value</i> ..	40
Tabel 13. Hasil Identifikasi Pencilan dan Pengamatan Berpengaruh TRES.....	41
Tabel 14. Hasil Pendugaan Parameter LTS	42
Tabel 15. Hasil Pengujian Simultan	43
Tabel 16. Hasil Pengujian Parsial.....	43

DAFTAR GAMBAR

Gambar 1. Diagram Alur Penelitian.....	26
Gambar 2. Grafik Garis Kemiskinan di Povinsi Lampung 2019-2023.....	28

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis regresi adalah salah satu metode dalam statistika yang digunakan untuk mempelajari pola serta mengukur hubungan statistik antara dua atau lebih variabel yang memiliki hubungan sebab-akibat (Montgomery dkk., 2012). Salah satu tujuan analisis regresi adalah menentukan nilai estimasi dari koefisien regresi dalam suatu model regresi. Metode Kuadrat Terkecil (MKT) merupakan pendekatan yang paling umum digunakan untuk menduga parameter dalam model regresi linier berganda. Namun, penerapan MKT memerlukan pemenuhan beberapa asumsi regresi, termasuk asumsi normalitas residual. Uji normalitas residual sering kali tidak dapat terpenuhi karena adanya pencilan pada data (Rahman & Wachidah, 2021; Chen, 2002). Pencilan merupakan amatan yang memiliki pola berbeda dengan kebanyakannya. Penyebab utama adanya pencilan antara lain kesalahan saat memasukkan data pada komputer, kesalahan saat mengambil contoh di lapangan, kejadian unik yang sesuai fakta di lapangan, dan sebagainya.

Data panel adalah jenis data yang terdiri dari pengamatan terhadap unit individu atau *cross-section* yang sama dalam beberapa periode waktu (Gujarati & Porter, 2009). Menurut Baltagi (2008), data panel memiliki sejumlah keunggulan, antara lain memberikan informasi yang lebih kaya, mengurangi kolinearitas antar variabel, dan meningkatkan efisiensi analisis. Data panel juga memungkinkan peneliti untuk menganalisis dinamika perubahan dalam variabel dari waktu ke waktu, sehingga

memberikan wawasan yang lebih mendalam tentang hubungan antar variabel. Melalui pendekatan ini, peneliti dapat memanfaatkan model regresi data panel untuk menganalisis pengaruh satu atau lebih variabel prediktor terhadap variabel respon secara lebih efektif.

Meskipun data panel menawarkan banyak keunggulan, keberadaan pencilan dalam data dapat menimbulkan permasalahan serius dalam analisis. Pencilan ini dapat mempengaruhi hasil estimasi dan interpretasi model, sehingga perlu ditangani dengan hati-hati. Menghapus pencilan dari data bukanlah langkah yang tepat, karena pencilan dapat mengandung informasi berharga yang tidak ditemukan dari pengamatan lainnya. Oleh karena itu, diperlukan metode estimasi parameter yang dapat menangani serta tetap andal terhadap keberadaan pencilan pada data, seperti regresi *robust* (Kutner dkk., 2004).

Regresi *robust* adalah metode analisis yang dirancang untuk memberikan estimasi yang lebih stabil dan akurat ketika data yang dianalisis mengandung pencilan atau ketika distribusi residual tidak normal. Metode ini sangat penting dalam situasi di mana pengamatan ekstrem dapat mempengaruhi hasil analisis secara signifikan. Regresi *robust* memungkinkan estimasi parameter tidak hanya bergantung pada seluruh data, tetapi juga memberikan bobot lebih kecil kepada pengamatan yang dianggap sebagai pencilan. Dengan demikian, model yang dihasilkan menjadi lebih tahan terhadap pengaruh data ekstrem. Beberapa metode estimasi dalam regresi *robust* meliputi Estimasi *Least Trimmed Square* (LTS), Estimasi *Scale* (S), dan Estimasi *Method Of Moment* (MM) (Chen, 2002; Rousseeuw & Leroy, 1987).

Beberapa penelitian sebelumnya telah membuktikan efektivitas menggunakan regresi *robust* dalam pemodelan data yang kompleks. Setyowati dkk. (2021) menemukan bahwa LTS memberikan estimasi stabil pada data produksi padi, sedangkan Akolo dan Nadjamuddin (2022) menunjukkan efektivitas LTS dalam pemodelan tekanan darah sistolik. Rahman & Wachidah (2021) menunjukkan bahwa metode LTS memiliki *residual standard error* yang lebih rendah serta

adjusted R-square yang lebih tinggi dibandingkan dengan metode S dan MM. Selain itu, Anggun dkk. (2021) menemukan bahwa regresi *ridge robust S-Estimation* menghasilkan estimasi yang lebih efisien dalam penelitian mengenai faktor apa saja yang mempengaruhi kemiskinan di Jawa Tengah. Penelitian-penelitian ini secara konsisten memperkuat efektivitas pendekatan *robust* dalam mencapai model yang lebih stabil dan akurat terutama dalam berbagai kondisi data yang mengandung pencilan.

Berdasarkan penelitian sebelumnya, regresi *robust* dengan penduga LTS, dinilai sangat sesuai untuk menganalisis data terkait garis kemiskinan yang kompleks dan rentan terhadap pencilan. Penelitian ini menggunakan data yang mencakup variabel-variabel penting seperti garis kemiskinan (GK), jumlah penduduk (JP), kepadatan penduduk (KP), dan lowongan kerja terdaftar (LK). Garis Kemiskinan mengukur batas minimal pendapatan yang diperlukan untuk memenuhi kebutuhan dasar (Sukirno, 2019). Kepadatan penduduk (KP) yang tinggi sering kali berkorelasi dengan peningkatan persaingan untuk pekerjaan dan sumber daya, sehingga berpotensi meningkatkan tingkat kemiskinan, sedangkan jumlah penduduk (JP) yang tinggi sering kali memicu meningkatnya kebutuhan akan sumber daya dan layanan dasar, seperti pendidikan, kesehatan, serta perumahan. Peningkatan jumlah penduduk tanpa diiringi dengan pertumbuhan ekonomi yang memadai dapat menyebabkan terbatasnya lapangan pekerjaan, penurunan pendapatan per kapita, dan pada akhirnya meningkatkan risiko kemiskinan. Selain itu, lowongan kerja terdaftar (LK) menunjukkan ketersediaan pekerjaan di suatu daerah, semakin banyak lowongan yang tersedia, Semakin tinggi peluang penduduk untuk memperoleh pekerjaan, semakin besar kesempatan mereka dalam meningkatkan taraf hidup.

Penggunaan regresi *robust* pada penelitian ini bertujuan untuk memahami hubungan antara variabel-variabel tersebut dan garis kemiskinan secara lebih akurat, meskipun terdapat pencilan dalam data. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam memahami lebih lanjut tentang

faktor apa saja yang mempengaruhi garis kemiskinan, serta memberikan alternatif metodologi yang lebih *robust* dalam analisis data sosio-ekonomi.

1.2 Tujuan Penelitian

Penelitian ini bertujuan untuk menganalisis pengaruh jumlah penduduk (JP), kepadatan penduduk (KP), dan lowongan kerja terdaftar (LK) terhadap garis kemiskinan di Provinsi Lampung. Selain itu, penelitian ini mengaplikasikan model regresi *robust* dengan penduga *Least Trimmed Square* (LTS) untuk memahami hubungan antar variabel serta mengatasi pencilan dalam data.

1.3 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah:

1. memberikan pemahaman yang lebih mendalam mengenai faktor-faktor yang memengaruhi garis kemiskinan sehingga dapat dijadikan referensi bagi pengambil kebijakan dalam merumuskan strategi pengentasan kemiskinan;
2. menambah pengalaman dan pengetahuan baru tentang penggunaan metode regresi *robust* dalam penelitian sosial ekonomi;
3. menghasilkan model regresi *robust* yang dapat diandalkan untuk analisis lebih lanjut dalam studi-studi serupa di masa depan.

II. TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi adalah teknik statistik yang digunakan untuk mengevaluasi hubungan antara variabel dependen dengan satu atau lebih variabel independen. Tujuan utama dari analisis ini adalah memahami serta mengukur sejauh mana perubahan pada suatu variabel dapat memprediksi perubahan pada variabel lainnya. Selain memodelkan dan menganalisis hubungan antar variabel, analisis regresi juga berperan sebagai alat bantu dalam proses prediksi dan pengambilan keputusan (Montgomery dkk., 2012). Batalgi (2008) menambahkan bahwa pemilihan model regresi yang tepat menjadi kunci untuk memperoleh estimasi yang akurat dan valid, terutama ketika menghadapi data dengan karakteristik khusus, seperti adanya heteroskedastisitas atau multikolinearitas.

Menurut Firdaus (2004), analisis regresi linear dibagi menjadi 2 kategori utama yaitu sebagai berikut:

1. Analisis regresi linear sederhana merupakan model yang digunakan untuk mengevaluasi hubungan antara satu variabel independen (bebas) dan satu variabel dependen (tak bebas). Model regresi linear sederhana adalah:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i ; i = 1, 2, \dots, n \quad (2.1)$$

Keterangan:

- Y_i : variabel dependen (terikat)
 X_i : variabel independen (bebas)
 β_0 : parameter intersep/konstanta regresi
 β_1 : parameter koefisien regresi
 ε_i : variabel galat
 n : banyaknya data penelitian

2. Analisis regresi linear berganda adalah model regresi yang melibatkan lebih dari satu variabel bebas (independen) dalam analisis hubungan dengan variabel tak bebas (dependen). Model regresi linear berganda:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + \varepsilon_i ; i = 1, 2, \dots, n \quad (2.2)$$

Keterangan:

- Y_i : variabel dependen (terikat)
 $X_1, X_2, X_3, \dots, X_k$: variabel independen (bebas)
 β_0 : parameter intersep/konstanta regresi
 $\beta_1, \beta_2, \beta_3, \dots, \beta_k$: parameter koefisien regresi
 ε_i : variabel galat
 k : banyaknya variabel bebas/faktor
 n : banyaknya data penelitian

2.2 Analisis Regresi Data Panel

Analisis regresi data panel merupakan metode statistik yang mengkombinasikan data *time series* (observasi sepanjang waktu) dan *cross-section* (observasi pada satu titik waktu) untuk menganalisis hubungan antara variabel dependen dan independen. Menurut Montgomery dkk. (2012), data panel dapat memungkinkan peneliti untuk

mengeksplorasi dinamika antar variabel dalam konteks waktu dan entitas yang berbeda, sehingga memberikan informasi yang lebih kaya dibandingkan dengan analisis regresi yang hanya menggunakan satu jenis data. Data panel juga membantu dalam mengatasi masalah multikolinearitas dan heteroskedastisitas yang sering muncul dalam analisis regresi tradisional, karena dapat memanfaatkan variasi antar individu dan waktu secara bersamaan (Gujarati & Porter, 2009).

Menurut Setiawan & Kusrini (2010), data panel memiliki beberapa keunggulan dibandingkan data *time series* dan *cross section*, yaitu:

1. data panel menawarkan informasi yang lebih kaya dan bervariasi, dengan korelasi antar variabel yang lebih rendah, derajat kebebasan yang lebih tinggi, serta efisiensi estimasi yang lebih baik;
2. data panel lebih sesuai untuk menganalisis perubahan dinamis, seperti studi mengenai pengangguran atau mobilitas tenaga kerja;
3. data panel juga berguna dalam mengamati pola perilaku, misalnya untuk memahami dinamika perubahan skala ekonomi dan perkembangan teknologi.

Menurut Baltagi (2008), secara umum model regresi data panel dapat dinyatakan sebagai berikut :

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \varepsilon_{it} \quad (2.3)$$

Keterangan:

Y_{it} : variabel terikat untuk unit individu ke- i pada waktu ke- t

X_{it} : variabel bebas untuk unit individu ke- i pada waktu ke- t

β_0 : parameter intersep atau konstanta regresi

β_1 : parameter koefisien *slope* yang berlaku untuk semua unit

ε_{it} : variabel galat untuk unit individu ke- i pada waktu ke- t

2.3 Estimasi Model Regresi Data Panel

Model analisis regresi data panel terdiri dari 3 jenis, yaitu *Common Effect Model* (CEM), *Fixed Effect Model* (FEM), dan *Random Effect Model* (REM). Ketiga model ini diklasifikasikan berdasarkan asumsi mengenai pengaruh yang digunakan pada analisis regresi data panel (Gujarati & Porter, 2009).

2.3.1 *Common Effect Model* (CEM)

Common Effect Model (CEM) merupakan metode paling dasar dalam analisis regresi data panel, dimana seluruh data dianalisis tanpa memperhitungkan variasi antar individu maupun perbedaan waktu. Model ini mengasumsikan bahwa semua pengamatan, baik dari variabel dependen maupun independen, berasal dari satu populasi yang sama sehingga hanya menghasilkan satu set parameter yang mencakup intersep dan koefisien regresi (Setiawan & Kusriani, 2010). Model *Common Effect Model* dapat dinyatakan sebagai berikut:

$$Y_{it} = \beta_0 + \sum_{k=1}^n \beta_k X_{kit} + \varepsilon_{it} \quad (2.4)$$

Keterangan:

Y_{it} : variabel terikat untuk unit individu ke- i pada waktu ke- t

X_{it} : variabel bebas untuk unit individu ke- i pada waktu ke- t

β_0 : parameter intersep atau konstanta regresi

β_1 : parameter koefisien *slope* yang berlaku untuk semua unit

ε_{it} : variabel galat untuk unit individu ke- i pada waktu ke- t

2.3.2 Fixed Effect Model (FEM)

Fixed Effect Model (FEM) merupakan metode regresi dalam analisis data panel yang digunakan untuk mengestimasi hubungan antar variabel dengan memperhitungkan perbedaan efek di setiap individu. Model ini memasukkan *dummy variables* untuk setiap individu, yang memungkinkan perbedaan intersep di antara mereka. Dengan cara ini, FEM dapat mengakomodasi variasi yang tidak teramati yang mungkin mempengaruhi variabel dependen tetapi tetap konstan sepanjang waktu untuk setiap individu (Gujarati & Porter, 2009).

Pada *Fixed Effect Model*, setiap individu diasumsikan mempunyai perbedaan parameter intersep. Secara umum, model ini dapat ditulis dalam bentuk persamaan sebagai berikut:

$$Y_{it} = \beta_{0it} + \sum_{k=1}^n \beta_k X_{kit} + \varepsilon_{it} \quad (2.5)$$

Keterangan:

Y_{it} : variabel dependen untuk unit individu ke- i pada waktu ke- t

X_{it} : variabel independen untuk unit individu ke- i pada waktu ke- t

β_{0it} : parameter intersep atau konstanta regresi

β_k : parameter koefisien *slope* berlaku untuk semua unit

ε_{it} : variabel galat untuk unit individu ke- i pada waktu ke- t

2.3.3 Random Effect Model (REM)

Fixed Effect Model (FEM) memiliki kelemahan dalam hal derajat kebebasan yang lebih rendah, karena penggunaan variabel *dummy* mengurangi jumlah observasi efektif (Winarno, 2007). Sebagai alternatif, *Random Effect Model* (REM)

mempertahankan derajat kebebasan lebih tinggi dan memungkinkan analisis terhadap variabel konstan sepanjang waktu (Baltagi, 2008). *Random Effect Model* berguna jika efek acak tidak berkorelasi dengan variabel independen, memberikan estimasi parameter yang lebih efisien daripada FEM. Jika asumsi ini tidak terpenuhi, estimasi REM bisa bias dan tidak konsisten (Widarjono, 2009). Pemilihan FEM dan REM umumnya dilakukan menggunakan uji *Hausman*, yang bertujuan menentukan model yang paling tepat sesuai dengan karakteristik data.

Random Effect Model mengasumsikan bahwa *error term* mungkin berhubungan antar waktu atau antar individu, memungkinkan variasi acak yang tidak teramati di antara individu. Pada model ini, efek individu dianggap sebagai komponen acak yang ditambahkan ke *error term*, dengan bentuk persamaan:

$$Y_{it} = \beta_{0it} + \sum_{k=1}^n \beta_k X_{kit} + \varepsilon_{it} + u_i \quad (2.6)$$

Keterangan:

Y_{it} : variabel terikat (dependen) untuk unit individu ke- i pada waktu ke- t

X_{it} : variabel bebas (independen) untuk unit individu ke- i pada waktu ke- t

β_{0it} : parameter intersep atau konstanta regresi

β_k : parameter koefisien *slope* berlaku untuk semua unit

ε_{it} : variabel galat untuk unit individu ke- i pada waktu ke- t

u_i : kombinasi variabel galat untuk unit individu ke- i pada waktu ke- t

2.4 Pemilihan Model Regresi Data Panel

Pemilihan model regresi data panel bertujuan untuk menentukan model yang paling sesuai di antara tiga jenis regresi, yaitu *Common Effect Model*, *Fixed Effect Model*, dan *Random Effect Model*. Untuk memilih model regresi data panel terbaik, diperlukan serangkaian pengujian sebagai berikut:

2.4.1 Uji Chow

Uji *Chow* merupakan metode yang diterapkan untuk membandingkan *Common Effect Model* dengan *Fixed Effect Model* (Sriyana, 2015). Hipotesis yang digunakan dalam uji ini sebagai berikut:

$H_0: \beta_1 = \beta_2 = \dots = \beta_n = 0$ (model yang sesuai adalah CEM)

H_1 : terdapat setidaknya satu $\beta_i \neq 0$, dengan $i = 1, 2, \dots, n$ (model yang sesuai adalah FEM)

Statistik uji:

$$F = \frac{\frac{RSS_1 - RSS_2}{(n-1)}}{\frac{RSS_2}{nT - n - K}} \quad (2.7)$$

Keterangan

RSS_1 : jumlah kuadrat residu dari model CEM

RSS_2 : jumlah kuadrat residu dari model FEM

n : jumlah individu dalam data *Cross Section*

T : jumlah periode waktu pada data *Time Series*

K : jumlah variabel independen

Jika nilai $Chow > F_{n-1, nT-n-K}$ atau $P\ value < \alpha$ maka H_0 ditolak, artinya *Fixed Effect Model* adalah yang lebih sesuai. Sebaliknya jika nilai $Chow < F_{n-1, nT-n-K}$ atau $P\ value > \alpha$ maka tidak ada cukup bukti untuk menolak H_0 , artinya *Common Effect Model* dianggap lebih tepat.

2.4.2 Uji Hausman

Uji *Hausman* merupakan salah satu metode yang diterapkan untuk melakukan pemilihan antara *Fixed Effect Model* (FEM) dan *Random Effect Model* (REM) dalam analisis regresi data panel. Menurut Gujarati & Porter (2009), uji *Hausman* bertujuan untuk menilai manakah model yang lebih sesuai adalah FEM, yang mengasumsikan bahwa efek individu bersifat tetap, atau REM, yang menganggap efek individu bersifat acak. Hipotesis dalam uji ini adalah sebagai berikut:

$$H_0: E(\mu_i, e_{it}) = 0 \quad (\text{model yang terbaik adalah REM})$$

$$H_1: E(\mu_i, e_{it}) \neq 0 \quad (\text{model yang terbaik adalah FEM})$$

Statistik uji:

$$W = \hat{q} \text{Var}(\hat{q})^{-1} \hat{q}$$

$$W = [\hat{\beta}_{FEM} - \hat{\beta}_{REM}] [\text{Var}[\hat{\beta}_{FEM} - \hat{\beta}_{REM}]]^{-1} [\hat{\beta}_{FEM} - \hat{\beta}_{REM}] \quad (2.8)$$

Keterangan:

$\hat{\beta}_{FEM}$: estimasi parameter koefisien dari FEM

$\hat{\beta}_{REM}$: estimasi parameter koefisien dari REM

Dalam uji *Hausman*, q adalah sebuah vektor yang merepresentasikan perbedaan antara dua vektor estimasi parameter regresi, yaitu $\hat{\beta}_{FEM}$ dan $\hat{\beta}_{REM}$. Secara matematis, $\hat{q} = \hat{\beta}_{FEM} - \hat{\beta}_{REM}$, dengan dimensi \hat{q} bergantung pada jumlah variabel independen dalam model (k), sehingga memiliki ukuran $k \times 1$. Dengan demikian, \hat{q} bukanlah konstanta atau matriks persegi, melainkan sebuah vektor yang berperan penting dalam menentukan model yang lebih sesuai antara FEM dan REM. Keputusan dalam uji Hausman didasarkan pada nilai probabilitasnya, jika probabilitas tersebut lebih kecil dari taraf signifikansi $\alpha = 5\%$ (0,05) maka H_0 ditolak, artinya model yang tepat adalah FEM. Sebaliknya jika nilai probabilitas lebih besar dari taraf signifikansi $\alpha = 5\%$ (0,05) maka tidak ada alasan untuk dapat menolak H_0 , artinya model yang tepat adalah REM.

2.5 Uji Asumsi Model

Dalam analisis regresi data panel, pemenuhan asumsi klasik memiliki peran krusial untuk memastikan bahwa model yang diestimasi memenuhi kriteria BLUE (*Best Linear Unbiased Estimator*). Menurut Gujarati & Porter (2009), asumsi yang harus dipenuhi meliputi linearitas, normalitas, homoskedastisitas, non-autokorelasi, dan tidak adanya multikolinearitas. Jika semua asumsi ini terpenuhi, maka estimasi parameter yang dihasilkan akan akurat dan dapat diandalkan.

2.5.1 Uji Normalitas

Uji normalitas bertujuan untuk menilai apakah residual yang diperoleh mengikuti distribusi normal atau tidak (Draper & Smith, 1992). Model regresi yang baik seharusnya menghasilkan residual yang berdistribusi secara normal. Salah satu metode yang sering diterapkan untuk menguji normalitas adalah uji *Jarque-Bera*. Menurut Mardiatmoko (2020), kriteria dalam pengujian normalitas adalah:

H_0 : residual terdistribusi secara normal

H_1 : residual tidak terdistribusi secara normal

Statistik uji:

$$JB = \frac{n}{2} \left(S^2 + \frac{(K-3)^2}{4} \right) \quad (2.10)$$

Dengan ,

$$S = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{\frac{3}{2}}}$$

$$K = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2}$$

Keterangan:

n : banyaknya data

S : *skewness*

K : *kurtosis*

Aturan keputusan:

- Jika $p\text{-value} >$ tingkat signifikansi α ($\alpha = 0,05$), maka gagal tolak H_0 sehingga residual terdistribusi secara normal
- Jika $p\text{-value} <$ tingkat signifikansi α ($\alpha = 0,05$), maka tolak H_0 sehingga residual tidak terdistribusi secara normal

2.5.2 Uji Multikolinearitas

Menurut Ghozali & Ratmono (2013), uji multikolinearitas dilakukan untuk mengevaluasi apakah terdapat hubungan di antara variabel independen dalam model regresi. Model regresi yang baik seharusnya tidak mengalami multikolinearitas, yang terjadi ketika variabel independen menunjukkan hubungan linear yang tinggi. Apabila variabel independen saling berkorelasi, maka timbul masalah multikolinearitas atau disebut juga variabel ortogonal. Variabel ortogonal mengacu pada variabel independen yang tidak berkorelasi satu sama lain, dimana nilai korelasi sama dengan nol. Deteksi multikolinearitas pada model regresi dapat dilakukan melalui analisis nilai VIF (*Variance Inflation Factor*). Umumnya, multikolinearitas dianggap terjadi jika nilai *tolerance* $< 0,10$ atau dengan nilai VIF > 10 . Sebaliknya jika VIF < 10 , maka model regresi dianggap bebas dari multikolinearitas. Statistik uji untuk VIF dapat dinyatakan sebagai berikut:

$$VIF_j = \frac{1}{1 - R_j^2} \quad (2.11)$$

Keterangan:

R_j^2 : koefisien determinasi yang menunjukkan sejauh mana variabel independen x_j berkorelasi dengan variabel independen lainnya.

2.5.3 Uji Heteroskedastisitas

Uji heteroskedastisitas dilakukan untuk mengidentifikasi adanya perbedaan varians residual di antara pengamatan dalam model regresi. Model regresi yang baik seharusnya memenuhi asumsi homoskedastisitas, yaitu varians residual yang konstan. Uji *Breusch-Pagan* merupakan salah satu metode yang dapat digunakan untuk mendeteksi heteroskedastisitas. Model regresi yang ideal tidak mengalami heteroskedastisitas (Tala & Karamoy, 2017). Hipotesis uji heteroskedastisitas adalah sebagai berikut:

H_0 : varians residual homogen (tidak terjadi heteroskedastisitas)

H_1 : varians residual tidak homogen (terjadi heteroskedastisitas)

Statistik uji *Breusch Pagan*, yang memiliki distribusi *chi-square* dari hipotesis tersebut adalah:

$$BP = \left(\frac{1}{2}\right) \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} \quad (2.12)$$

Keterangan:

f : vektor berukuran $n \times 1$ dengan $f = \left(\frac{\varepsilon_i^2}{\sigma^2} - 1\right)$ dan $\varepsilon_i = y_i - \hat{y}_i$

Z : matriks berukuran $n \times (p + 1)$ berisi vektor yang telah dinormalkan

Apabila $p \text{ value} > \alpha(0,05)$ maka tidak terdapat cukup bukti untuk menolak H_0 , berarti varians residual tetap konstan atau tidak terjadi heteroskedastisitas. Sebaliknya jika $p \text{ value} < \alpha(0,05)$ maka H_0 ditolak, sehingga dapat disimpulkan bahwa varians residual tidak homogen, atau terdapat heteroskedastisitas.

2.5.4 Uji Autokorelasi

Uji autokorelasi bertujuan untuk mendeteksi adanya keterkaitan antara *error term* pada periode t dengan *error term* pada periode sebelumnya ($t - 1$) dalam model regresi linear. Jika terdapat hubungan antara keduanya, maka model

mengalami masalah autokorelasi. Model regresi yang ideal tidak mengalami autokorelasi (Ghozali & Ratmono, 2013). *Breusch-Godfrey Test* merupakan salah satu metode yang digunakan untuk mendeteksi autokorelasi.

H_0 : tidak terjadi gejala autokorelasi

H_1 : terjadi gejala autokorelasi

Statistik uji *Breusch Godfrey*, yang memiliki distribusi *chi-square* dari hipotesis tersebut adalah:

$$BG = N \times R^2 \quad (2.13)$$

Keterangan:

N : jumlah observasi

R^2 : koefisien determinasi dari model tambahan

Apabila $p \text{ value} > \alpha(0,05)$ maka tidak menolak H_0 , yang berarti tidak terdapat gejala autokorelasi dalam model. Sebaliknya jika $p \text{ value} < \alpha(0,05)$ maka H_0 ditolak, yang menunjukkan adanya gejala autokorelasi.

2.6 Pendeteksian Pencilan

Pencilan merupakan nilai yang sangat ekstrem dan berada cukup jauh dari kelompok data lainnya, umumnya terletak sekitar tiga hingga empat kali simpangan baku dari rata-ratanya (Aunuddin, 1989). Menurut Bowerman & O'Connell (1990), beberapa metode yang bisa digunakan untuk mendeteksi pencilan dalam data di antaranya *leverage value* dan nilai *Studentized Deleted Residual* (TRES). Metode ini memungkinkan peneliti untuk mengidentifikasi data yang mungkin tidak sesuai dengan pola umum dalam data set, yang dapat memengaruhi hasil analisis regresi.

2.6.1 Nilai Pengaruh (*Leverage Value*)

Menurut Kutner dkk. (2004), *leverage value* digunakan untuk mengidentifikasi pencilan pada variabel prediktor dalam analisis regresi. Pendeteksian pencilan ini

dilakukan dengan menggunakan matriks \mathbf{H} , yang dikenal sebagai matriks hat atau *influence matrix*. Matriks \mathbf{H} dapat dirumuskan sebagai berikut:

$$H = X(X'X)^{-1}X' \quad (2.13)$$

Dengan \mathbf{H} merupakan matriks berisi nilai pengaruh (h_{it}) untuk subjek ke- i ($i= 1, 2, \dots, N$) dan periode ke- t ($t= 1, 2, \dots, T$). Serta $\mathbf{X} = [1, x_{i1}, x_{i2}, \dots, x_{np}]$ merupakan vektor baris yang terdiri dari nilai variabel pada pengamatan ke- i . Nilai *leverage* setiap observasi dapat diperoleh dari diagonal matriks \mathbf{H} , yang menunjukkan sejauh mana pengaruh suatu observasi terhadap estimasi parameter dalam model regresi (Widarjono, 2009).

2.6.2 TRES (*Studentized Deleted Residual*)

Menurut Kutner, dkk., (2004), TRES (*Studentized Deleted Residual*) merupakan statistik uji yang digunakan untuk mengidentifikasi pencilan pada variabel respons Y . Pengujian ini bertujuan untuk menentukan apakah suatu pengamatan memiliki nilai residual yang cukup ekstrem sehingga dapat dianggap sebagai pencilan. Hipotesis yang digunakan dalam pengujian ini sebagai berikut:

H_0 : Pengamatan ke- i tidak termasuk pencilan.

H_1 : Pengamatan ke- i termasuk pencilan

Dengan statistik uji sebagai berikut:

$$TRES_i = \frac{d_i}{s_{d_i}} = e_i \left[\frac{n-p-2}{JKS(1-h_{ii})-e^2} \right]^{\frac{1}{2}} \quad (2.14)$$

Keterangan:

e_i : $y_i - \hat{y}_i$

d_i : $y_i - \hat{y}_{i(i)}$

s_{d_i} : simpangan baku beda (d_i)

h_{ii} : $X_i(X'X)^{-1}X_i'$

JKS : jumlah kuadrat sisa
 p : banyaknya peubah prediktor
 n : banyaknya pengamatan ($i = 1, 2, \dots, n$)

Kriteria pengujiannya adalah:

$$|TRES_i| \begin{cases} \leq t_{\frac{\alpha}{2}, n-p-2}, H_0 \text{ diterima} \\ > t_{\frac{\alpha}{2}, n-p-2}, H_0 \text{ ditolak} \end{cases} \quad (2.15)$$

Semua kemungkinan nilai dari $|TRES_i|$ mengikuti distribusi t dengan derajat bebas sebesar $n - p - 2$, di mana p merupakan jumlah variabel independen, dan h_{ii} adalah elemen diagonal ke- i dari matriks H . Distribusi ini digunakan untuk mengevaluasi apakah suatu pengamatan merupakan pencilan dalam model regresi (Montgomery dkk., 2012).

2.7 Regresi *Robust*

Regresi *robust* merupakan pendekatan alternatif yang digunakan untuk memperoleh model yang lebih akurat dibanding estimasi yang menggunakan *Ordinary Least Squares* (OLS). Pendekatan ini berperan penting dalam analisis data yang mengandung pencilan (Chen, 2002). Menurut Rousseeuw & Hubert (1997), regresi *robust* bertujuan menyesuaikan model regresi dengan mempertimbangkan sebagian besar data, sambil menangani pencilan yang memiliki nilai sisaan besar. Dengan demikian, metode ini tidak menghilangkan data pencilan, tetapi mencari model yang tetap valid berdasarkan mayoritas data yang tersedia.

Menurut Chen (2002), terdapat beberapa metode pada regresi *robust*, antara lain:

1. Estimasi *Maximum Likelihood Type* (M): Diperkenalkan oleh Huber (1977), metode ini merupakan pendekatan yang simpel untuk menganalisis data yang berpotensi mengandung pencilan dalam variabel independen.

2. Estimasi *Least Median Squares* (LMS): Dikembangkan oleh Andrews (1974), metode ini menghasilkan penduga dengan cara mengabaikan sisaan yang signifikan, sehingga lebih optimal dalam menangani pencilan.
3. Estimasi *Least Trimmed Squares* (LTS): Metode ini memiliki *high breakdown point*, yang berarti lebih tahan terhadap pengaruh pencilan dalam proporsi tertentu dari keseluruhan data (Rousseeuw, 1984).
4. Estimasi *Scale* (S): Mirip dengan LTS, metode ini juga memiliki *high breakdown point*, tetapi lebih efisien dalam estimasi parameter (Rousseeuw, 1984).
5. Estimasi *Method of Moment* (MM): Merupakan kombinasi dari metode estimasi M dan S, yang bertujuan meningkatkan ketahanan terhadap pencilan serta menjaga efisiensi estimasi (Yohai, 1987).

2.8 Estimasi *Least Trimmed Square* (LTS)

Penduga *Least Trimmed Squares* (LTS) adalah metode dalam regresi *robust* yang bertujuan meminimalkan jumlah kuadrat galat dengan mempertimbangkan ketahanan terhadap pencilan (Rousseeuw & Hubert, 1997). Metode ini termasuk dalam kategori pendugaan parameter yang memiliki ketahanan tinggi terhadap pencilan, karena memiliki *breakdown point* (BDP) yang lebih besar dibandingkan metode penduga lainnya.

BDP merupakan ukuran seberapa besar proporsi data yang dapat mengandung pencilan sebelum suatu metode penduga kehilangan keandalannya. Nilai maksimum BDP untuk penduga *robust* adalah 50%, jika BDP melebihi 50%, penduga tidak dapat merepresentasikan mayoritas data secara efektif. Chen (2002) mendefinisikan penduga LTS sebagai berikut:

$$\hat{\beta}_{LTS} = \arg \min_{\beta} Q_{LTS}(\beta) \quad (2.16)$$

Dengan:

$Q_{LTS}(\beta)$: jumlah kuadrat galat dari h pengamatan yang telah dipilih

$$\sum_{i=1}^h e_{(i)}^2 = e_{(1)}^2 \leq e_{(2)}^2 \leq \dots \leq e_{(n)}^2$$

h : konstanta pemotongan

Nilai h dalam metode LTS merepresentasikan jumlah pengamatan yang dipilih untuk mengestimasi parameter dalam model regresi. Sebanyak $(n - h)$ pengamatan dengan galat terbesar tidak diperhitungkan dalam proses estimasi karena diberi bobot nol. Jika nilai h sama dengan n , maka metode LTS akan menghasilkan penduga yang sama dengan metode kuadrat terkecil (OLS). Menurut Chen (2002), semakin kecil nilai h , semakin tinggi ketahanan metode terhadap pencilan, tetapi hal ini dapat mengurangi efisiensi estimasi parameter. Sebaliknya, semakin besar nilai h mendekati n , estimasi parameter menjadi lebih efisien, tetapi metode ini menjadi kurang tahan terhadap pencilan. Rentang nilai h adalah:

$$\frac{n}{2} \leq h \leq n \quad (2.17)$$

Nilai h yang optimal dalam metode LTS adalah:

$$h = \left\lceil \frac{n}{2} \right\rceil + \left\lceil \frac{p+1}{2} \right\rceil \text{ atau } \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{p+2}{2} \right\rfloor \quad (2.18)$$

dengan p merupakan jumlah variabel prediktor dan n adalah jumlah pengamatan. Pemilihan nilai h yang optimal sangat berpengaruh terhadap nilai *breakdown point* (BDP), yaitu proporsi terkecil dari data yang dapat dipengaruhi oleh pencilan hingga menyebabkan penduga mengalami penyimpangan yang signifikan dibandingkan dengan penduga yang diperoleh dari data tanpa pencilan. Nilai *breakdown point* ini dapat dihitung menggunakan persamaan berikut:

$$\varepsilon_n^* = \frac{n-h}{n} \quad (2.19)$$

Nilai *breakdown point* maksimum adalah 50%. Jika pencilan terdapat 50% dari data, maka penduga akan sulit membedakan antara data asli dan pencilan. Metode MKT memiliki *breakdown point* sebesar 0%, artinya satu pencilan saja dapat mempengaruhi hasil estimasi secara signifikan. Dengan semakin tingginya nilai

breakdown point, maka penduga akan semakin tahan terhadap pengaruh pencilan, menjadikannya lebih *robust* (Rousseeuw & Leroy, 1987).

2.9 Pemilihan Model Terbaik Menggunakan Nilai R^2

Koefisien determinasi (R^2) digunakan untuk mengukur sejauh mana variabel independen (X) memberikan kontribusi terhadap model yang digunakan (Sugiyono, 2010). Analisis ini bertujuan untuk mengetahui persentase pengaruh variabel bebas secara simultan terhadap variabel terikat. Semakin tinggi nilai R^2 dan mendekati 1, maka semakin baik model tersebut dalam menjelaskan variasi data yang diamati. Namun, nilai R^2 dapat meningkat hanya dengan menambahkan variabel independen, meskipun variabel tersebut mungkin tidak signifikan. Untuk mengatasi masalah ini, sering kali disarankan untuk menggunakan *Adjusted R^2* , yang mengoreksi nilai R^2 dengan menghitung jumlah variabel independen, memberi penilaian yang lebih akurat tentang kebaikan model (Ghozali & Ratmono, 2013). Nilai R^2 dihitung berdasarkan rumus pada persamaan berikut:

$$R^2 = 1 - \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \hat{y}_{it})^2}{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_{it})^2} \quad (2.20)$$

2.10 Garis Kemiskinan

Garis kemiskinan berfungsi sebagai batas untuk menentukan apakah seseorang atau sekelompok orang tergolong miskin berdasarkan pengeluarannya. Badan Pusat Statistik (BPS) mengukur garis kemiskinan berdasar dari pengeluaran minimum untuk memenuhi kebutuhan dasar, yaitu makanan dan non-makanan. Terdapat dua komponen utama dalam penghitungan garis kemiskinan, yaitu Garis Kemiskinan Makanan (GKM) dan Garis Kemiskinan Bukan Makanan (GKBM). GKM dihitung

berdasarkan kebutuhan energi minimum 2.100 kalori per kapita per hari dari 52 jenis komoditas pangan, sedangkan GKBM mencakup kebutuhan dasar seperti perumahan, pakaian, pendidikan, dan kesehatan. Pada Maret 2024, BPS mencatat garis kemiskinan per kapita di Indonesia sebesar Rp582.932 per bulan, dengan 74,44% atau Rp433.906 dialokasikan untuk kebutuhan makanan.

Berdasarkan penelitian-penelitian sebelumnya, garis kemiskinan digunakan sebagai variabel dependen (Y), dengan variabel independen berupa jumlah penduduk (JP), kepadatan penduduk (KP), dan lowongan kerja terdaftar (LK). Jumlah penduduk per Kabupaten/kota mengukur total populasi yang tinggal di setiap wilayah administratif Kabupaten atau kota dalam suatu provinsi. Jumlah penduduk dihitung dengan menjumlahkan seluruh penduduk yang menetap di suatu Kabupaten atau kota selama periode tertentu, termasuk penduduk tetap dan pendatang yang memenuhi syarat tertentu. Informasi ini penting untuk analisis demografis, perencanaan pembangunan daerah, serta penentuan kebijakan pemerintah. Kepadatan penduduk mengukur jumlah penduduk yang tinggal di setiap kilometer persegi wilayah tertentu. Cara menghitungnya adalah dengan membagi jumlah total penduduk dengan luas wilayah (dalam km^2). Data lowongan kerja terdaftar menurut gender dihitung dengan memisahkan jumlah lowongan kerja berdasarkan jenis kelamin, misalnya laki-laki dan perempuan. Persentase lowongan kerja terdaftar untuk setiap gender dapat dihitung dengan membagi jumlah lowongan kerja terdaftar untuk gender tertentu dengan total lowongan kerja, kemudian dikalikan 100%.

III. METODE PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan pada semester ganjil tahun 2024/2025 bertempat di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini mencakup Garis Kemiskinan pada tingkat Kabupaten/kota di Provinsi Lampung. Sumber data diperoleh melalui laman resmi Badan Pusat Statistik (BPS) Provinsi Lampung (<https://lampung.bps.go.id/id>). Data yang digunakan merupakan data tahun 2019-2023 dengan Kabupaten/kota sebanyak 15 (Lampung Barat, Tanggamus, Lampung Selatan, Lampung Timur, Lampung Tengah, Lampung Utara, Way Kanan, Tulang Bawang, Pesawaran, Pringsewu, Mesuji, Tulang Bawang Barat, Pesisir Barat, Bandar Lampung, dan Metro). Data Garis kemiskinan (Y) sebagai variabel dependen dan tiga variabel independen meliputi jumlah penduduk (JP) (X_1), Kepadatan Penduduk (KP) (X_2), dan lowongan kerja terdaftar (LK) (X_3).

3.3 Metode Penelitian

Penelitian ini bertujuan untuk mengestimasi parameter serta menerapkan model regresi *robust* menggunakan penduga *Least Trimmed Square* (LTS) pada data Garis Kemiskinan di Provinsi Lampung periode 2019-2023. Proses analisis data dilakukan dengan bantuan software *R-Studio*. Adapun langkah-langkah yang digunakan dalam penelitian ini sebagai berikut:

1. Melakukan uji multikolinearitas untuk melihat nilai VIF apakah ada pada peubah prediktor yang berkorelasi tinggi.
2. Melakukan estimasi awal parameter regresi data panel menggunakan metode *Ordinary Least Squares* (OLS) untuk memperoleh gambaran awal hubungan antar variabel.
3. Menentukan model data panel yang paling sesuai dengan melakukan Uji *Chow* dan Uji *Hausman*.
4. Menguji asumsi klasik pada model data panel yang terpilih, meliputi:
 - a. Memeriksa normalitas residu dengan Uji *Jarque-Berra*.
 - b. Memeriksa homogenitas varians residu menggunakan Uji *Breusch-Pagan*.
 - c. Memeriksa bebasnya residu dari autokorelasi menggunakan Uji *Breusch-Godfrey*.
5. Melakukan deteksi pencilan menggunakan *Leverage Value* dan TRES (*Studentized Deleted Residual*) pada variabel prediktor maupun respon untuk mengidentifikasi pengamatan yang berpengaruh terhadap hasil analisis.
6. Melakukan pendugaan parameter regresi *robust* dengan metode LTS, melalui tahapan berikut:
 - a. Menentukan penduga awal β_0 menggunakan (MKT).
 - b. Menghitung residu kemudian menghitung nilai e_{it}^2 .
 - c. Menghitung jumlah kuadrat galat untuk h pengamatan awal dengan rumus $\frac{NT+p+1}{2}$.
 - d. Menghitung $\sum_{i=1}^h e_i^2$.

7. Melakukan uji parameter untuk mengukur pengaruh variabel prediktor terhadap variabel respon, baik secara parsial maupun simultan.
8. Validasi model dengan penduga LTS menggunakan koefisien determinasi R^2 serta efisiensi relatif model tersebut.

V. KESIMPULAN

Berdasarkan hasil analisis , dapat dilihat bahwa garis kemiskinan di Kabupaten/kota di Provinsi Lampung bisa dipengaruhi oleh perbedaan dalam biaya hidup, upah minimum, serta tingkat infrastruktur dan layanan sosial yang tersedia di masing-masing daerah. Berdasarkan analisis tersebut, model hasil regresi *robust* dengan penduga *Least Trimmed Square* (LTS) yang diperoleh adalah sebagai berikut:

$$Y_{LTS} = 44650 + 0.01931X_{1it} + 1.853X_{2it} - 1.5921X_{3it}$$

Dengan nilai R^2 sebesar 0.8909, menunjukkan bahwa ketiga variabel prediktor mampu menjelaskan 89,09% variasi garis kemiskinan. Oleh karena itu, dalam kondisi data yang mengandung pencilan, dapat disimpulkan bahwa penduga LTS menghasilkan model yang baik dalam pendugaan parameter regresi untuk data garis kemiskinan di Provinsi Lampung pada rentang tahun 2019-2023. Berdasarkan hasil pengujian simultan dan parsial, dapat disimpulkan bahwa pada taraf nyata 5%, jumlah penduduk (JP), kepadatan penduduk (KP), dan lowongan kerja (LK) secara bersama-sama berpengaruh signifikan terhadap garis kemiskinan di Provinsi Lampung.

DAFTAR PUSTAKA

- Akolo, I. R., & Nadjamuddin, A. (2022). Analisis Regresi *Robust* Estimasi Least Trimmed Square dan Estimasi Maximum Likelihood pada pemodelan IPM di Pulau Sulawesi. *Euler : Jurnal Ilmiah Matematika, Sains, dan Teknologi*, **10**(2), 211-221.
- Andrews, D. F. (1974). A *robust* method for multiple linear regression. *Technometrics*, **16**(4), 523-531.
- Anggun, P., Sugito, & Yasin, H. (2021). Pemodelan Regresi Ridge *Robust* S, M, MM-Estimator Dalam Penanganan Multikolinearitas dan Pencilan. *Jurnal Gaussian*, **10**(3), 402-412.
- Aunuddin. (1989). *Analisa Data*. Bogor: Institut Pertanian Bogor.
- Badan Pusat Statistik. (2024). *Persentase Penduduk Miskin Maret 2024*.
- Baltagi, B. H. (2008). *Econometric Analysis of Panel Data*. John Willey & Sons Inc.
- Bowerman, B. L. & R. T. O'Connell. (1990). *Linier Statistical Models An applied Approach*. USA: PWS Kent Publishing Company.
- Chen, C. H. (2002). *Robust* Statistical Methods in Linear Models with Outliers. *Communications in Statistics—Theory and Methods*, 1733-1746.
- Drapper, N. R. & Smith, H. (1992). *Analisis Regresi Terapan*. Jakarta: Alih Bahasa oleh PT. Gramedia Pustaka Utama.

- Firdaus, M. (2004). *Ekonometrika Suatu Pendekatan Aplikatif*. Jakarta: Bumi Aksara.
- Ghozali, I., & Ratmono, D. (2013). *Analisis Multivariat dan Ekonometrika. Teori, Konsep, dan Aplikasi dengan EViews 8*. Semarang: Penerbit Universitas Diponegoro.
- Gujarati, D. N. & Porter, D. C. (2009). *Basic Econometrics*. McGraw-Hill Education.
- Huber, P. J. (1977). Robust Method of Estimation of Regression Coefficients. *Statistics: A Journal of Theoretical and Applied Statistic*, **8**(1), 41-53.
- Kutner, Michael H., Nactsheim, Christopher J., Neter, J., & Li, W. (2004). *Applied Linear Statistical Models*. McGraw-Hill Education.
- Mardiatmoko, G. (2020). Pentingnya Uji Asumsi Klasik Pada Analisis Regresi Linier Berganda (Studi Kasus Penyusunan Persamaan Allometrik Kenari Muda [Canarium Indicum L.]). *Barekeng: Jurnal Ilmu Matematika & Terapan*, **14**(3), 333-342.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. John Wiley & Sons Inc.
- Rahman, Nur Fatimah & Wachidah, Lisnur. (2021). Perbandingan Estimasi Least Trimmed Square (LTS), Estimasi Scale (S), dan Estimasi Method Of Moment (MM) pada Analisis Regresi Robust. *Prosiding Statistika*, 613-621.
- Rousseeuw, P. J. (1984). Least Median of Square Regression. *Journal of American Statistical Association*, **79**(388), 871-880.
- Rousseeuw, P. J., & Hubert, M. (1997). Robust regression with both continuous and binary regressors. *Journal of Statistical Planning and Inference*, **57**(1), 153-163.
- Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust Regression and Outlier Detection*. John Wiley & Sons, Inc.

- Setiawan & Kusriani. (2010). *Ekonometrika*. Yogyakarta: Andi.
- Setyowati, E., Akbarita, R., & Robby, R. R. (2021). Perbandingan Regresi *Robust* Metode Least Trimmed Square (Lts) Dan Metode Estimasi-S Pada Produksi Padi Di Kabupaten Blitar. *Jurnal Matematika UNAND*, **10**(3), 329-341.
- Sriyana, J. (2015). *Metode Regresi Data Panel : Dilengkapi Analisis Kinerja Bank Syariah di Indonesia*. Yogyakarta: Ekonisia, FE UII.
- Sugiyono. (2010). *Statistika Untuk Penelitian*. Bandung: Alfabeta.
- Sukirno, S. (2019). *Pengantar Ekonomi Makro*. Jakarta: Rajawali Pers.
- Tala, O., & Karamoy, H. (2017). Analisis Profitabilitas dan Leverage Terhadap Manajemen Laba Pada Perusahaan Manufaktur di Bursa Efek Indonesia. *Jurnal Accountability*, **6**, 57-64.
- Widarjono, A. (2009). *Ekonometrika : Pengantar dan Aplikasinya*. Yogyakarta: Ekonisia.
- Winarno, W. W. (2007). *Analisis Ekonometrika dan Statistika Dengan Eviews*. Yogyakarta: UPP STIM YKPN.
- Yohai, V. J. (1987). High breakdown-point and high efficiency *robust* estimates for regression. *The Annals of statistics*, 642-656.