

**PEMODELAN REGRESI DATA PANEL DENGAN PENDEKATAN
RANDOM EFFECT MODEL UNTUK MENGANALISIS FAKTOR
PERTUMBUHAN EKONOMI DI PULAU SUMATERA**

(Skripsi)

Oleh

WINDA APRILIYANTI



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024**

ABSTRAK

PEMODELAN REGRESI DATA PANEL DENGAN PENDEKATAN *RANDOM EFFECT MODEL* UNTUK MENGANALISIS FAKTOR PERTUMBUHAN EKONOMI DI PULAU SUMATERA

Oleh

Winda Apriliyanti

Analisis regresi data panel merupakan metode statistika untuk menganalisis hubungan antara variabel dependen dan independen dengan mempertimbangkan variasi antar individu (*cross-section*) dan waktu (*time-series*). Dalam penelitian ini, digunakan *Random Effect Model* (REM) untuk menentukan model pertumbuhan ekonomi di Pulau Sumatera berdasarkan data Produk Domestik Regional Bruto (PDRB). Metode penelitian mencakup estimasi parameter regresi, pemilihan model terbaik, pengujian asumsi klasik, dan signifikansi koefisien regresi. Hasilnya menunjukkan bahwa *Random Effect Model* mampu menjelaskan sekitar 79,67% variabel dependen PDRB dengan variabel independen Penanaman Modal Dalam Negeri, Ekspor, dan Pengeluaran Konsumsi Pemerintah. Sisanya, sebesar 20,33%, diatribusikan pada faktor-faktor lain yang tidak dimasukkan dalam penelitian. Model persamaan yang dihasilkan adalah $\hat{Y}_{it}^* = 5,0995 + 0,0137X_{1it}^* + 0,0608X_{2it}^* + 0,6327X_{3it}^*$. Hasil ini memberikan wawasan tentang faktor-faktor yang signifikan dalam mempengaruhi pertumbuhan ekonomi regional.

Kata kunci: Data Panel, Produk Domestik Regional Bruto, *Random Effect Model*

ABSTRACT

PANEL DATA REGRESSION MODELING WITH A RANDOM EFFECT MODEL APPROACH TO ANALYZE ECONOMIC GROWTH FACTORS IN THE ISLAND OF SUMATRA

By

Winda Apriliyanti

Panel data regression analysis is a statistical method to analyze the relationship between dependent and independent variables by considering variations between individuals (cross-section) and time (time-series). This study uses the Random Effect Model (REM) to determine the economic growth model in the island of Sumatera based on Gross Regional Domestic Product (GRDP) data. The research method includes regression parameter estimation, selection of the best model, classical assumption testing, and significance of regression coefficients. The results show that the Random Effect Model is able to explain about 79,67% of the dependent variable of GRDP with the independent variables of Domestic Investment, Export, and Government Consumption Expenditure. The remainder 20,33% is attributed to additional elements omitted from the research. The resulting equation model is $\hat{Y}_{it}^* = 5,0995 + 0,0137X_{1it}^* + 0,0608X_{2it}^* + 0,6327X_{3it}^*$. These results provide insight into the factors that are significant in influencing regional economic growth.

Keywords: Panel Data, Gross Regional Domestic Product, Random Effect Model

**PEMODELAN REGRESI DATA PANEL DENGAN PENDEKATAN
RANDOM EFFECT MODEL UNTUK MENGANALISIS FAKTOR
PERTUMBUHAN EKONOMI DI PULAU SUMATERA**

Oleh

WINDA APRILIYANTI

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lampung



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024**

Judul Skripsi

: PEMODELAN REGRESI DATA PANEL
DENGAN PENDEKATAN *RANDOM*
EFFECT MODEL UNTUK MENGANALISIS
FAKTOR PERTUMBUHAN EKONOMI DI
PULAU SUMATERA

Nama Mahasiswa

: Winda Apriliyanti

Nomor Pokok Mahasiswa

: 2017031017

Jurusan

: Matematika

Fakultas

: Matematika dan Ilmu Pengetahuan Alam



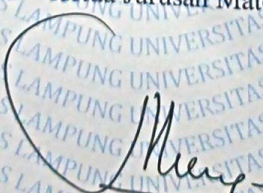

Drs. Nusrirwan, M.Si.

NIP. 196610101992051001


Dr. Notiragayu, S.Si., M.Si.

NIP. 197311092000122001

2. Ketua Jurusan Matematika


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

NIP. 197403162005011001

MENGESAHKAN

1. Tim Penguji

Ketua

: Drs. Nusyirwan, M.Si.



Sekretaris

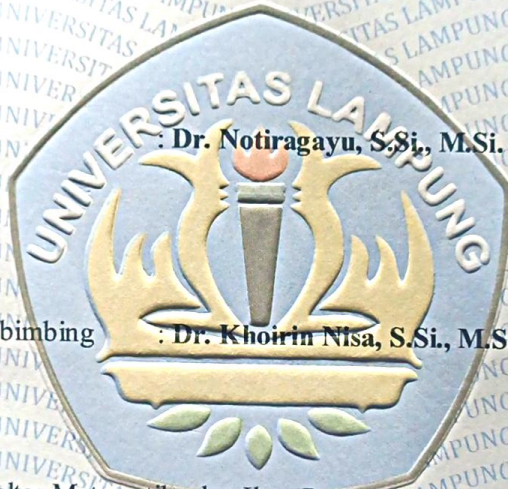
: Dr. Notiragayu, S.Si., M.Si.



Penguji

Bukan Pembimbing

: Dr. Khoirin Nisa, S.Si., M.Si.



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



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

NIP. 19711001 200501 1 002

Tanggal Lulus Ujian Skripsi: 21 Mei 2024

PERNYATAAN SKRIPSI MAHASISWA

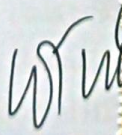
Saya yang bertanda tangan di bawah ini:

Nama : **Winda Apriliyanti**
Nomor Pokok Mahasiswa : **2017031017**
Jurusan : **Matematika**
Judul Skripsi : **PEMODELAN REGRESI DATA PANEL
DENGAN PENDEKATAN *RANDOM
EFFECT MODEL* UNTUK MENGANALISIS
FAKTOR PERTUMBUHAN EKONOMI DI
PULAU SUMATERA**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri dan semua tulisan yang tertuang dalam skripsi ini telah mengikuti kaidah karya penulisan ilmiah Universitas Lampung.

Bandar Lampung, 21 Mei 2024

Penulis



Winda Apriliyanti
NPM. 2017031017

RIWAYAT HIDUP

Penulis memiliki nama lengkap Winda Apriliyanti yang lahir di Kota Bandar Lampung pada tanggal 04 April 2003. Penulis merupakan anak pertama dari pasangan Bapak Madali dan Ibu Wiwin Ningsih.

Penulis menyelesaikan Pendidikan Sekolah Dasar di SD Negeri 3 Keteguhan pada tahun 2014, Pendidikan Sekolah Menengah Pertama di SMP Negeri 3 Bandar Lampung yang pada tahun 2017, dan Pendidikan Sekolah Menengah Atas di SMA Negeri 2 Bandar Lampung pada tahun 2020.

Penulis melanjutkan Pendidikan Strata Satu (S1) di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Lampung pada tahun 2020 melalui jalur SNMPTN. Sebagai bentuk penerapan ilmu perkuliahan, pada tahun 2023, penulis melaksanakan Kuliah Kerja Nyata (KKN) di Desa Bina Karya Utama Kecamatan Putra Rumbia Kabupaten Lampung Tengah dan Kerja Praktik (KP) di Badan Pusat Statistik Provinsi Lampung.

Selama menjadi mahasiswa, penulis aktif mengikuti berbagai organisasi. Pada tahun 2021, penulis bergabung dengan HIMATIKA Unila sebagai Anggota Bidang Kaderisasi dan Kepemimpinan. Pada tahun 2022, penulis menjabat sebagai Asisten Media di BEM FMIPA Unila. Kemudian, tahun 2023 penulis menjabat sebagai Bendahara Dinas Media dan Informasi BEM FMIPA Unila. Selain itu, penulis juga aktif mengikuti salah satu Program Merdeka Belajar-Kampus Merdeka yaitu Studi Independen Bangkit Academy by Google, GoTo, Traveloka dengan fokus pada pembelajaran *machine learning* pada tahun 2023.

PERSEMBAHAN

Alhamdulillah Rabbil' Alamin Puji dan Syukur tiada hentinya terpanjatkan kepada Allah SWT atas ridhonya sehingga penulis dapat menyelesaikan skripsi ini. Kupersembahkan karya kecil dan sederhana ini kepada:

Orang Tua Tercinta

Terima kasih Ibu dan Bapak atas kasih sayang, dukungan, motivasi, nasihat dan doa yang tidak berhenti sampai saat ini, karena doa dan didikan kalianlah yang membawaku bertahan dan kuat sampai sejauh ini.

Dosen Pembimbing dan Pembahas

Terima kasih kepada dosen pembimbing dan pembahas yang sudah sangat membantu, memberikan motivasi, memberikan kritik dan saran serta ilmu yang berharga.

Sahabat-Sahabatku

Kepada sahabat-sahabatku yang selalu memberikan keceriaan, semangat dan doa serta kenangan canda tawa selama masa perkuliahan.

Almamater Tercinta, Universitas Lampung

KATA INSPIRASI

“For indeed, with hardship (will be) ease. Indeed with hardship (will be) ease.”

[Q.S. Al-Insyirah (94) : 5-6]

"Most of the important things in the world have been accomplished by people who have kept on trying when there seemed to be no hope at all."

[Dale Carnegie]

“There is no failure. You either win or you learn.”

[Penulis]

SANWACANA

Puji syukur kehadirat Allah SWT, atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “**Pemodelan Regresi Data Panel dengan Pendekatan *Random Effect Model* untuk Menganalisis Faktor Pertumbuhan Ekonomi di Pulau Sumatera**”.

Selesainya skripsi ini adalah berkat motivasi, arahan, dan dukungan banyak pihak. Untuk itu, dengan segala kerendahan hati penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Drs. Nusyirwan, M.Si., selaku dosen pembimbing utama yang telah banyak memberikan bimbingan dan arahan, memberikan evaluasi dan saran yang membangun dalam proses penyusunan skripsi ini.
2. Ibu Dr. Notiragayu, S.Si., M.Si., selaku pembimbing kedua yang telah banyak membantu memberikan bimbingan, saran, dan nasehat selama proses penyusunan skripsi dan selama proses perkuliahan.
3. Ibu Dr. Khoirin Nisa, S.Si., M.Si., selaku dosen penguji atas kritik dan saran yang membangun guna penyempurnaan skripsi ini.
4. Ibu Dr. Asmiati, S.Si., M.Si., selaku dosen pembimbing akademik yang telah membimbing penulis dalam menyelesaikan permasalahan akademik.
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Bapak Dr. Eng. Heri Satria, S.Si., M.Si., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Bapak dan Ibuku yang tak henti-hentinya mencurahkan doa, ridho, kasih sayang, dan cinta untuk penulis hingga saat ini.

8. Saudara kandungku tersayang dan keluarga besar yang selalu mendukung penulis dalam menempuh pendidikan.
9. Achmad Daffa Muntashir yang selalu menemani, membantu, dan memberikan dukungan kepada penulis dalam menyelesaikan skripsi ini.
10. Teman-teman terbaikku Anisa Aprilia, Hilal Aulia, Yulian Khairunisa, Fegy Herlina, Ahmad Hanafi, Sisilia Amara, Muhtarom Ahkam, dan Dhifa Zhafirah yang selalu memberikan dukungan, semangat, dan motivasi selama menjalani pendidikan di Jurusan Matematika FMIPA Universitas Lampung.
11. Seluruh pihak yang tidak bisa penulis sebutkan satu per satu yang telah membantu penulis dalam menyelesaikan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna dan masih terdapat banyak kekurangan baik dalam penyajian maupun teknik penulisan. Oleh sebab itu, saran dan kritikan yang membangun senantiasa penulis harapkan agar dapat digunakan sebagai bahan perbaikan kedepannya.

Bandar Lampung, 21 Mei 2024

Penulis

Winda Apriliyanti

NPM. 2017031017

DAFTAR ISI

Halaman

DAFTAR TABEL	xv
DAFTAR GAMBAR.....	xvi
I. PENDAHULUAN.....	1
1.1 Latar Belakang dan Masalah	1
1.2 Tujuan Penelitian	3
1.3 Manfaat Penelitian	4
II. TINJAUAN PUSTAKA.....	5
2.1 Analisis Regresi.....	5
2.2 Analisis Regresi Data Panel	8
2.3 Estimasi Parameter Model Regresi Data Panel	9
2.3.1 <i>Common Effect Model</i> (CEM).....	9
2.3.2 <i>Fixed Effect Model</i> (FEM).....	11
2.3.3 <i>Random Effect Model</i> (REM).....	12
2.4 Pemilihan Model Regresi Data Panel.....	15
2.4.1 Uji <i>Chow</i>	15
2.4.2 Uji <i>Hausman</i>	16
2.4.3 Uji <i>Lagrange Multiplier</i>	17
2.5 Uji Asumsi Klasik Model Regresi Data Panel	17
2.5.1 Uji Normalitas	18
2.5.2 Uji Multikolinearitas	19
2.5.3 Uji Heteroskedastisitas	20
2.6 Pengujian Signifikansi Koefisien Regresi	20
2.6.1 Uji Simultan (Uji F)	21
2.6.2 Uji Parsial (Uji t).....	21
2.7 Koefisien Determinasi.....	22
2.8 Produk Domestik Regional Bruto	23
2.9 Transformasi Logaritma	24

III. METODELOGI PENELITIAN	25
3.1 Waktu dan Tempat Penelitian	25
3.2 Data Penelitian	25
3.3 Metode Penelitian	26
IV. HASIL DAN PEMBAHASAN	29
4.1 Analisis Deskriptif	29
4.2 Estimasi Parameter β dengan Metode <i>Generalized Least Square</i>	31
4.3 Estimasi Model Regresi Data Panel	32
4.3.1 <i>Common Effect Model</i> (CEM).....	32
4.3.2 <i>Fixed Effect Model</i> (FEM).....	33
4.3.3 <i>Random Effect Model</i> (REM).....	34
4.4 Pemilihan Model Regresi Data Panel.....	35
4.4.1 Uji <i>Chow</i>	35
4.4.2 Uji <i>Hausman</i>	36
4.4.3 Uji <i>Lagrange Multiplier</i>	37
4.5 Pengujian Asumsi Klasik Model Regresi Data Panel	38
4.5.1 Uji Normalitas	38
4.5.2 Uji Multikolinearitas	39
4.5.3 Uji Heteroskedastisitas	40
4.6 Uji Signifikansi Koefisien Regresi Data Panel	41
4.6.1 Uji Simultan (Uji F)	41
4.6.2 Uji Parsial (Uji t).....	42
4.6.3 Koefisien Determinasi	43
4.7 Interpretasi	43
V. KESIMPULAN.....	45
DAFTAR PUSTAKA.....	46
LAMPIRAN.....	49
Lampiran 1. Data Penelitian.....	50
Lampiran 2. Hasil Penelitian	51

DAFTAR TABEL

Tabel	Halaman
1. Data Penelitian.....	25
2. Analisis Deskriptif Variabel Penelitian	30
3. Estimasi Parameter <i>Common Effect Model</i>	32
4. Estimasi Parameter <i>Fixed Effect Model</i>	33
5. Variabel <i>Dummy</i> pada <i>Fixed Effect Model</i>	33
6. Estimasi Parameter <i>Random Effect Model</i>	34
7. Hasil Uji <i>Chow</i>	35
8. Hasil Uji <i>Hausman</i>	36
9. Hasil Uji <i>Breusch-Pagan</i>	37
10. Hipotesis Uji <i>Breusch-Pagan</i>	38
11. Hasil Uji <i>Breush-Pagan</i>	38
12. Uji Normalitas	39
13. Uji Normalitas Tahap 2	39
14. Uji Multikolinearitas	40
15. Uji Heteroskedastisitas	40
16. Uji Simultan	41
17. Uji Parsial.....	42
18. Koefisien Determinasi.....	43
19. Hasil <i>Random Effect Model</i>	43

DAFTAR GAMBAR

Gambar	Halaman
1. <i>Flowchart</i> Penelitian	28
2. Grafik PDRB Tahun 2016-2022	29

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis regresi merupakan teknik statistika yang bertujuan guna mempelajari korelasi antara satu atau lebih variabel bebas, yang juga dikenal sebagai variabel prediktor atau variabel eksplanatori dan variabel terikat yang disebut sebagai variabel respons dalam suatu model matematis. Tujuan analisis regresi yaitu untuk mengevaluasi dan mengukur seberapa besar pengaruh variabel bebas terhadap variabel terikat. Selain itu, dengan mempertimbangkan variabel independen yang ada, nilai variabel dependen dapat diprediksi.

Analisis regresi dapat membantu mengukur dan mengidentifikasi bagaimana masing-masing variabel independen mempengaruhi masing-masing variabel dependen. Terdapat beberapa jenis analisis regresi, termasuk regresi linear sederhana, berganda, dan regresi data panel. Data panel sendiri merupakan kombinasi data *cross-section* dan *time-series*. *Cross-section* mencakup observasi dari berbagai variabel pada satu saat atau periode waktu tertentu di antara berbagai unit atau individu sedangkan *time-series* mencakup variabel-variabel tertentu pada titik-titik waktu yang berurutan dan teratur. Menurut Nachrowi & Usman (2006), analisis regresi yang menggunakan gabungan data *cross-section* dan *time-series* cenderung lebih baik daripada yang menggunakan hanya data *cross-section* atau *time-series*.

Analisis regresi data panel adalah metode statistika yang digunakan untuk menggabungkan data lintas waktu (*time-series*) dan data lintas unit (*cross-*

sectional) dalam satu analisis. Dalam analisis regresi data panel, variabel-variabel independen dan variabel dependen diamati sepanjang waktu di berbagai unit atau individu. Mengacu pada Baltagi (2005), ada beberapa kelebihan dalam menggunakan data panel. Pertama, metode ini memungkinkan pengendalian terhadap heterogenitas variabel yang tidak disertakan pada model (*unobserved heterogeneity*). Kedua, pendekatan ini memungkinkan informasi yang lebih mendalam, meminimalisir tingkat kolinearitas yang ada antar variabel, meningkatkan derajat kebebasan, serta meningkatkan efisiensi analisis. Ketiga, data panel lebih cocok guna penelitian *dynamics of adjustment*. Keempat, pendekatan ini dapat mengukur dan mendeteksi efek yang tidak mudah diselesaikan pada data *cross-section* atau data *time series* saja. Terakhir, data panel mampu mengurangi bias yang timbul akibat menggabungkan individu dikarenakan melibatkan lebih banyak data. Keuntungan-keuntungan ini membuat pendekatan data panel menjadi pilihan yang sangat relevan dan efektif dalam analisis ekonomi dan sosial.

Menurut Widarjono (2009), terdapat tiga pendekatan utama untuk mengestimasi parameter model regresi. Pertama adalah *Common effect Model* (CEM) yang menggunakan metode *Ordinary Least Square* (OLS). Pendekatan ini mengasumsikan bahwa terdapat efek umum yang sama untuk semua unit dalam data panel. Kedua adalah *Fixed effect Model* (FEM) yang menggunakan metode *Least Square Dummy Variable* (LSDV). FEM memperhitungkan efek tetap atau karakteristik khusus dari setiap unit dengan memasukkan variabel *dummy* ke dalam model. Pendekatan terakhir adalah *Random effect Model* (REM) yang menggunakan metode *Generalized Least Square* (GLS). REM mengasumsikan bahwa efek antar unit bersifat acak, dan pendekatan ini memperhitungkan variasi dari efek-efek tersebut dalam estimasi parameter model.

Menurut Wooldridge (2010), *Random Effect Model* memiliki keuntungan signifikan, termasuk pengakomodasian variabilitas antar unit secara acak, fleksibilitas dalam mengatasi efek tetap tanpa asumsi tetap untuk setiap unit, penanganan masalah endogenitas, dan penanganan heteroskedastisitas serta

korelasi antar unit. Keuntungan-keuntungan ini membuat REM menjadi pendekatan yang kuat dan akurat dalam analisis data panel.

Analisis data panel kerap kali digunakan dalam melihat faktor-faktor yang mempengaruhi pertumbuhan ekonomi. Peran pertumbuhan ekonomi regional terhadap pertumbuhan ekonomi nasional sangat penting. Tingkat pertumbuhan produk domestik regional bruto (PDRB) adalah salah satu indikator yang digunakan untuk mengukur pertumbuhan ekonomi suatu wilayah. Terdapat beberapa faktor yang memungkinkan menjadi faktor penunjang pertumbuhan ekonomi diantaranya yaitu penanaman modal dalam negeri yang merupakan investasi yang dilakukan oleh investor domestik, baik individu ataupun perusahaan, di dalam negeri. Kemudian, ekspor yang merupakan kegiatan menjual barang atau jasa antar negara. Selain itu, pengeluaran konsumsi pemerintah yang merupakan total belanja yang dilakukan oleh pemerintah suatu negara guna mendanai berbagai program dan proyek yang mencakup layanan publik seperti infrastruktur, kesehatan, pendidikan, infrastruktur, dan lainnya.

Berdasarkan penjelasan yang telah disajikan, penelitian ini akan menggunakan pendekatan *Random Effect Model* untuk menganalisis data panel Produk Domestik Regional Bruto di Pulau Sumatera dari tahun 2016 hingga 2022.

1.2 Tujuan Penelitian

Berdasarkan latar belakang yang telah dipaparkan, penelitian ini bertujuan sebagai berikut:

1. Menentukan model Produk Domestik Regional Bruto di Pulau Sumatera dengan pendekatan *Random Effect Model* pada regresi data panel.
2. Menentukan variabel-variabel yang memiliki pengaruh dalam Produk Domestik Regional Bruto di Pulau Sumatera.

1.3 Manfaat Penelitian

Adapun manfaat dari penelitian ini yaitu sebagai berikut:

1. Meningkatkan pemahaman dan pengetahuan mengenai analisis regresi data panel.
2. Penelitian ini diharapkan dapat menjadi referensi bagi penelitian berikutnya dalam pengolahan data menggunakan data panel.

II. TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi adalah teknik statistik yang berguna untuk memeriksa serta membentuk model matematika yang menggambarkan relasi antar variabel (Montgomery & Peck, 1992). Dalam analisis regresi, terdapat dua jenis variabel yang umumnya dikenal, yaitu variabel independen dan variabel dependen. Variabel independen merujuk pada faktor-faktor yang mempengaruhi atau menjadi penyebab perubahan pada variabel dependen. Sebaliknya, variabel dependen merupakan variabel yang dipengaruhi oleh atau merupakan hasil dari variasi pada variabel independen.

Dalam analisis regresi, terdapat dua jenis utama, yaitu analisis regresi linear dan analisis regresi non-linear. Analisis regresi linear mencakup dua bentuk, yakni regresi linear sederhana dan regresi linear berganda. Menurut Montgomery & Peck (1992), regresi linear sederhana menggambarkan hubungan antara satu variabel independen dan variabel dependen dalam bentuk garis lurus. Model regresi linear sederhana dapat dijelaskan secara matematis sebagai berikut:

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (2.1)$$

dengan:

Y = variabel dependen

X = variabel independen

β_0 = intersep model

β_1 = *slope* regresi

ε = *error*

Jika hanya satu variabel independen yang digunakan maka disebut sebagai regresi linear sederhana. Namun, jika lebih dari satu variabel independen yang digunakan disebut sebagai regresi linear berganda. Analisis regresi linear berganda adalah metode statistik yang digunakan untuk mengidentifikasi hubungan sebab-akibat antara satu variabel dependen dengan beberapa variabel independen (Supranto, 2001). Berikut ini adalah rumus persamaan regresi linear berganda (Montgomery & Peck, 1992):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \varepsilon_i \quad (2.2)$$

dengan:

Y_i = variabel dependen pada pengamatan ke- i , dimana $i = 1, 2, \dots, n$

X_{ik} = variabel independen ke- k dengan $k = 1, 2, \dots, n$

β_0 = intersep model

$\beta_1, \beta_2, \dots, \beta_k$ = *slope* regresi

ε_i = *error*

oleh karena i menunjukkan pengamatan, maka n persamaan:

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_k X_{1k} + \varepsilon_1 \\ Y_2 &= \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_k X_{2k} + \varepsilon_2 \\ &\vdots \\ Y_n &= \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_k X_{nk} + \varepsilon_n \end{aligned} \quad (2.3)$$

Dalam bentuk matriks, persamaan (2.3) menjadi:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2.4)$$

Persamaan (2.4) dapat ditulis dalam matriks:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.5)$$

Menurut Gujarati (2009), terdapat beberapa asumsi pada model regresi linear berganda, adalah sebagai berikut:

1. Hubungan antara variabel dependen (Y) dan variabel independen (X) bersifat linear dalam parameter.
2. Tidak ada hubungan linier antara variabel independen atau tidak ada multikolinearitas antara variabel independen.
3. Nilai harapan atau rata-rata dari *error* ε_i adalah nol (0).

$$E(\varepsilon_i) = 0 ; i = 1, 2, \dots, n \quad (2.6)$$

Jika diekspresikan dalam bentuk matriks, persamaan (2.6) dapat dituliskan sebagai berikut:

$$E(\boldsymbol{\varepsilon}) = \begin{bmatrix} E(\varepsilon_1) \\ E(\varepsilon_2) \\ \vdots \\ E(\varepsilon_i) \\ \vdots \\ E(\varepsilon_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0 \quad (2.7)$$

4. Tidak terdapat korelasi antar *error*.

$$E(\varepsilon_i \varepsilon_j) = 0 ; i \neq j \quad (2.8)$$

5. Varian dari setiap *error* adalah sama (homoskedastisitas) atau tidak mengandung heteroskedastisitas.

$$\begin{aligned} Var(\varepsilon_i) &= E[\varepsilon_i - E(\varepsilon_i^2)] ; i = 1, 2, \dots, n \\ &= E(\varepsilon_i^2) \\ &= \sigma^2 \end{aligned} \quad (2.9)$$

Jika diekspresikan dalam bentuk matriks, persamaan (2.9) dapat dituliskan sebagai berikut:

$$\begin{aligned} Var(\boldsymbol{\varepsilon}) &= E(\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}') \\ &= E \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \dots & \varepsilon_n \end{bmatrix} \\ &= \begin{bmatrix} E(\varepsilon_1^2) & E(\varepsilon_1 \varepsilon_2) & \dots & E(\varepsilon_1 \varepsilon_n) \\ E(\varepsilon_2 \varepsilon_1) & E(\varepsilon_2^2) & \dots & E(\varepsilon_2 \varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(\varepsilon_i \varepsilon_1) & E(\varepsilon_i \varepsilon_2) & \dots & E(\varepsilon_i \varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(\varepsilon_n \varepsilon_1) & E(\varepsilon_n \varepsilon_2) & \dots & E(\varepsilon_n^2) \end{bmatrix} \end{aligned}$$

$$\begin{aligned} Var(\boldsymbol{\varepsilon}) &= \begin{bmatrix} \sigma_\varepsilon^2 & 0 & \dots & 0 & \dots & 0 \\ 0 & \sigma_\varepsilon^2 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_\varepsilon^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & \sigma_\varepsilon^2 \end{bmatrix} \\ &= \sigma_\varepsilon^2 \mathbf{I} \end{aligned} \quad (2.10)$$

6. *Error* berdistribusi normal.

$$\varepsilon_i \sim N(0, \sigma^2); i = 1, 2, \dots, n \quad (2.11)$$

2.2 Analisis Regresi Data Panel

Menurut Gujarati (2009), analisis regresi data panel adalah metode statistika yang digunakan untuk menganalisis hubungan antara variabel dependen dan variabel independen dengan mempertimbangkan variasi antar individu (*cross-section*) dan variasi waktu (*time-series*). Berikut adalah bentuk persamaan umumnya:

$$Y_{it} = \beta_{0it} + \beta_{1it}X_{1it} + \beta_{2it}X_{2it} + \dots + \beta_{kit}X_{kit} + \varepsilon_{it} \quad (2.12)$$

dengan:

- Y_{it} = variabel dependen untuk unit *cross-section* ke- i pada *time-series* ke- t
- X_{it} = variabel independen untuk unit *cross-section* ke- i pada *time-series* ke- t
- β = *intercept*
- β_i = *slope*, untuk $i = 1, 2, \dots, k$
- ε_{it} = *error* unit *cross-section* ke- i pada *time-series* ke- t
- i = 1, 2, ..., N yaitu banyaknya unit *cross-section*
- t = 1, 2, ..., N yaitu banyaknya unit *time-series*

Menurut Hsiao (2003), terdapat beberapa hal yang perlu dipertimbangkan untuk berbagai model yang mempertimbangkan perbedaan perilaku antar individu dan waktu. Beberapa model dituliskan sebagai berikut:

1. Koefisien *slope* tetap, sedangkan intersep bervariasi antar individu:

$$Y_{it} = \beta_{0i} + \sum_{k=1}^K \beta_k X_{kit} + \varepsilon_{it} \quad (2.13)$$

2. Koefisien *slope* tetap, sementara intersep bervariasi antar individu dan waktu:

$$Y_{it} = \beta_{0it} + \sum_{k=1}^K \beta_k X_{kit} + \varepsilon_{it} \quad (2.14)$$

3. Semua koefisien bervariasi antar individu:

$$Y_{it} = \beta_{0i} + \sum_{k=1}^K \beta_{ki} X_{kit} + \varepsilon_{it} \quad (2.15)$$

4. Semua koefisien bervariasi seiring waktu dan antar individu:

$$Y_{it} = \beta_{0it} + \sum_{k=1}^K \beta_{kit} X_{kit} + \varepsilon_{it} \quad (2.16)$$

dengan:

$i = 1, 2, \dots, N$

$t = 1, 2, \dots, T$

N = banyaknya unit *cross-section*

T = banyaknya unit *time-series*

Y_{it} = nilai variabel terikat unit *cross-section* ke- i dan *time-series* ke- t

X_{kit} = nilai variabel bebas ke- k unit *cross-section* ke- i dan *time-series* ke- t

β_{kit} = parameter yang akan diduga dari variabel bebas ke- k untuk unit *cross section* ke- i dan *time-series* ke- t

ε_{it} = *error* untuk unit *cross-section* ke- i dan *time-series* ke- t

K = jumlah parameter dalam regresi yang akan diduga

2.3 Estimasi Parameter Model Regresi Data Panel

Dalam analisis ini, terdapat tiga teknik yang umum diterapkan untuk mengestimasi parameter model yaitu *Common Effect Model* (CEM), *Fixed Effect Model* (FEM), dan *Random Effect Model* (REM).

2.3.1 *Common Effect Model* (CEM)

Dalam analisis data panel, CEM merupakan pendekatan yang paling mendasar. Teknik ini mengkombinasikan data *time-series* dan data *cross-section* tanpa memperhatikan varian antar unit *time-series* dan *cross-section*. Dalam model ini,

diasumsikan bahwa intersep dan *slope* memiliki nilai tetap bagi seluruh unit *time-series* dan *cross-section* (Gujarati, 2009). Berikut adalah persamaan CEM:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_k X_{kit} + \varepsilon_{it} \quad (2.17)$$

dengan:

Y_{it} = variabel terikat unit *cross-section* ke- i pada *time-series* ke- t

X_{it} = variabel bebas unit *cross-section* ke- i pada *time-series* ke- t

β_0 = intersep

β_i = *slope* untuk $i = 1, 2, \dots, k$

ε_{it} = *error* untuk unit *cross-section* ke- i pada *time-series* ke- t

2.3.1.1 Ordinary Least Square (OLS)

Metode estimasi parameter dalam model *Common effect* dilakukan menggunakan *Ordinary Least Square* atau Metode Kuadrat Terkecil (MKT). OLS adalah teknik yang mencari nilai residual sekecil mungkin dengan menjumlahkan kuadrat residual (Widarjono, 2009). Prinsip dasar metode kuadrat terkecil adalah sebagai berikut:

$$\varepsilon = Y - X\beta \quad (2.18)$$

Sehingga diperoleh jumlah kuadrat galat sebagai berikut:

$$\begin{aligned} \sum_{i=1}^n \varepsilon_i^2 &= \varepsilon' \varepsilon \\ &= (Y - X\beta)'(Y - X\beta) \\ &= (Y' - X'\beta')(Y - X\beta) \\ &= Y'Y - \beta'XY' - \beta'X'Y + \beta'X'X\beta \\ &= Y'Y - 2\beta'X'Y + \beta'X'X\beta \end{aligned} \quad (2.19)$$

Jika matriks $(X\beta)' = \beta'X'$, maka skalar $\beta'X'Y = Y'X\beta$. Untuk mendapatkan estimasi parameter β , hasil turunan diatur sama dengan nol, menghasilkan persamaan berikut:

$$\begin{aligned}
\frac{\partial(\varepsilon'\varepsilon)}{\partial\beta} &= \frac{\partial(Y'Y - 2\beta'X'Y + \beta'X'X\beta)}{\partial\beta} = 0 \\
-2X'Y + 2X'X\hat{\beta} &= 0 \\
2X'X\hat{\beta} &= 2X'Y \\
X'X\hat{\beta} &= X'Y \\
(X'X)^{-1}(X'X)\hat{\beta} &= (X'X)^{-1}X'Y \\
I\hat{\beta} &= (X'X)^{-1}X'Y \\
\hat{\beta} &= (X'X)^{-1}X'Y \tag{2.20}
\end{aligned}$$

Jika diasumsikan bahwa *slope* dan intersep memiliki nilai yang konstan untuk setiap data *time-series* dan *cross-section*, maka dapat diestimasi menggunakan metode OLS. Sehingga, model *common effect* dengan $N \times T$ pengamatan dinyatakan sebagai berikut:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \varepsilon_{it} ; i = 1, 2, \dots, N ; t = 1, 2, \dots, T \tag{2.21}$$

2.3.2 Fixed Effect Model (FEM)

Menurut Gujarati (2009), FEM merupakan suatu metode pada analisis data panel di mana setiap individu memiliki nilai intersepnnya sendiri, yang mencerminkan perbedaan antar individu. Dalam model ini, setiap individu memiliki intersep yang berbeda, tetapi intersep dianggap tetap sepanjang waktu (*time-invariant*), dapat ditulis sebagai berikut:

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_k X_{kit} + \varepsilon_{it} \tag{2.22}$$

dengan:

Y_{it} = variabel terikat unit *cross-section* ke- i pada *time-series* ke- t

X_{it} = variabel bebas unit *cross-section* ke- i pada *time-series* ke- t

β_{0i} = intersep

β_i = *slope* untuk $i = 1, 2, \dots, k$

ε_{it} = *error* untuk unit *cross-section* ke- i pada *time-series* ke- t

2.3.2.1 Least Squares Dummy Variable (LSDV)

Dalam FEM metode umum yang diterapkan yaitu *Least Squares Dummy Variable*. Menurut Widarjono (2009), LSDV menggunakan variabel *dummy* untuk memperhitungkan perbedaan nilai intersep antar unit dalam data panel, tetapi tetap mempertahankan koefisien *slope* variabel independen yang konstan. Persamaan model *fixed effect* dengan *dummy* variabel ditulis sebagai berikut:

$$Y_{it} = \beta_0 + \sum_{k=1}^K \beta_k X_{kit} + \sum_{i=2}^N \beta_{0i} D_{ki} + \varepsilon_{it} \quad (2.23)$$

dengan:

- Y_{it} = variabel terikat untuk unit *cross-section* ke- i dan *time-series* ke- t
- X_{kit} = variabel bebas ke- k untuk unit *cross-section* ke- i dan *time-series* ke- t
- D_{ki} = variable *dummy* ke- k untuk unit *cross-section* ke- i
- β_k = parameter untuk variabel ke- k
- ε_{it} = *error* untuk unit *cross-section* ke- i dan *time-series* ke- t

Menurut Singh & Sachdeva (2020), β_{0i} menunjukkan bahwa setiap objek memiliki intersep yang berbeda. Variabel *dummy* D_{1i} bernilai 1 untuk objek pertama dan 0 untuk objek lainnya. Jumlah variabel *dummy* yang dibuat adalah $N - 1$ dengan β_0 sebagai intersep untuk unit *cross-section* pertama.

2.3.3 Random Effect Model (REM)

REM diterapkan guna menyelesaikan masalah yang timbul dari FEM. Penggunaan variabel semu (*dummy*) pada pendekatan FEM menyebabkan kehilangan derajat kebebasan model dan dapat mengaburkan model asli. Dengan demikian, estimasi lebih baik dengan menggunakan *error components model* (ECM) atau *random effect model*. Persamaan *random effect model* dapat dirumuskan sebagai berikut:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \beta_k X_{kit} + W_{it} \quad (2.24)$$

dengan:

Y_{it} = variabel terikat untuk unit *cross-section* ke- i dan *time-series* ke- t

X_{kit} = variabel bebas untuk unit *cross-section* ke- i dan *time-series* ke- t

β_0 = intersep unit *cross-section*

β_k = *slope*

W_{it} = komponen *error* untuk unit *cross-section* ke- i dan *time-series* ke- t .

Diasumsikan bahwa β_{0i} merupakan variabel acak dengan mean β_0 . Oleh karena itu, intersep dapat ditulis:

$$\beta_{0i} = \beta_0 + \mu_i \quad (2.25)$$

Dengan μ_i merupakan *error random* yang mempunyai mean bernilai nol dan varian σ_μ^2 .

$$W_{it} = \mu_i + \varepsilon_{it} \quad (2.26)$$

Suku *error* gabungan W_{it} terdiri dari dua *error components*, yakni μ_i sebagai *error* pada unit individu serta ε_{it} sebagai gabungan *error* pada unit *time-series* serta *cross-section*. Karena hal ini, *random effect model* juga dikenal sebagai Model Komponen Galat (*Error Components Model/ECM*). Terdapat asumsi tertentu dalam *random effect model*, antara lain:

$$\begin{aligned} \mu_i &\sim N(0, \sigma_\mu^2) \\ \varepsilon_{it} &\sim N(0, \sigma_\varepsilon^2) \\ E(\mu_i \varepsilon_{it}) &= 0 \\ E(\mu_i \mu_j) &= 0 ; (i \neq j) \\ E(\varepsilon_{it} \varepsilon_{is}) &= E(\varepsilon_{it} \varepsilon_{jt}) = E(\varepsilon_{it} \varepsilon_{js}) = 0 ; (i \neq j ; t \neq s) \end{aligned} \quad (2.27)$$

2.3.3.1 Generalized Least Square (GLS)

Heteroskedastisitas sering muncul pada data *cross-section*. Menurut Greene (1997), untuk menangani heteroskedastisitas pada data *cross-section*, penggunaan estimasi dengan pembobotan (*weighted*) dapat digunakan, yang juga dikenal

sebagai *Generalized Least Squares*. Gujarati (2009) menyatakan bahwa pada data panel, metode GLS lebih baik dan konsisten daripada metode OLS. GLS menggunakan informasi secara eksplisit dan mampu menghasilkan *best linier unbiased estimator* (BLUE).

Diketahui:

$$\Omega_i = \begin{bmatrix} \sigma_\varepsilon^2 + \sigma_u^2 & \sigma_\varepsilon^2 & \cdots & \sigma_\varepsilon^2 \\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2 + \sigma_u^2 & \cdots & \sigma_\varepsilon^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2 & \cdots & \sigma_\varepsilon^2 + \sigma_u^2 \end{bmatrix} \quad (2.28)$$

Varian komponen Ω identik untuk setiap *cross section*. Sehingga varian komponen dari unsur gangguan untuk seluruh observasi dapat dituliskan:

$$S = \begin{bmatrix} \Omega_1 & 0 & \cdots & 0 \\ 0 & \Omega_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Omega_n \end{bmatrix} \quad (2.29)$$

Diberikan model regresi linear sebagai berikut:

$$Y = X\beta + \varepsilon; E(\varepsilon\varepsilon') = W \quad (2.30)$$

Persamaan (2.30) dikalikan dengan matriks V maka diperoleh:

$$VY = VX\beta + V\varepsilon \quad (2.31)$$

Menurut Taboga (2021), dengan $E[(V\varepsilon)(V\varepsilon)'] = I$ atau secara ekuivalen $VWV' = I$. $V'(VWV')V = V'V$ sehingga $V'V = W^{-1}$. Dari persamaan (2.18) maka dapat dibentuk persamaan baru untuk *error* yaitu:

$$V\varepsilon = VY - VX\beta \quad (2.32)$$

Kemudian turunan pertama dari jumlah kuadrat $V\varepsilon$ disamakan dengan 0, maka diperoleh:

$$\begin{aligned} \hat{\beta} &= [(VX)'(VX)]^{-1}(VX)'(VY) \\ &= (X'V'VX)^{-1}(X'V'VY) \\ &= (X'W^{-1}X)^{-1}(X'W^{-1}Y) \end{aligned} \quad (2.33)$$

yang merupakan *best linier unbiased estimator* (BLUE).

2.4 Pemilihan Model Regresi Data Panel

Beberapa pengujian yang bisa digunakan untuk memilih model yang optimal meliputi Uji *Chow*, Uji *Hausman*, dan Uji *Lagrange Multiplier*.

2.4.1 Uji *Chow*

Uji ini merupakan metode guna memilih apakah *common effect model* atau *fixed effect model* adalah yang terbaik untuk mengestimasi data panel (Widarjono, 2009). Pengujian ini didasarkan dengan hipotesis:

H_0 : model yang diterapkan yaitu *common effect model*

H_1 : model yang diterapkan yaitu *fixed effect model*

Statistik uji yang digunakan adalah uji F, yaitu:

$$F_{hitung} = \frac{(SSE_1 - SSE_2) / (n - 1)}{SSE_2 / (nt - n - k)} \quad (2.34)$$

dengan:

SSE_1 = sum square error dari model *common effect*

SSE_2 = sum square error dari model *fixed effect*

n = jumlah *cross-section*

nt = jumlah *cross-section* x jumlah *time-series*

k = jumlah variabel independen

Nilai F_{hitung} akan dibandingkan dengan nilai kritis $F_{(N-1, NT-N-K)}$ atau nilai $p - value$ yang dibandingkan dengan nilai α . Jika F_{hitung} lebih besar dari $F_{(N-1, NT-N-K)}$ atau $p - value$ lebih kecil dari α , maka hipotesis nol ditolak. Ini berarti terdapat cukup bukti untuk mendukung bahwa model yang paling sesuai untuk digunakan yaitu FEM. Sebaliknya, jika F_{hitung} lebih kecil dari

$F_{(N-1, NT-N-K)}$ atau $p - value$ lebih besar dari α , maka hipotesis nol tidak ditolak. Dalam kasus ini, terdapat cukup bukti untuk menyatakan bahwa model *common effect* lebih sesuai digunakan.

2.4.2 Uji Hausman

Uji ini merupakan metode yang diterapkan guna menentukan apakah model yang paling sesuai guna mengestimasi data panel adalah REM atau FEM (Gujarati, 2009). Uji ini didasarkan dengan hipotesis sebagai berikut:

H_0 : model yang digunakan yaitu *random effect model*

H_1 : model yang digunakan yaitu *fixed effect model*

Statistik uji yang diterapkan mengikuti distribusi *chi-squared* berdasarkan kriteria Wald, yaitu:

$$W = (\hat{\beta}_{MFE} - \hat{\beta}_{MRE})' [\text{var}(\hat{\beta}_{MFE} - \hat{\beta}_{MRE})]^{-1} (\hat{\beta}_{MFE} - \hat{\beta}_{MRE}) \quad (2.35)$$

dengan:

$\hat{\beta}_{MFE}$ = vektor *slope* model *fixed effect*

$\hat{\beta}_{MRE}$ = vektor *slope* model *random effect*

Nilai W akan dibandingkan dengan nilai kritis $\chi^2_{(\alpha, k)}$ atau perbandingan $p - value$ dengan nilai α . Jika W lebih kecil dari $\chi^2_{(\alpha, k)}$ atau $p - value$ lebih besar dari α , maka hipotesis nol tidak ditolak. Hal tersebut membuktikan bahwa model *random effect* lebih tepat untuk digunakan. Sebaliknya, jika W lebih besar dari $\chi^2_{(\alpha, k)}$ atau $p - value$ lebih kecil dari α , maka hipotesis nol ditolak. Dalam hal ini, model *fixed effect* lebih sesuai digunakan.

2.4.3 Uji Lagrange Multiplier

Uji ini bertujuan guna menentukan *common effect model* atau *random effect model* yang lebih sesuai guna mengestimasi data panel (Septianingsih, 2022). Uji ini didasarkan dengan hipotesis sebagai berikut:

H_0 : model yang diterapkan yaitu *common effect model*

H_1 : model yang diterapkan yaitu *random effect model*

Statistik uji yang digunakan adalah:

$$LM = \frac{nT}{2(T-1)} \left[\frac{\sum_{i=1}^n (T - e_i)^2}{\sum_{i=1}^n \sum_{t=1}^T e_{it}^2} \right] \quad (2.36)$$

dengan:

n = Jumlah individu

T = Jumlah periode waktu

e = Residual metode *Common Effect* (OLS)

Nilai LM akan dibandingkan dengan nilai kritis $\chi^2_{(\alpha,1)}$ atau perbandingan $p - value$ dengan nilai α . Jika LM lebih besar dari $\chi^2_{(\alpha,1)}$ atau $p - value$ lebih kecil dari α , maka hipotesis nol ditolak. Sehingga cukup bukti bahwa *random effect model* lebih sesuai untuk diterapkan. Sebaliknya, jika LM lebih kecil dari $\chi^2_{(\alpha,1)}$ atau $p - value$ lebih besar dari α , maka hipotesis nol tidak ditolak. Dalam hal ini, model *common effect* lebih sesuai digunakan.

2.5 Uji Asumsi Klasik Model Regresi Data Panel

Pada analisis data panel terdapat pengujian asumsi klasik, termasuk normalitas, multikolinearitas, dan heteroskedastisitas, dilakukan untuk memeriksa apakah model-model tersebut memenuhi asumsi-asumsi dasar.

2.5.1 Uji Normalitas

Untuk mengetahui apakah nilai *error* persamaan regresi memiliki distribusi normal, uji normalitas digunakan. Jika nilai *error* tersebut mendekati nilai rata-rata secara signifikan, maka dikatakan bahwa nilai *error* tersebut berdistribusi normal (Rahmadeni & Wulandari, 2017). Terdapat berbagai metode untuk membuktikan normalitas data, baik secara visual maupun melalui analisis statistik. Secara visual, normalitas dapat dinilai melalui histogram dan *boxplot*. Di sisi lain, analisis statistik dapat menggunakan Uji *Jarque Bera* (JB) yang dihitung dengan rumus berikut:

$$JB = \frac{n}{6} \left[S^2 + \frac{(Kr - 3)^2}{4} \right] \quad (2.37)$$

$$S = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{3/2}} \quad (2.38)$$

$$Kr = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^2} \quad (2.39)$$

dengan:

n = banyaknya data pengamatan

S = *skewness*

Kr = kurtosis.

Uji ini didasarkan dengan hipotesis sebagai berikut:

H_0 : residual berdistribusi normal

H_1 : residual tidak berdistribusi normal

Penolakan hipotesis dalam uji *Jarque Bera* bergantung pada perbandingan nilai JB dengan nilai kritis $x_{(\alpha,2)}^2$ atau perbandingan $p - value$ dengan nilai α . Jika nilai JB lebih kecil dari $x_{(\alpha,2)}^2$ atau $p - value$ lebih besar dari α maka hipotesis nol tidak ditolak. Dalam konteks ini, cukup bukti bahwa residual berdistribusi

normal. Sebaliknya, jika nilai JB lebih besar dari $\chi^2_{(\alpha,2)}$ atau $p - value$ lebih kecil dari α , maka hipotesis nol ditolak, sehingga dapat dikatakan bahwa residual tidak berdistribusi normal.

2.5.2 Uji Multikolinearitas

Uji ini digunakan untuk memeriksa ada atau tidaknya hubungan yang kuat antar variabel bebas dalam suatu model regresi (Rahmadeni & Wulandari, 2017). Jika model regresi melibatkan lebih dari satu variabel bebas, uji ini diperlukan.

Sebuah model regresi yang baik tidak seharusnya memiliki korelasi yang kuat antar variabel bebas. Korelasi yang kuat dapat mengakibatkan ketidakakuratan dalam estimasi parameter, sehingga sebaiknya dihindari. Untuk mendeteksi multikolinearitas, langkah-langkahnya sebagai berikut:

1. Memeriksa nilai *Variance Inflation Factor* (VIF) dari setiap variabel bebas. Multikolinearitas diindikasikan oleh nilai VIF yang lebih besar dari 10.

Menurut Gujarati (2009), rumus VIF yaitu:

$$VIF = \frac{1}{1 - R_j^2} \quad (2.40)$$

dimana R_j^2 merupakan koefisien determinasi yang diperoleh dari regresi variabel X_j terhadap variabel bebas lainnya. Jika variabel X_j tidak berkorelasi dengan variabel bebas lainnya, nilai R_j^2 akan kecil dan nilai VIF mendekati 1. Sebaliknya, jika X_j berkorelasi dengan variabel bebas lainnya, R_j^2 mendekati 1 dan nilai VIF akan besar (Montgomery & Peck, 1992).

2. Multikolinearitas dapat diidentifikasi dengan menggunakan matriks korelasi variabel bebas. Jika terdapat korelasi dengan nilai lebih dari 0,8 antar variabel bebas, ini mengindikasikan kemungkinan multikolinearitas (Gujarati, 2009).

2.5.3 Uji Heteroskedastisitas

Heteroskedastisitas terjadi ketika variabilitas *error* dalam model tidak stabil atau berbeda dari satu pengamatan ke pengamatan lainnya (Rahmadeni & Wulandari, 2017). Pemeriksaan heteroskedastisitas sangat penting, karena model yang baik seharusnya memiliki varians residual yang konstan. Heteroskedastisitas terjadi ketika asumsi ini tidak terpenuhi. Keberadaan heteroskedastisitas dapat membuat estimasi parameter menjadi tidak efisien (Nachrowi & Usman, 2006). Pengujian heteroskedastisitas dapat dilakukan dengan menggunakan Uji *Breusch-Pagan* (BP). Hipotesis yang digunakan dalam uji heteroskedastisitas yakni:

H_0 : tidak terdapat heteroskedastisitas pada data

H_1 : terdapat heteroskedastisitas pada data

Kriteria penolakan dalam uji *Breusch-Pagan* adalah jika nilai uji BP lebih kecil dari nilai kritis $\chi^2_{(\alpha,k)}$ atau *p-value* lebih besar dari α , maka hipotesis nol tidak ditolak. Berdasarkan hal ini, cukup bukti bahwa tidak ada heteroskedastisitas dalam data. Sebaliknya, jika nilai uji BP lebih besar dari $\chi^2_{(\alpha,k)}$ atau *p-value* lebih kecil dari α , hipotesis nol ditolak, menunjukkan adanya heteroskedastisitas dalam data.

2.6 Pengujian Signifikansi Koefisien Regresi

Dalam menguji signifikansi koefisien regresi, terdapat dua uji yang umum digunakan, yaitu Uji Simultan (Uji F) dan Uji Parsial (Uji t).

2.6.1 Uji Simultan (Uji F)

Uji Simultan bertujuan guna menilai apakah variabel bebas secara bersama-sama memiliki pengaruh terhadap variabel terikat (Alamsyah, *et al.*, 2022). Uji ini didasarkan pada hipotesis:

$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$ (secara simultan variabel bebas tidak memberikan pengaruh signifikan terhadap variabel terikat)

$H_1 : \text{Minimal ada satu } \beta_j \neq 0 ; j = 1, 2, \dots, k$ (secara simultan variabel bebas memberikan pengaruh signifikan terhadap variabel terikat).

Statistik uji yang digunakan adalah:

$$F_{hitung} = \frac{R^2 / (k - 1)}{(1 - R^2) / (n - k)} \quad (2.41)$$

dengan:

N = banyaknya unit *cross-section*

T = banyaknya periode waktu (unit *time-series*)

k = banyaknya variabel bebas

Kriteria uji yang digunakan adalah Jika F_{hitung} lebih dari $F_{(N+K-1, NT-N-K)}$, hipotesis nol akan ditolak. Ini menandakan bahwa variabel bebas memiliki pengaruh keseluruhan yang signifikan terhadap variabel terikat. Sebaliknya, jika F_{hitung} kurang dari $F_{(N+K-1, NT-N-K)}$, maka hipotesis nol tidak ditolak, menunjukkan bahwa variabel bebas tidak memiliki pengaruh keseluruhan yang signifikan terhadap variabel terikat.

2.6.2 Uji Parsial (Uji t)

Uji Parsial bertujuan guna menilai pengaruh variabel bebas secara individu terhadap variabel terikat (Alamsyah, *et al.*, 2022). Hipotesis yang digunakan dalam Uji t yakni:

$H_0 : \beta_j = 0$ (secara individual variabel bebas tidak memberikan pengaruh signifikan terhadap variabel terikat)

$H_1 : \beta_j \neq 0 ; j = 1, 2, \dots, k$ (secara individual variabel bebas memberikan pengaruh signifikan terhadap variabel terikat).

Uji t didefinisikan sebagai berikut:

$$t = \frac{\beta_i}{Se(\beta_i)} \quad (2.42)$$

dengan:

β_i = merupakan koefisien regresi ke- i

$Se(\beta_i)$ = standar *error* koefisien regresi.

Kriteria uji yang digunakan adalah jika $|t_{hitung}| > t_{(\alpha/2, NT-N-K)}$ maka tolak H_0 , itu menandakan adanya pengaruh secara individu dari variabel bebas terhadap variabel terikat. Sebaliknya, jika H_0 tidak ditolak, itu menunjukkan bahwa tidak ada pengaruh signifikan secara individu dari variabel bebas terhadap variabel terikat.

2.7 Koefisien Determinasi

Koefisien determinasi atau R^2 bertujuan guna mengevaluasi sejauh mana variabel independen mempengaruhi variabel dependen atau seberapa besar variabilitas dalam variabel independen dapat dijelaskan oleh variabel dependen (Rahmadeni & Wulandari, 2017). Nilai R^2 mendekati 1 menunjukkan bahwa variabel bebas memiliki kemampuan yang baik dalam menjelaskan variabel terikat. Sebaliknya, nilai mendekati 0 menunjukkan bahwa kemampuan variabel bebas dalam menjelaskan variabel terikat terbatas. Oleh karena itu, kualitas suatu model regresi dapat dinilai berdasarkan nilai R^2 , yang berada dalam rentang antara 0 dan 1 (Silalahi, *et al.*, 2011). Koefisien determinasi dapat ditulis:

$$R^2 = 1 - \frac{\sum e_i^2 / (n - k)}{\sum (Y_i - \bar{Y})^2 / (n - 1)} \quad (2.43)$$

dengan:

k = jumlah parameter (termasuk intersep)

n = jumlah observasi.

2.8 Produk Domestik Regional Bruto

PDRB adalah ukuran nilai total barang serta jasa yang diperoleh dalam suatu wilayah atau regional, biasanya dalam jangka waktu satu tahun. PDRB mencakup nilai tambah dari semua sektor ekonomi di suatu wilayah, termasuk pertanian, industri, perdagangan, jasa, dan sektor-sektor lainnya. PDRB dapat dihitung dengan menggunakan harga berlaku atau harga konstan. PDRB atas dasar harga konstan menggunakan harga pasar pada tahun tertentu sebagai dasar, memungkinkan perbandingan data dari tahun ke tahun tanpa terpengaruh oleh fluktuasi harga. Di sisi lain, PDRB atas dasar harga berlaku mengukur nilai barang dan jasa dengan menggunakan harga pasar saat ini (BPS, 2023).

Untuk menganalisis pertumbuhan ekonomi suatu wilayah, kita memperhatikan PDRB sebagai indikator utama. Pertumbuhan ekonomi daerah ini dipengaruhi oleh beberapa faktor kunci, termasuk investasi melalui penanaman modal dalam negeri, aktivitas ekspor, serta tingkat konsumsi yang dilakukan oleh pemerintah.

- a. Penanaman Modal Dalam Negeri (PMDN) di Indonesia diatur oleh Undang-Undang Nomor 25 Tahun 2007 tentang Penanaman Modal. Menurut Undang-Undang ini, PMDN adalah kegiatan penanaman modal yang dilakukan oleh investor dalam bentuk uang, barang modal, bahan baku, sarana produksi, merek, paten, dan/atau pengetahuan, serta keahlian yang bernilai ekonomis untuk melakukan usaha di dalam negeri.
- b. Ekspor merupakan kegiatan penjualan barang atau jasa dari Indonesia ke negara lain, termasuk hasil produksi dalam negeri dan barang impor yang diolah ulang atau diolah lebih lanjut di Indonesia sebelum diekspor. Ekspor

adalah salah satu indikator penting dalam mengukur kinerja ekonomi suatu negara dan berperan dalam meningkatkan pendapatan dan pertumbuhan ekonomi nasional (BPS, 2023).

- c. Pengeluaran Konsumsi Pemerintah merujuk pada total belanja yang dilakukan oleh pemerintah suatu negara untuk membiayai berbagai program dan proyek, termasuk pendidikan, kesehatan, infrastruktur, dan layanan publik lainnya (BPS, 2023).

2.9 Transformasi Logaritma

Transformasi logaritma adalah teknik analisis statistik yang digunakan untuk memperbaiki sifat distribusi data, menstabilkan varians, dan memenuhi asumsi model regresi. Teknik ini sangat berguna untuk data yang memiliki distribusi skewed atau rentang nilai yang luas (Wooldridge, 2010). Transformasi logaritma mengubah nilai data X menjadi $\log(X)$, umumnya menggunakan logaritma natural berbasis e . Transformasi ini sering diterapkan pada data untuk membuatnya lebih mendekati distribusi normal dan untuk mengatasi heteroskedastisitas. Berikut adalah penerapan transformasi logaritma dalam model regresi linear:

1. Model Log-Linier: Transformasi logaritma pada variabel dependen.

$$\log(Y) = \beta_0 + \beta_1 X + \varepsilon \quad (2.44)$$

2. Model Lin-Log: Transformasi logaritma pada variabel independen.

$$Y = \beta_0 + \beta_1 \log(X) + \varepsilon \quad (2.45)$$

3. Model Log-Log: Transformasi logaritma pada variabel dependen dan independen.

$$\log(Y) = \beta_0 + \beta_1 \log(X) + \varepsilon \quad (2.46)$$

III. METODELOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada Semester Ganjil tahun akademik 2023/2024 dengan melakukan penelitian secara studi pustaka di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan pada penelitian ini merupakan data sekunder sebanyak 70 observasi yang terdiri dari Produk Domestik Regional Bruto (PDRB) di Provinsi Aceh, Bangka Belitung, Bengkulu, Jambi, Kep. Riau, Lampung, Riau, Sumatera Barat, Sumatera Selatan, dan Sumatera Utara pada tahun 2016 sampai 2022 yang diperoleh dari website BPS Republik Indonesia. Dengan variabel Penanaman Modal Dalam Negeri (X_1), Ekspor (X_2) dan Pengeluaran Konsumsi Pemerintah (X_3) sebagai variabel bebas, sedangkan variabel PDRB (Y) sebagai variabel terikat. Pengolahan data pada penelitian ini dibantu menggunakan *software* R.

Tabel 1. Data Penelitian

No.	Provinsi	Tahun	Y	X_1	X_2	X_3
1.	Aceh	2016	116374	2456	2899	31803
2.	Aceh	2017	121241	783	3931	34058
3.	Aceh	2018	126824	970	2969	35423
4.	Aceh	2019	132070	3607	1312	38121
5.	Aceh	2020	131581	8241	2605	36496

6.	Aceh	2021	135251	7905	5369	37919
7.	Aceh	2022	140948	4424	7458	37623
⋮	⋮	⋮	⋮	⋮	⋮	⋮
64.	Kep. Riau	2016	162853	493	10090	13810
65.	Kep. Riau	2017	166082	1398	12182	14737
66.	Kep. Riau	2018	173499	4386	13193	14733
67.	Kep. Riau	2019	181878	5656	12790	15131
68.	Kep. Riau	2020	174959	14249	11952	14163
69.	Kep. Riau	2021	180952	9769	15969	14238
70.	Kep. Riau	2022	190164	4817	19626	14558

3.3 Metode Penelitian

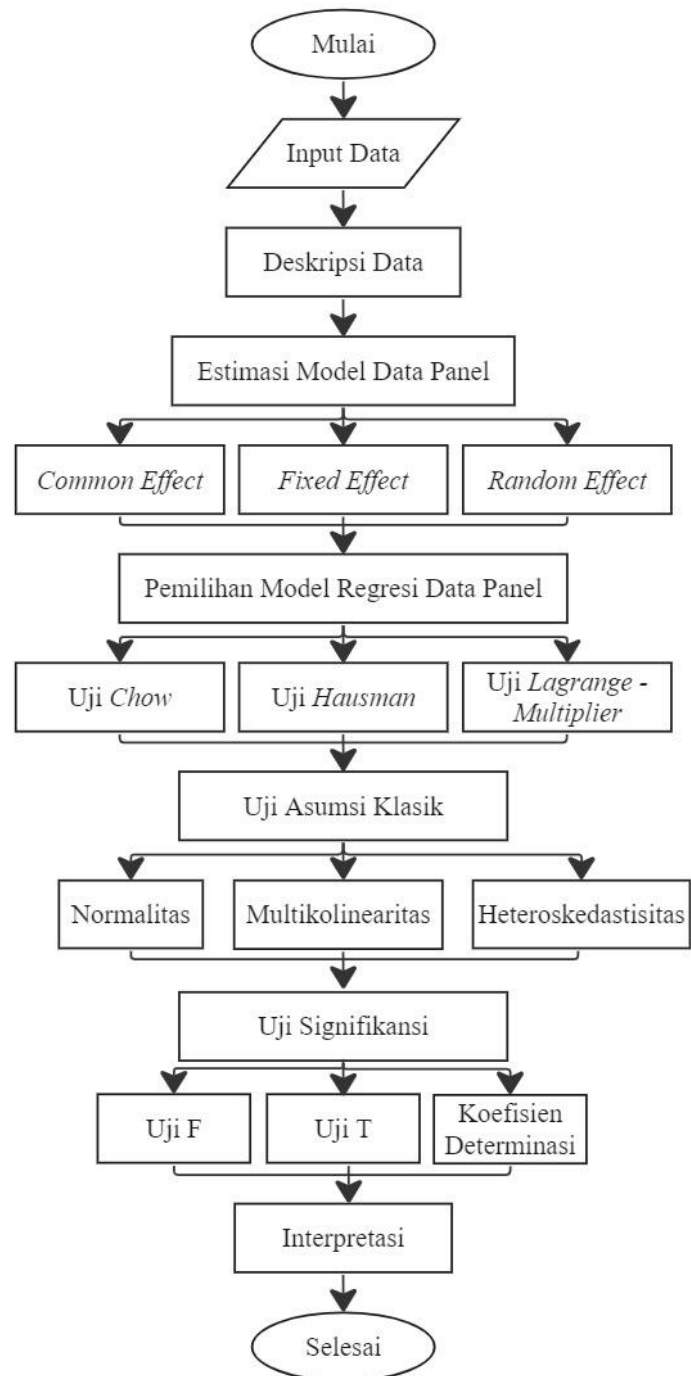
Berikut adalah langkah-langkah yang diambil dalam penelitian ini:

1. Menganalisis data penelitian secara deskriptif.
2. Menganalisis model regresi data panel dengan tiga pendekatan, yaitu *Common Effect Model*, *Fixed Effect*, dan *Random Effect Model*.
3. Memilih model terbaik regresi data panel dengan menggunakan Uji *Chow*, Uji *Hausman*, dan uji *Lagrange Multiplier*.
4. Menguji asumsi klasik pada model terbaik regresi data panel yang meliputi:
 - a. Uji normalitas dengan menggunakan Uji *Jarque Bera*. Jika
 - b. Uji multikolineritas dengan melihat nilai VIF dari setiap variabel bebas.
 - c. Uji heteroskedastisitas dengan menggunakan Uji *Breusch Pagan*.

Pada uji asumsi klasik banyak cara yang dapat dilakukan jika terdapat masalah pada uji normalitas, multikolinearitas, dan heteroskedastisitas, yaitu dengan cara sebagai berikut:

1. Transformasi variabel, dapat dilakukan jika model tidak berdistribusi normal atau terdapat masalah multikolinearitas. Transformasi data dapat dilakukan dengan membentuk data menjadi *log* atau *ln* (logaritma natural).
2. Menghapus *outlier* atau menambahkan data baru juga dapat mengatasi masalah-masalah yang biasa terjadi pada uji asumsi klasik.

5. Menguji signifikansi koefisien model regresi data panel yang terbaik mencakup:
 - a. Uji Simultan (Uji F) yang dilakukan untuk menilai pengaruh secara bersama-sama dari semua variabel bebas terhadap variabel terikat.
 - b. Uji Parsial (Uji T) dilakukan untuk menilai pengaruh variabel bebas secara individu terhadap variabel terikat.
 - c. Koefisien determinasi (R^2) dilakukan untuk mengukur sejauh mana variabel bebas mempengaruhi variabel terikat.
3. Menginterpretasi dan menarik kesimpulan pada model regresi data panel yang diperoleh.



Gambar 1. *Flowchart* Penelitian

V. KESIMPULAN

Berdasarkan uraian hasil analisis dan pembahasan pada penelitian ini, disimpulkan bahwa model regresi data panel yang lebih sesuai untuk memodelkan variabel Produk Domestik Regional Bruto di Pulau Sumatera dari tahun 2016-2022 yang telah ditransformasi logaritma adalah *Random Effect Model* dengan efek individu, diperoleh model persamaan hasil estimasi adalah

$$\hat{Y}_{it}^* = 5,0995 + 0,0137X_{1it}^* + 0,0608X_{2it}^* + 0,6327X_{3it}^*$$

Dengan diperoleh nilai *R-Squared* sebesar 0,7967, artinya variabel Penanaman Modal Dalam Negeri, Ekspor, dan Pengeluaran Konsumsi Pemerintah mampu menjelaskan variabel Produk Domestik Regional Bruto sebesar 79,67%, sedangkan sisanya yakni sebesar 20,33% dijelaskan oleh variabel lain di luar model.

DAFTAR PUSTAKA

- Haryati, A. O. 2016. *Regresi Data Panel dengan Metode Cross Section Weighted*. Fakultas Sains Dan Teknologi, UIN Suska Riau.
- Alamsyah, I. F., Esra, R., Awalia, S., & Nohe, D. A. 2022. Analisis Regresi Data Panel Untuk Mengetahui Faktor Yang Memengaruhi Jumlah Penduduk Miskin Di Kalimantan Timur. In *Prosiding Seminar Nasional Matematika dan Statistika* (Vol. 2).
- Aziz, A. 2010. *Ekonometrika Teori dan Praktek Eksperimen dengan Matlab*. UIN-Malang Press, Malang.
- Baltagi, B. H. 2005. *Econometric Analysis of Panel Data*. 3th Edition. John Willey & Sons, Southern Gate.
- BPS. 2023. *Produk Domestik Regional Bruto (Lapangan Usaha)*. Retrieved from Badan Pusat Statistik: <https://www.bps.go.id/subject/52/produk-domestik-regional-bruto--lapangan-usaha-.html#subjekViewTab1>
- BPS. 2023. *Produk Domestik Regional Bruto (Pengeluaran)*. Retrieved from Badan Pusat Statistik: <https://www.bps.go.id/subject/171/produk-domestik-regional-bruto--pengeluaran-.html#subjekViewTab1>
- Ghozali, I. 2016. *Aplikasi Analisis Multivariate dengan Program IBM SPSS 23*. Ed. ke-8. Badan Penerbit Universitas Diponegoro, Semarang.
- Greene, W.H. 1997. *Economic Analysis*. Prentice-Hall International, Lnc.USA.
- Gujarati, D.N. & Porter, D.C. 2009. *Basic Econometrics*. 5th Edition. McGraw Hill Companies, New York.

- Hsiao, C. 2003. *Analysis of Panel Data*. 2nd Edition. Cambridge University Press, California.
- Montgomery, D. C. & Peck, E. A. 1992. *Introduction to Linear Regression Analysis*. 2nd Edition. John Wiley & Sons, Toronto.
- Nachrowi, D. N. & Usman, H. 2006. *Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan*. Lembaga Penerbit FE UI, Jakarta.
- Rahmadeni, R., & Wulandari, N. 2017. Analisis Faktor-Faktor yang Mempengaruhi Inflasi pada Kota Metropolitan di Indonesia dengan Menggunakan Analisis Data Panel. *Jurnal Sains Matematika Dan Statistika*. 3(2): 34-42.
- Septianingsih, A. 2022. Pemodelan Data Panel Menggunakan Random Effect Model untuk Mengetahui Faktor yang Mempengaruhi Umur Harapan Hidup di Indonesia. *Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*. 3(3): 525-536.
- Silalahi, D., Sitepu, R., & Tarigan, G. 2014. Analisis ketahanan pangan provinsi Sumatera Utara dengan metode regresi data panel. *Jurnal Saintia Matematika*. 2(3): 237-251.
- Singh, S. & Sachdeva, T. K. 2020. Financial Performance of Selected IT Companies in India: A Panel Data Approach. *International Journal of Arts, Science and Humanities*. 7(3): 7-14.
- Supranto, J. 2001. *Statistika Teori dan Aplikasi*. Erlangga, Jakarta.
- Taboga, Marco. 2021. *Generalized Least Squares, Lectures on probability theory and mathematical statistics*. K indle Direct Publishing.
- Widarjono, A. 2009. *Ekonometrika Pengantar dan Aplikasinya*. Ed. ke-3. Ekonisia Fakultas Ekonomi Universitas Islam Indonesia, Yogyakarta.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, London.