

**PENERAPAN METODE *ORDINARY LEAST SQUARE* (OLS) DALAM  
ESTIMASI PARAMETER MODEL REGRESI DATA PANEL *COMMON  
EFFECT* PADA DATA INDEKS PEMBANGUNAN MANUSIA (IPM)  
PROVINSI SUMATERA UTARA TAHUN 2012-2020**

**(SKRIPSI)**

**Oleh**

**SILVIA RISMA FEBIOLA MARBUN  
2017031046**



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



## **ABSTRACT**

### **APPLICATION OF THE ORDINARY LEAST SQUARE (OLS) METHOD IN ESTIMATING PARAMETERS OF THE COMMON EFFECT PANEL DATA REGRESSION MODEL ON HUMAN DEVELOPMENT INDEX (HDI) DATA IN NORTH SUMATRA FOR THE YEARS 2012-2020.**

**BY**

**SILVIA RISMA FEBIOLA MARBUN**

Finding causal correlations between variables is one of the benefit and objectives of regression analysis. A common problem that often arises in regression analysis is the availability of sufficient data to represent the variables to be analyzed. Data is often available in a short time span and limited in number of individuals or units. To overcome this challenge, panel data is utilized because the concept is to combine unit or individual data with time series data. The Common Effect Model is one of several approaches in panel data that implemented on the Human Development Index data in North Sumatra from 2012 to 2020, where the model parameters are estimated using the Ordinary Least Square method. The results on parameters show a significant influence between the dependent variables on the independent variables is 91.06% and the following model is obtained:  $\hat{Y}_{it} = 25.78 + 0.2178X_{1it} + 1.283X_{2it} + 0.6024X_{3it} + 0.001086 X_{4it} - 0.1043X_{5it}$

**Keywords:** common effect model, ordinary least square, panel data



## ABSTRAK

### **PENERAPAN METODE *ORDINARY LEAST SQUARE* (OLS) DALAM ESTIMASI PARAMETER MODEL REGRESI DATA PANEL *COMMON EFFECT* PADA DATA INDEKS PEMBANGUNAN MANUSIA (IPM) PROVINSI SUMATERA UTARA TAHUN 2012-2020**

**OLEH**

**SILVIA RISMA FEBIOLA MARBUN**

Menemukan korelasi sebab-akibat antara berbagai variabel merupakan salah satu manfaat dan tujuan dalam analisis regresi. Masalah umum yang sering muncul dalam analisis regresi adalah ketersediaan data yang memadai untuk mewakili variabel yang akan dianalisis. Data sering kali tersedia dalam rentang waktu yang pendek dan terbatas dalam jumlah individu atau unit. Untuk mengatasi tantangan ini, dilakukan pemanfaatan data panel karena konsepnya ialah menggabungkan data data unit atau individu dengan data rangkaian waktu. Model *Common Effect* adalah satu dari beberapa pendekatan dalam data panel yang akan diimplementasikan pada data Indeks Pembangunan Manusia di Provinsi Sumatera Utara tahun 2012-2020 yang mana parameter modelnya diestimasi dengan menggunakan metode *Ordinary Least Square*. Parameter yang dihasilkan memperlihatkan pengaruh signifikan antara variabel-variabel dependen terhadap variabel independen sebesar 91.06% serta diperoleh model sebagai berikut:  $\hat{Y}_{it} = 25.78 + 0.2178X_{1it} + 1.283X_{2it} + 0.6024X_{3it} + 0.001086 X_{4it} - 0.1043X_{5it}$

**Kata kunci:** *common effect model, ordinary least square, data panel*



**PENERAPAN METODE *ORDINARY LEAST SQUARE* (OLS) DALAM  
ESTIMASI PARAMETER MODEL REGRESI DATA PANEL *COMMON  
EFFECT* PADA DATA INDEKS PEMBANGUNAN MANUSIA (IPM)  
PROVINSI SUMATERA UTARA TAHUN 2012-2020**

**Oleh**

**SILVIA RISMA FEBIOLA MARBUN  
2017031046**

**Skripsi**

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

**Pada**

Jurusan Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Lampung



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



Judul Skripsi

: **PENERAPAN METODE *ORDINARY LEAST SQUARE* (OLS) DALAM ESTIMASI PARAMETER MODEL REGRESI DATA PANEL *COMMON EFFECT* PADA DATA INDEKS PEMBANGUNAN MANUSIA (IPM) PROVINSI SUMATERA UTARA TAHUN 2012-2020**

Nama Mahasiswa

: **Silvia Risma Febiola Marbun**

Nomor Pokok Mahasiswa

: **2017031046**

Program Studi

: **Matematika**

Fakultas

: **Matematika dan Ilmu Pengetahuan Alam**



  
**Drs. Nusyirwan, M.Si.**  
NIP. 19661010 199203 1 028

  
**Dr. Notiragayu, S.Si., M.Si.**  
NIP. 19731109 200012 2 001

**Mengetahui,**  
**Ketua Jurusan Matematika**

  
**Dr. Aang Nuryaman, S.Si., M.Si.**  
NIP. 19740316 200501 1 001



## 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: 2 April 2024



## PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Silvia Risma Febiola Marbun  
Nomor Pokok Mahasiswa : 2017031046  
Jurusan : Matematika  
Judul Skripsi : **Penerapan Metode *Ordinary Least Square* (OLS) Dalam Estimasi Parameter Model Regresi Data Panel *Common Effect* Pada Data Indeks Pembangunan Manusia (IPM) Provinsi Sumatera Utara Tahun 2012-2020**

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. Baik gagasan maupun pembahasannya adalah benar karya saya sendiri yang saya susun dengan mengikuti norma dan etika yang berlaku dan saya memastikan bahwa tingkat similaritas tidak lebih dari 25%. Apabila kemudian hari terbukti bahwa pernyataan saya tidak benar, saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku

Bandar Lampung, 2 April 2024  
Penulis.



Silvia Risma Febiola Marbun  
NPM. 2017031046



## **RIWAYAT HIDUP**

Penulis bernama Silvia Risma Febiola Marbun, anak ketiga dari Bapak Jason Marbun dan Ibu Lomoria Arithonang, lahir di Rantauprapat pada hari Sabtu, 23 Februari 2002. Penulis memiliki dua kakak perempuan yang bernama Revi Marbun dan Siska Marbun serta memiliki satu adik perempuan bernama Jesica Marbun. Penulis menyelesaikan pendidikan Sekolah Dasar di SD Negeri 112137 Rantauprapat pada tahun 2008 s.d 2014, sekolah menengah pertama di SMP Negeri 02 Rantauprapat pada tahun 2014 s.d 2017, dan sekolah menengah atas di SMA Negeri 03 Rantauprapat pada tahun 2017 s.d 2020.

Pada tahun 2020 penulis diterima sebagai mahasiswa S1 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Lampung melalui jalur SBMPTN. Selama aktif menjadi mahasiswa, penulis ikut serta dalam Unit Kegiatan Mahasiswa Sains dan Teknologi (UKM-SAINTEK) Universitas Lampung sebagai anggota Departemen Kominfo pada tahun 2021.

Pada tahun 2023, sebagai bentuk penerapan bidang ilmu di dunia kerja, penulis melaksanakan Kerja Praktek (KP) di Badan Pengelolaan Keuangan Daerah (BPKD) Jakarta Pusat dan sebagai bentuk pengabdian kepada masyarakat penulis melaksanakan Kuliah Kerja Nyata (KKN) di Desa Mekar Jaya, Kecamatan Putra Rumbia, Kabupaten Lampung Tengah.



## **KATA INSPIRASI**

“Serahkanlah perbuatanmu kepada Tuhan, maka terlaksanalah segala rencanamu”  
(Amsal 16:3)

“Sebab itu janganlah kamu kuatir akan hari besok, karena hari besok mempunyai kesusahannya sendiri, kesusahan sehari cukuplah untuk sehari”  
(Matius 6:34 TB)

“Janganlah takut, sebab Aku menyertai engkau, janganlah bimbang, sebab Aku ini Allahmu; Aku akan meneguhkan, bahkan akan menolong engkau; Aku akan memegang engkau dengan tangan kanan-Ku yang membawa kemenangan”  
(Yesaya 41:10)

“Aku berdiri karena punya mimpi. Tak akan ada keberhasilan tanpa doa dan usaha. Jangan pernah merasa lemah. Masalah adalah anugerah Tuhan dengan bentuknya yang lain. Terkadang kita diuji bukan untuk menunjukkan kelemahan kita, tetapi untuk menemukan kekuatan kita”  
(Penulis)



## **PERSEMBAHAN**

Puji dan Syukur tiada henti kepada Yesus Kristus  
atas segala Berkah, Penyertaan, Kekuatan dan Pertolongan-Nya sehingga penulis  
mampu menyelesaikan skripsi ini dengan sebaik mungkin.  
Kupersembahkan karya ini kepada:

### **Diri Sendiri**

Terima kasih atas kegigihan dan keberanian yang dimiliki dalam menghadapi  
segala tantangan dan rintangan yang telah datang. Meskipun terkadang rasanya  
berat, ingatlah bahwa setiap langkah yang dijalani telah membentuk kekuatan  
yang luar biasa. Teruslah berjuang dan percaya bahwa diri ini mampu mengatasi  
segala hal dengan baik.

### **Mama dan Keluarga**

Dengan penuh cinta dan rasa terima kasih, saya ingin mengucapkan penghargaan  
kepada Mama tercinta, Lomoria Arithonang, kedua kakak saya, Revi Marbun dan  
Siska Marbun, Bang Manik, adik saya, Jesica Marbun, serta kepada ketiga  
keponakan saya, Audrey, Jonathan, dan Eirene juga kepada Tulang Robis yang  
senantiasa selalu memberikan doa, dukungan, dan kasih sayang yang tak  
terhingga. Keberadaan dan cinta kalian adalah anugerah yang diberikan Tuhan  
dalam hidup saya.

### **Dosen**

Bapak dan Ibu Dosen pembimbing dan penguji yang sangat berjasa dalam  
membimbing dan memberikan masukan yang membangun serta menyampaikan  
banyak ilmu kepada saya.

### **Sahabat-sahabatku**

Para sahabat tersayang yang selalu mendukung, menolong, serta menemani  
penulis selama merantau dan berkuliah di Universitas Lampung.

**Almamater kebanggaan, Universitas Lampung.**



## SANWACANA

Segala puji dan syukur, penulis panjatkan kepada Yesus Kristus, atas segala berkat, penyertaan, pertolongan, serta kekuatan yang selalu diberikan-Nya sehingga penulis mampu menyelesaikan Skripsi ini yang berjudul “**Penerapan Metode *Ordinary Least Square* (OLS) Dalam Estimasi Parameter Model Regresi Data Panel *Common Effect* Pada Data Indeks Pembangunan Manusia (IPM) Provinsi Sumatera Utara Tahun 2012-2020**”

Dalam proses penyusunan dan penyelesaian Skripsi ini, Penulis banyak memperoleh bantuan baik pengajaran, bimbingan dan arahan dari berbagai pihak baik secara langsung maupun tidak langsung. Untuk itu penulis menyampaikan penghargaan dan terima kasih yang tak terhingga kepada:

1. Bapak Drs. Nusyirwan, M.Si., selaku Dosen Pembimbing I yang telah memberikan kritik, saran dan pengarahan kepada penulis dalam menyelesaikan skripsi ini.
2. Ibu Dr. Notiragayu, S.Si., M.Si., selaku Dosen Pembimbing II yang telah meluangkan waktu ditengah kesibukan beliau untuk memberikan bimbingan, kritik, saran dan pengarahan dalam menyelesaikan skripsi ini.
3. Ibu Dr. Khoirin Nisa, S.Si., M.Si., selaku dosen Pembahas skripsi yang telah memberikan evaluasi dan saran bagi perbaikan skripsi Penulis.
4. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
5. Bapak Dr. Ahmad Faisol, S.Si., M.Sc., selaku dosen pembimbing akademik.
6. Bapak Dr. Eng. Heri Satria, S.Si., M.Si., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.



7. Seluruh Dosen, Staff, dan Karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Kedua Orang Tua terkhusus Mama, Kak Revi dan Bang Manik, Kak Siska, Jesica, Keponakan tersayang Audrey, Jonathan, dan Eirene serta Tulang Robis yang telah memotivasi penulis untuk terus berjuang menyelesaikan Laporan ini. Terima Kasih atas segala doa, nasihat, dukungan, semangat dan perjuangan yang besar untuk penulis.
9. Sahabat-sahabat Penulis, Mawar, Nadia, Ocha, Rini, Devanisa dan Ziyad yang telah membantu, memotivasi, memberikan nasihat, dan menjadi tempat keluh kesah penulis selama mengerjakan skripsi ini hingga selesai.
10. Seluruh teman-teman Matematika angkatan 2020 yang telah membantu penulis untuk menyelesaikan laporan ini.
11. Almamater Tercinta, Universitas Lampung
12. Seluruh pihak yang telah membantu dan terlibat dalam menyelesaikan skripsi ini.

Sebagai manusia biasa, penulis menyadari penyusunan skripsi ini jauh dari kata sempurna dan masih banyak kekurangan baik dalam penyajian maupun penulisan. Oleh sebab itu, penulis memohon maaf dan menerima kritik serta saran yang dapat membangun agar menjadi pelajaran serta perbaikan pada skripsi ini untuk kedepannya.

Bandar Lampung, 2 April 2024  
Penulis

**Silvia Risma Febiola Marbun**  
**NPM. 2017031046**



## DAFTAR ISI

	Halaman
<b>ABSTRAK</b> .....	iii
<b>LEMBAR PERSETUJUAN</b> .....	iv
<b>LEMBAR PENGESAHAN</b> .....	v
<b>LEMBAR PERNYATAAN</b> .....	vi
<b>RIWAYAT HIDUP</b> .....	vii
<b>KATA INSPIRASI</b> .....	viii
<b>SANWACANA</b> .....	ix
<b>DAFTAR ISI</b> .....	xi
<b>DAFTAR TABEL</b> .....	xiii
<b>I. PENDAHULUAN</b> .....	1
1.1 Latar Belakang dan Masalah .....	1
1.2 Tujuan .....	3
1.3 Manfaat .....	4
<b>II. TINJAUAN PUSTAKA</b> .....	5
2.1 Analisis Regresi .....	5
2.1.1 Analisis Regresi Sederhana .....	5
2.1.2 Analisis Regresi Berganda .....	6
2.2 Data Panel .....	6
2.3 Model Regresi Data Panel .....	8
2.3.1 Model <i>Common Effect</i> (CEM) .....	9
2.3.2 Model <i>Fixed Effect</i> (FEM) .....	10
2.3.3 Model <i>Random Effect</i> (REM) .....	11
2.4 Uji Spesifikasi Model .....	12
2.4.1 Uji <i>Chow</i> .....	12
2.4.2 Uji <i>Hausman</i> .....	12
2.4.3 Uji <i>Lagrange Multiplier</i> .....	13
2.5 Estimasi Parameter .....	14
2.6 Metode <i>Ordinary Least Square</i> (OLS) .....	14
2.7 Sifat-sifat BLUE .....	18



2.8 Uji Asumsi Model .....	18
2.8.1 Uji Normalitas .....	19
2.8.2 Uji Autokorelasi .....	20
2.8.3 Uji <i>Heterokedastisitas</i> .....	20
2.8.4 Uji Multikolinearitas .....	21
2.9 Uji Statistik Model .....	22
2.9.1 Uji F .....	22
2.9.2 Uji t .....	23
2.9.3 Koefisien Determinasi .....	24
2.10 Indeks Pembangunan Manusia .....	24
2.11 Persentase Penduduk Miskin .....	26
<b>III. METODOLOGI PENELITIAN .....</b>	<b>27</b>
3.1 Waktu dan Tempat Penelitian .....	27
3.2 Data Penelitian .....	27
3.3 Metode Penelitian .....	27
<b>IV. HASIL DAN PEMBAHASAN .....</b>	<b>29</b>
4.1 Analisis Deskriptif .....	29
4.2 Estimasi Parameter $\beta$ Menggunakan <i>Ordinary Least Square</i> .....	30
4.3 Analisis Model Regresi Data Panel .....	33
4.3.1 Model <i>Common Effect</i> (CEM) .....	33
4.3.2 Model <i>Fixed Effect</i> (FEM) .....	34
4.3.3 Model <i>Random Effect</i> (REM) .....	36
4.4 Pemilihan Model Regresi Data Panel .....	37
4.4.1 Uji <i>Chow</i> .....	37
4.5 Penerapan Model Regresi Data Panel <i>Common Effect</i> .....	38
4.5.1 Estimasi Parameter Model <i>Common Effect</i> .....	38
4.5.2 Asumsi Gauss-Markov .....	40
4.6 Uji Asumsi Model .....	42
4.6.1 Uji Normalitas Residual .....	42
4.6.2 Uji <i>Heteroskedastisitas</i> .....	42
4.6.3 Uji Multikolinearitas .....	43
4.7 Uji Statistik Model .....	43
4.7.1 Uji F .....	43
4.7.2 Uji t .....	44
4.7.3 Koefisien Determinasi .....	45
<b>V. KESIMPULAN .....</b>	<b>46</b>
<b>DAFTAR PUSTAKA .....</b>	<b>47</b>
<b>LAMPIRAN .....</b>	<b>51</b>



## DAFTAR TABEL

Tabel	Halaman
1. Susunan Data <i>Cross section</i> .....	7
2. Susunan Data <i>Time series</i> .....	7
3. Susunan Data Panel.....	8
4. Statistika deskriptif data penelitian .....	29
5. Hasil estimasi CEM .....	33
6. Hasil estimasi FEM.....	34
7. Efek individu .....	34
8. Hasil estimasi REM .....	36
9. Hasil uji <i>Chow</i> .....	37
10. Hasil analisis <i>common effect</i> .....	39
11. Nilai prediksi dan <i>error</i> .....	40
12. Uji <i>Kolmogorov-Smirnov</i> .....	42
13. Uji Breusch-Pagan .....	42
14. Nilai VIF .....	43
15. Hasil uji F.....	44
16. Hasil uji t.....	44
17. Nilai koefisien determinasi .....	45



## I. PENDAHULUAN

### 1.1 Latar Belakang dan Masalah

Analisis regresi ialah alat utama dan penting dalam ekonometrika yang digunakan untuk menemukan dan mengevaluasi korelasi antara berbagai variabel. Teknik ini membantu dalam memahami pola hubungan dari satu atau beberapa variabel prediktor dengan variabel respons. Dengan demikian, analisis regresi memberikan kerangka kerja matematis untuk menjelaskan dan memodelkan bagaimana perubahan dalam satu variabel dapat memengaruhi variabel yang lain. Dikatakan, baik dalam penelitian maupun riset, terkadang akan ditemui suatu permasalahan mengenai terbatasnya data dalam merepresentasikan sebuah variabel. Pada contohnya, data dapat ditemukan dalam rangkaian waktu yang singkat. Akibatnya, jumlah data yang diperlukan sangat sedikit, sehingga proses pengolahan pada rangkaian waktu atau runtun waktu sulit maupun tidak dapat dilakukan. Selain masalah tersebut, bentuk data dapat memiliki jumlah individu atau unit yang terbatas, yang membuat proses pengolahan pada rangkaian data sulit dipahami dan diperoleh informasi yang relevan. Untuk mengatasi kedua kondisi tersebut, digunakan data panel.

Data panel ialah penggabungan data rangkaian waktu (*time series*) dengan data silang (*cross section*). Keunggulannya menurut Gujarati & Porter (2009), (1) dapat secara *explicit* mengendalikan dan mengontrol perbedaan individu sehingga memungkinkan pembentukan model yang lebih kompleks, (2) dapat menyediakan tambahan informasi yang tidak bisa diberikan oleh metode konvensional, seperti hanya data unit/individu (saja atau data rangkaian waktu



saja, selain itu (3) menghasilkan estimasi yang lebih baik, memberikan tambahan informasi, meningkatkan variabilitas, mengurangi korelasi antar variabel penjelas, dan memberikan lebih banyak derajat kebebasan.

Tiga jenis metode dalam estimasi model regresi data panel yang banyak ditawarkan, ialah model *common effect*, *fixed effect*, dan *random effect*. Namun, pada penelitian yang akan dilakukan dan dibahas nanti, peneliti memilih untuk menggunakan model *common effect* dimana menganggap koefisien regresi dan intersep konstan sepanjang waktu dan individu,

Merujuk pada penelitian sebelumnya, Aulina & Mirtawati (2021), mengenai model regresi yang estimasi parameternya dilakukan dengan *ordinary least square* melalui pendekatan *fixed effect* pada Data Kemiskinan di Indonesia Tahun 2015-2019. Dilakukan melalui variabel *dummy* untuk menentukan perbedaan *intercpt* di setiap provinsi, yang menjelaskan dampak dari perbedaan wilayah. Selama periode 2015 hingga 2019, diperoleh bahwa tingkat kemiskinan di Indonesia mengalami pengaruh negatif dan signifikan dari variabel pertumbuhan ekonomi (PDRB). Sebaliknya, variabel tingkat pengangguran terbuka berpengaruh secara positif dan signifikan. Selanjutnya, penelitian oleh Hutagalung & Darnius (2022), mengenai pemilihan model data panel terbaik dan aspek-aspek utama yang memiliki pengaruh terhadap Indeks Pembangunan Manusia (IPM) di Sumatera Utara. Berdasarkan penelitian tersebut, ditemukan bahwa *Common Effect Model* merupakan model terbaik dan variabel-variabelnya memiliki pengaruh yang signifikan terhadap IPM di Sumatera Utara.

*Ordinary Least Square* (OLS) merupakan salah satu metode dalam estimasi parameter. Ketika mengolah data menggunakan metode statistika, tidak ada kepastian mutlak, tetapi merupakan hasil perkiraan yang memperhitungkan variasi dalam fenomena tertentu. Oleh karena itu, perlu dilakukan estimasi. Dalam hal ini, peneliti menjalankan proses estimasi parameter dengan tujuan memahami



karakteristik parameter suatu populasi dan memperoleh penduga parameter yang tidak bias.

Dalam mengevaluasi tingkat kemajuan pembangunan manusia di suatu daerah, digunakan suatu instrumen, yakni Indeks Pembangunan Manusia. Instrumen ini diperhitungkan berdasarkan tiga faktor utama, yakni harapan hidup, taraf hidup yang memadai, dan tingkat pendidikan. Tiap-tiap faktor ini direpresentasikan oleh dimensi tertentu; harapan hidup mewakili usia rata-rata penduduk, taraf hidup yang memadai direpresentasikan oleh pengeluaran per kapita, dan tingkat pendidikan diwakili dengan rata-rata tahun sekolah yang ditempuh dan angka harapan lama sekolah. Pembangunan manusia di suatu wilayah umumnya diklasifikasikan dalam empat kelompok menurut capaiannya, yaitu dikatakan sangat tinggi jika  $IPM \geq 80$ , tinggi jika  $70 \leq IPM < 80$ , sedang jika  $60 \leq IPM < 70$  dan rendah jika  $IPM < 60$  (BPS, 2022).

Berdasarkan penjelasan sebelumnya, peneliti akan mengimplementasikan *Ordinary Least Square* untuk mencari estimasi parameter model *common effect* pada data IPM di Provinsi Sumatera Utara dalam rentang waktu delapan tahun, dari tahun 2012 hingga 2020 berdasarkan faktor yang mempengaruhinya.

## 1.2 Tujuan

1. Mengestimasi parameter model data panel melalui metode *Ordinary Least Square*.
2. Menentukan model data panel *common effect* pada data Indeks Pembangunan Manusia di Provinsi Sumatera Utara pada tahun 2012-2020.



### 1.3 Manfaat

1. Mendapatkan nilai parameter model regresi yang diestimasi dengan *Ordinary Least Square*.
2. Mendapatkan model *common effect* pada data Indeks Pembangunan Manusia di Provinsi Sumatera Utara pada tahun 2012-2020.
3. Memperoleh pemahaman terkait model *common effect* yang parameteranya diestimasi menggunakan *Ordinary Least Square*.



## II. TINJAUAN PUSTAKA

### 2.1 Analisis Regresi

Tujuan analisis regresi ialah mengetahui bagaimana variabel tertentu berpengaruh satu sama lain. Francis Galton, seorang antropolog dan ahli meteorologi, pertama kali menggunakan istilah regresi dalam artikelnya pada tahun 1886 “*Family Likeness in Stature*” (Basuki & Prawoto, 2017). Analisis regresi dapat digunakan untuk banyak hal, seperti membentuk model hubungan numerik, menjelaskan fenomena data, mengendalikan kasus atau hal yang diamati dengan menggunakan model yang dihasilkannya, dan memprediksi variabel terikat (Basuki & Prawoto, 2017). Terdapat dua bentuk fungsi regresi, regresi non-linear dan regresi linear. Analisis regresi linear sederhana dan berganda adalah dua jenis analisis dalam regresi linear. Menurut Sunyoto (2020), dikatakan linear karena setiap estimasi nilai diharapkan meningkat atau menurun sesuai dengan garis lurus.

#### 2.1.1 Analisis Regresi Sederhana

Analisis ini mengacu pada keterkaitan fungsional linear antara satu variabel prediktor ( $X$ ) dengan satu variabel respon ( $Y$ ) (Sulistiyowati & Astuti, 2017).

Model umum regresi linear sederhana adalah:

$$y_i = \beta_0 + \beta_1 X_1 + \varepsilon_i \quad (2.1)$$



keterangan:

$y_i$  = variabel dependen (nilai yang diprediksikan)

$X_1$  = variabel independen

$\beta_0$  = parameter konstanta atau *intercept*

$\beta_1$  = parameter koefisien regresi

$\varepsilon$  = variabel galat/kesalahan regresi (*error term*), dengan  $\varepsilon \sim N(0; \sigma^2)$

$i = 1, 2, 3, \dots, n$

### 2.1.2 Analisis Regresi Berganda

Menurut Sulistiyowati & Astuti (2017), regresi ini digunakan dalam menentukan keterkaitan fungsional linear diantara dua atau beberapa variabel prediktor ( $X$ ) dengan satu variabel respons ( $Y$ ). Dengan bentuk model umumnya adalah:

$$y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K + \varepsilon_i \quad (2.2)$$

keterangan:

$y_i$  = variabel dependen (nilai yang diprediksikan)

$X_1, X_2, \dots, X_K$  = variabel independen

$\beta_0$  = parameter konstanta atau *intercept*

$\beta_1, \dots, \beta_K$  = parameter koefisien regresi

$\varepsilon$  = variabel galat/kesalahan regresi (*error term*), dengan  $\varepsilon \sim N(0; \sigma^2)$

$k$  = banyak variabel bebas

$i = 1, 2, 3, \dots, n$

### 2.2 Data Panel

Data Panel menurut Gujarati & Porter (2009), kombinasi dari data rangkaian waktu dan data silang atau unit. *Time series* merupakan data dari sejumlah periode, sementara *cross section* merupakan data dari sejumlah objek (Winarno,



2007). Data panel memiliki tujuan untuk mengetahui aspek-aspek penting yang didasarkan oleh pengamatan yang dilakukan berulang-ulang pada satu atau lebih objek pada berbagai waktu (Aulina & Mirtawati, 2021). Memberikan sejumlah pengamatan yang besar, meningkatkan derajat kebebasan (tingkat kebebasan), dan mampu mengurangi kolinieritas antara variabel penjela adalah beberapa keunggulan penggunaan data panel. Data panel juga memiliki variabilitas yang besar dan dapat memberikan informasi tambahan serta dapat memberikan inferensi perubahan dinamis yang lebih baik daripada data *cross-section* (Hsiao, 2014). Dengan regresi data panel, kita dapat mempelajari model yang lebih kompleks (Chairunnisa & Fauzan, 2023). Susunan dari data panel dibentuk melalui tabel:

Tabel 1. Susunan Data *Cross Section*

Individu ( $i$ )	Variabel Respons ( $Y$ )	Variabel Prediktor ( $X_1$ )	Variabel Prediktor ( $X_2$ )	...	Variabel Prediktor ( $X_j$ )
1	...	...	...	...	...
2	...	...	...	...	...
$\vdots$	...	...	...	...	...
N	...	...	...	...	...

Tabel 2. Susunan Data *Time Series*

Waktu ( $t$ )	Variabel Respons ( $Y$ )	Variabel Prediktor ( $X_1$ )	Variabel Prediktor ( $X_2$ )	...	Variabel Prediktor ( $X_j$ )
1	...	...	...	...	...
2	...	...	...	...	...
$\vdots$	...	...	...	...	...
T	...	...	...	...	...



Tabel 3. Susunan Data Panel

Individu ( <i>i</i> )	Waktu ( <i>t</i> )	Variabel Respons ( <i>Y</i> )	Variabel Prediktor ( <i>X</i> <sub>1</sub> )	Variabel Prediktor ( <i>X</i> <sub>2</sub> )	...	Variabel Prediktor ( <i>X</i> <sub><i>j</i></sub> )
1	1	...	...	...	...	...
1	2	...	...	...	...	...
⋮	⋮	...	...	...	...	...
1	T	...	...	...	...	...
2	1	...	...	...	...	...
2	2	...	...	...	...	...
⋮	⋮	...	...	...	...	...
2	T	...	...	...	...	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮
N	1	...	...	...	...	...
N	2	...	...	...	...	...
⋮	⋮	...	...	...	...	...
N	T	...	...	...	...	...

### 2.3 Model Regresi Data Panel

Analisis data yang mencatat kejadian berulang pada unit yang sama selama beberapa waktu dapat dilakukan dengan menggunakan model regresi data panel dengan persamaan umumnya adalah:

$$y_{it} = \beta_{0it} + \sum_{j=1}^K \beta_{jit} X_{jit} + \varepsilon_{it} \quad (2.3)$$

keterangan:

$y_{it}$  = variabel terikat untuk unit individu ke-*i* dan waktu ke-*t*

$X_{jit}$  = variabel bebas ke-*k* untuk unit individu ke-*i* dan waktu ke-*t*

$\beta$  = parameter yang tidak diketahui atau koefisien variabel bebas



$\varepsilon_{it}$  = error untuk individu ke- $i$  dan waktu ke- $t$

$i = 1, 2, \dots, n$  untuk unit individu

$t = 1, 2, \dots, t$  untuk waktu

Semua variabel bebas tidak stokastiks serta *error term* mengikuti asumsi klasik, yaitu berdistribusi normal  $\varepsilon_{it} \sim N(0; \sigma^2)$  (Hill, *et al.*, 2011). Menurut Gujarati & Porter (2009), kemungkinan-kemungkinan yang dapat terjadi antar intersep, koefisien *slope* dan *error term* adalah:

1. Koefisien *slope* dan *intercept* tetap tidak berubah searing waktu dan antar individu, dengan *error* bervariasi sepanjang waktu dan antar individu.
2. Koefisien *slope* tetap, namun *intercept* beragam antar individu.
3. Koefisien *slope* tetap, namun *intercept* beragam antar individu serta waktu.
4. *Intercept* dan koefisien *slope* bervariasi sepanjang individu.
5. *Intercept* dan koefisien *slope* beragam seiring waktu dan antar individu.

Hal-hal yang disebutkan di atas menampilkan jika semakin banyak variabel yang dijelaskan, estimasi parameternya akan semakin kompleks.

### 2.3.1 Model *Common Effect* (CEM)

*Common Effect Model*, dikenal model efek umum, mengasumsikan bahwa karakteristik individu atau objek konsisten sepanjang waktu sehingga *intercept* dan koefisien *slope* yang tetap tidak berubah, serta terdapat *error term* untuk menjelaskan variasi antar individu (Basuki & Prawoto, 2017). Model ini memanfaatkan metode *Ordinary Least Square* untuk memperkirakan parameter-parameternya dengan menggabungkan (*pooling*) keseluruhan set data individu dan waktu. Secara umum, persamaan model ini adalah:

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

keterangan:

$i$  = unit individu sebanyak  $n$



$t$  = unit waktu sebanyak  $t$

$k$  = jumlah variabel prediktor sebanyak  $p$

$Y_{it}$  = variabel respon individu ke-  $i$  pada waktu ke- $t$

$X_{kit}$  = variabel prediktor ke-  $k$  pada waktu ke- $t$  untuk individu ke-  $i$

$\beta_0$  = *intercept* model

$\beta_k$  = koefisien *slope*

$\varepsilon_{it}$  = *error* untuk waktu ke- $t$  dan unit ke-  $i$ ,  $\varepsilon_{it} \sim N(0; \sigma^2)$

### 2.3.2 Mode *Fixed Effect* (FEM)

Model *fixed effect*, atau model efek tetap, digunakan sebagai solusi dalam mengatasi kelemahan analisis yang timbul dari metode *common effect* karena ketidakmampuan *common effect* untuk mengubah *intercept*, baik pada tingkat individu maupun waktu. Oleh karena itu, model ini akan mengatasi masalah tersebut dengan memberikan perbedaan konstanta antara objek meskipun koefisien *regresor*-nya sama (Alviani, 2021). Diasumsikan dalam model ini bahwa variasi efek antar unit atau individu dapat dijelaskan melalui perbedaan dalam *intercept*-nya (Alamsyah, dkk., 2022). Melalui estimasi yang melibatkan variabel *dummy* untuk menangkap variasi *intercept*-nya (Widarjono, 2007). Persamaan model ini secara umum adalah:

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

keterangan:

$i$  = unit individ sebanyak  $n$

$t$  = unit waktu sebanyak  $t$

$k$  = jumlah variabel prediktor sebanyak  $p$

$Y_{it}$  = variabel respon individu ke-  $i$  pada waktu ke- $t$

$X_{kit}$  = variabel prediktor ke-  $k$  untuk individu ke-  $i$  pada waktu ke- $t$

$\beta_0$  = *intercept* model

$\beta_k$  = koefisien *slope*



$\varepsilon_{it}$  = nilai *error* untuk unit ke-  $i$  dan waktu ke- $t$

### 2.3.3 Model *Random Effect* (REM)

Estimasi pada model ini dilakukan dimana variabel gangguan dapat saling berkorelasi baik pada tingkat individu maupun waktu (Alamsyah, dkk., 2022). Model ini tidak secara langsung memperhitungkan variasi intersep dalam estimasinya, tetapi mengambil adanya efek acak yang berasal dari variabel yang tidak diamati yang memengaruhi variabel tersebut dan dengan manfaat mengatasi *heteroskedastisitas*. Model REM juga dikenal sebagai *Error Component Model* (ECM) atau teknik *Generalized Least Square* (GLS) (Basuki & Prawoto, 2017). Secara umum, persamaan model ini adalah:

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

keterangan:

$i$  = unit individu sebanyak  $n$

$t$  = unit waktu sebanyak  $t$

$k$  = jumlah variabel prediktor sebanyak  $p$

$Y_{it}$  = variabel respon individu ke-  $i$  pada waktu ke- $t$

$X_{kit}$  = variabel prediktor ke-  $k$  untuk individu ke-  $i$  pada waktu ke- $t$

$\beta_0$  = *intercept* model

$\beta_k$  = koefisien *slope*

$\mu_i$  = nilai *error* pada uniy observasi ke-  $i$

$\varepsilon_{it}$  = nilai *error* untuk unit ke-  $i$  dan waktu ke- $t$



## 2.4 Uji Spesifikasi Model

### 2.4.1 Uji Chow

Uji ini dilakukan dalam melakukan pemilihan model regresi data panel khususnya antara model *fixed effect* dan model *common effect*. Menggunakan hipotesis:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_N = \beta \text{ (Model Common Effect)}$$

$$H_1 : \text{setidaknya terdapat satu } \beta_i \text{ yang berbeda (Model Fixed Effect)}$$

Dengan statistik uji:

$$F = \frac{\frac{RSS - URSS}{N-1}}{\frac{URSS}{NT-N-K}} \quad (2.7)$$

dimana:

$RSS$  = sum square residual model OLS

$URSS$  = sum square residual model fix

$N$  = jumlah unit *cross section*

$T$  = jumlah unit waktu

$K$  = jumlah parameter yang akan diestimasi

Apabila  $|F| > F_{0.05(N-1, NT-N-K)}$ , maka  $H_0$  ditolak, artinya *intercept* untuk semua individu tidak sama dan model yang terpilih adalah *fixed effect*, jika sebaliknya, maka  $H_0$  diterima, artinya *intercept* untuk semua individu sama dan model yang terpilih ialah *common effect* (Ghozi & Hermansyah, 2018).

### 2.4.2 Uji Hausman

Salah satu kegunaan dari uji ini ialah memilih antara model *fixed effect* dan model *random effect*, dan juga untuk menentukan apakah ada hubungan antara galat



dalam model dengan satu atau lebih variabel penjelas (Madany, dkk., 2022).

Menggunakan hipotesis:

$$H_0 : \text{corr}(x_{it}, \varepsilon_i) = 0 \text{ (Model Random Effect)}$$

$$H_1 : \text{corr}(x_{it}, \varepsilon_i) \neq 0 \text{ (Model Fixed Effect Model)}$$

Statistik *Hausman* akan mengikuti distribusi *Chi-Square* dengan menggunakan kriteria Wald.

$$W = \chi^2_{(P)} = [\hat{\beta}_{FEM} - \hat{\beta}_{REM}]' \hat{\psi}^{-1} [b - \beta] \quad (2.8)$$

dimana

$$\hat{\psi} = \text{Var}[\hat{\beta}_{FEM}] - \text{Var}[\hat{\beta}_{REM}] \quad (2.9)$$

keterangan:

$$\hat{\beta}_{FEM} = \text{Koefisien fixed effect}$$

$$\hat{\beta}_{REM} = \text{Koefisien random effect}$$

Jika nilai statistik *Hausman* yang diperoleh lebih besar dari nilai kritisnya,  $H_0$  ditolak, artinya model yang terpilih ialah *fixed effect*. Namun, jika nilai statistik yang diperoleh lebih rendah dari nilai kritisnya, maka model yang dipilih ialah *random effect* (Madany, dkk., 2022).

#### 2.4.3 Uji Lagrange Multiplier

Uji ini dilakukan ketika hasil pada uji *Hausman* menerima hipotesis nol.

Tujuannya adalah untuk mencari tahu antara model dengan pendekatan *random effect* jauh lebih baik daripada model *common effect*. Menggunakan hipotesis berikut:

$$H_0 : \sigma_1^2 = 0 \text{ (Model mengikuti common effect)}$$

$$H_1 : \sigma_1^2 \neq 0 \text{ (Model mengikuti random effect)}$$



Dengan statistik uji:

$$LM = \frac{NT}{2(T-1)} \left[ \frac{\sum_{i=2}^K [\sum_{t=1}^T e_{it}]^2}{\sum_{i=2}^K \sum_{t=1}^T e_{it}^2} - 1 \right]^2 \quad (2.10)$$

Jika  $LM > X_{(K;0.05)}^2$  ( $K$  = jumlah variabel), maka tolak  $H_0$  dan *random effect* ialah model terpilih, jika sebaliknya, model yang lebih tepat adalah *common effect* (Greene, 2011).

## 2.5 Estimasi Parameter

Dengan menggunakan data sampel, tujuan analisis regresi adalah untuk menghasilkan estimasi parameter model, yang menunjukkan bagaimana variabel independen mempengaruhi variabel dependen. Estimasi parameter merujuk pada metode statistika untuk memperkirakan nilai parameter populasi menggunakan statistik sampel, dapat berupa estimasi titik atau interval (Nuryadi, dkk., 2018). Memiliki hubungan linear, *residual* yang dihasilkan dari model identik (*homoskedastik*) dan independen antara variabel independen dan dependen memungkinkan pencapaian tujuan dengan akurat dan tepat. Jika terpenuhi, maka teorema Gaus-Markov menjamin bahwa estimasi parameter yang dihasilkan adalah *Best Linear Unbiased Estimator* (BLUE) (Best & Wolf, 2015).

## 2.6 Metode *Ordinary Least Square* (OLS)

*Ordinary Least Square* juga biasa disebut Metode Kuadrat Terkecil ialah pendekatan yang dapat diterapkan dalam mengestimasi koefisien regresi dengan meminimalkan jumlah kuadrat galat (Nurhalizah & Sitompul, 2022). Agar taksiran koefisien regresi itu bersifat BLUE, beberapa asumsi dasar digunakan untuk memastikan estimator tidak bias linear (*unbiased*) terbaik dari model yang didapat. Jika estimator tidak cukup efisien, prediksi regresi akan memberikan



hasil yang tidak efisien. Karena ketidakcocokan dalam matriks kovarians dari regresi yang diperkirakan koefisien, uji hipotesis seperti uji-t dan uji-F tidak lagi layak (Petronilla, 2022). Dimisalkan model statistik linear dengan data sebanyak  $n$ :

$$y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i \quad (2.11)$$

Model tersebut dapat diuraikan ke dalam bentuk matriks sebagai berikut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2.12)$$

Sehingga secara sederhana menjadi:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.13)$$

Variabel  $\boldsymbol{\varepsilon}$  tidak dapat diobservasi serta tidak ada informasi mengenai kemungkinan distribusinya. Untuk variabel  $\boldsymbol{\varepsilon}$ , Gauss membuat asumsi sebagai berikut menggunakan model regresi yang disebutkan sebelumnya:

1. *Expected value* (nilai harapan) dari variabel  $\boldsymbol{\varepsilon}$  memiliki rata-rata 0, atau

$$E(\boldsymbol{\varepsilon}) = 0 \quad (2.14)$$

Hal ini menandakan jika tergantung pada nilai  $x$ , nilai kondisional  $\varepsilon$  yang diharapkan adalah nol. Sehingga, untuk banyak nilai  $x$ , nilai harapan  $\varepsilon$  diharapkan menjadi nol.

2. Tidak ditemukan korelasi serial antar variabel dalam setiap observasi, menunjukkan  $\boldsymbol{\varepsilon}_i$  dan  $\boldsymbol{\varepsilon}_j$  tidak memiliki hubungan baik positif maupun negatif.

Selain itu, tidak ada *heteroskedastisitas* antar variabel  $\boldsymbol{\varepsilon}$ , sehingga syarat *homoskedastisitas* terpenuhi. Dengan demikian, variabel  $\boldsymbol{\varepsilon}$  memiliki varian yang konstan dan positif dan nilainya  $\sigma^2$ , yaitu:

$$Var(\varepsilon_i) = \sigma^2 \quad (2.15)$$

$$Cov(\varepsilon_i, \varepsilon_j) = 0 \quad (2.16)$$

Persamaan (2.15) dan (2.16) dapat dibentuk ke dalam matriks menjadi:

$$\begin{bmatrix} var(\varepsilon_1) & cov(\varepsilon_1, \varepsilon_2) & \cdots & cov(\varepsilon_1, \varepsilon_n) \\ cov(\varepsilon_2, \varepsilon_1) & var(\varepsilon_2) & \cdots & cov(\varepsilon_2, \varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(\varepsilon_n, \varepsilon_1) & cov(\varepsilon_n, \varepsilon_2) & \cdots & var(\varepsilon_n) \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}$$



Sehingga asumsi kedua ini ditulis sebagai berikut:

$$\begin{aligned}
 Cov(\boldsymbol{\varepsilon}) &= E[(\boldsymbol{\varepsilon} - E(\boldsymbol{\varepsilon}))(\boldsymbol{\varepsilon} - E(\boldsymbol{\varepsilon}))'] \\
 &= E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') \\
 &= \sigma^2 I
 \end{aligned} \tag{2.17}$$

3. Variabel  $x_i$  dan  $\varepsilon_i$  adalah independen satu sama lain dalam setiap observasi sehingga:

$$\begin{aligned}
 Cov(x_i, \varepsilon_i) &= E[(x_i - E(x_i))(\varepsilon_i - E(\varepsilon_i))] \\
 &= E[(x_i - \bar{x})(\varepsilon_i - 0)] \\
 &= E[(x_i - \bar{x})\varepsilon_i] \\
 &= (x_i - \bar{x})E(\varepsilon_i) \\
 &= 0
 \end{aligned} \tag{2.18}$$

dan diperoleh dari ketiga asumsi ini adalah:

$$\begin{aligned}
 E(\mathbf{y}) &= E(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\
 &= E(\mathbf{X}\boldsymbol{\beta}) + E(\boldsymbol{\varepsilon}) \\
 &= \mathbf{X}\boldsymbol{\beta} + \mathbf{0} \\
 &= \mathbf{X}\boldsymbol{\beta}
 \end{aligned} \tag{2.19}$$

sehingga,

$$\begin{aligned}
 E(\mathbf{y}) &= \mathbf{X}\boldsymbol{\beta} \\
 Cov(\mathbf{y}) &= \sigma^2 I
 \end{aligned} \tag{2.20}$$

4. Dalam parameter, hubungan antara  $\mathbf{Y}$  dan  $\mathbf{X}$  adalah linier.
5. Variabel  $\mathbf{X}$  merupakan variabel tak stokastik yang nilainya tetap. Untuk sejumlah observasi yang berbeda, nilai  $\mathbf{X}$  tidak berubah. Nilai variabel independen adalah tetap dalam sampel berulang atau variabel independen adalah variabel yang dikontrol.

Misalkan diberikan sampel  $\mathbf{y}$ . Dengan menggunakan sampel ini untuk dalam mendapatkan estimasi  $\boldsymbol{\beta}$  adalah untuk meminimalkan  $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ . Ini bertujuan untuk menghasilkan komponen sistematis yang lebih signifikan daripada komponen stokastiknya (Wasiaturrahma & Rohmawati, 2021). Jika peran komponen stokastik meningkat, ini mengindikasikan bahwa informasi mengenai  $\mathbf{y}$  yang diperoleh menjadi terbatas, dan  $\mathbf{X}$  tidak mampu menjelaskan  $\mathbf{y}$ . Oleh karena itu, diperlukan pemilihan parameter  $\boldsymbol{\beta}$  sedemikian rupa sehingga:



$$JKG = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (2.21)$$

seminimal mungkin. Karena persamaan (2.21) skalar, maka komponennya juga skalar. Oleh karena itu, nilai skalar tidak diubah oleh transpose skalar.  $JKG$  ditulis sebagai berikut:

$$\begin{aligned} JKG &= \sum_{i=1}^n \varepsilon_i^2 \\ &= \varepsilon_1^2 + \varepsilon_2^2 + \cdots + \varepsilon_n^2 \\ &= [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n] \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \\ &= \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \\ &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= (\mathbf{y}^T - (\mathbf{X}\boldsymbol{\beta})^T) (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= (\mathbf{y}^T - \boldsymbol{\beta}^T \mathbf{X}^T) (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{y}^T \mathbf{y} - \mathbf{y}^T \mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.22)$$

karena  $\mathbf{y}^T \mathbf{X}\boldsymbol{\beta}$  merupakan skalar, maka matriks transpose-nya adalah:

$$\begin{aligned} (\mathbf{y}^T \mathbf{X}\boldsymbol{\beta})^T &= \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} \\ JKG &= \mathbf{y}^T \mathbf{y} - 2 \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.23)$$

Selanjutnya, turunan pertama  $JKG$  terhadap  $\boldsymbol{\beta}$  dicari dan disamakan dengan nol untuk meminimumkannya:

$$\begin{aligned} \frac{dJKG}{d\boldsymbol{\beta}} &= \frac{d[(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})]}{d\boldsymbol{\beta}} \\ &= -2\mathbf{X}^T \mathbf{Y} + 2\mathbf{X}^T \mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.24)$$

disamakan dengan nol maka akan diperoleh

$$\mathbf{X}^T \mathbf{Y} = \mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}}_{OLS} \quad (2.25)$$

atau

$$\hat{\boldsymbol{\beta}}_{OLS} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (2.26)$$

yang disebut *estimator* (penaksir) parameter  $\boldsymbol{\beta}$  secara *Ordinary Least Square* (Aziz, 2010).



## 2.7 Sifat-sifat BLUE

Estimator OLS dikatakan BLUE apabila memenuhi sifat-sifat ideal atau optimum yang terdapat dalam teorema Gauss-Markov sebagai berikut:

1. Merupakan fungsi linear dari sebuah variabel acak. Pada persamaan (2.26) karena  $(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$  merupakan matriks bernilai tetap, maka  $\hat{\boldsymbol{\beta}}_{OLS}$  fungsi linear dari  $\mathbf{y}$ .
2. Tidak bias dimana nilai ekspektasinya,  $E(\hat{\boldsymbol{\beta}})$  sama dengan nilai sebenarnya  $\boldsymbol{\beta}$ .
3. Estimator dikatakan efisien jika mempunyai varians terkecil dari semua kelompok estimator tidak bias dan linear (Gujarati & Porter, 2009).

## 2.8 Uji Asumsi Model

Terdapat asumsi-asumsi model yang wajib terpenuhi sehingga dapat memenuhi aturan BLUE pada estimasi model yang dihasilkan, yaitu asumsi normalitas, autokorelasi, *heteroskedastisita* dan multikolinearitas (Wardhana & Indawati, 2021). Namun, menurut Basuki & Prawoto (2017), dalam uji asumsi model data panel, beberapa tidak harus dipenuhi dimana:

1. Karena model diasumsikan bersifat linear, uji linearitas nyaris tidak dilakukan pada setiap model.
2. Uji normalitas karena dasarnya bukan syarat BLUE.
3. Ketika regresi menggunakan variabel bebas lebih dari satu, uji multikolinearitas harus dilakukan.
4. Uji *heteroskedastisitas* harus dilakukan dalam memastikan asumsi homoskedastisitas terpenuhi dan memastikan estimasi parameter model tidak bias.
5. Uji autokorelasi hanya relevan untuk data *sequence time*. Melakukan uji ini pada data *sequence non-time*, seperti *cross section* atau panel, tidak bermakna atau tidak berarti.



### 2.8.1 Uji Normalitas

Uji ini bertujuan memeriksa dalam pada residual mengikuti distribusi normal atau tidak (Ghozali, 2021). Menurut Gujarati & Porter (2009), jika asumsi ini dilanggar, tidak akan menyebabkan estimasi parameter tidak BLUE. Untuk mendeteksi kenormalan residual dapat dilakukan melalui analisis grafik ataupun uji statistik. Pola sebaran normal dapat diamati dengan melihat histogram residu atau sebaran titik pada sumbu diagonal grafik dimana pola sebaran normal akan terlihat jika mengikuti arah garis pada histogram atau data tersebar di sekitar diagonal. Jika data menyimpang dari diagonal, pola sebaran normal tidak akan terlihat (Ghozali, 2021). Normalitas residual dapat diuji melalui *Kolmogorov-Smirnov test*. Uji *Kolmogorv-Smirnov* digunakan untuk mengevaluasi kesesuaian antara distribusi sampel dan distribusi lainnya, terutama distribusi normal dengan mean dan standar deviasi yang sama (Sintia, dkk., 2022). Menggunakan statistik uji:

$$L = |F(z_i) - S(z_i)|; z_i = \frac{x_i - \bar{x}}{sd} \quad (2.27)$$

dimana:

$$F(z_i) = \text{peluang teoritis nilai-nilai} \leq z_{hit} (P(Z \leq z_{hit}))$$

$$S(z_i) = \text{frekuensi kumulatif empiris nilai-nilai} \leq z_{hit} (P(Z \leq z_{hit}))$$

Menggunakan hipotesis:

$H_0$ : data residual berdistribusi normal

$H_1$ : data residual tidak berdistribusi normal

Dengan kriteria keputusan apabila probabilitasnya melebihi  $\alpha$ ,  $H_0$  akan diterima.

Normalitas residual dapat diuji melalui *Kolmogorov-Smirnov test*. Jika hasil yang diperoleh menunjukkan bahwa residual tidak berdistribusi normal, maka data dapat diubah atau diatasi dengan transformasi atau dengan metode regresi non-parametrik.



### 2.8.2 Uji Autokorelasi

Autokorelasi terjadi karena adanya korelasi linear antara residual pada periode  $t - 1$  (sebelumnya) dan pada periode  $t$  (saat ini). Uji ini dilakukan hanya pada model regresi linear data rangkaian waktu, dan tidak relevan untuk data unit maupun data panel yang mengkombinasikan keduanya. Uji Durbin-Watson (DW) ialah salah satu cara untuk memeriksa masalah autokorelasi. Statistik uji DW adalah:

$$DW = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2} \quad (2.28)$$

Menggunakan hipotesis:

$H_0$ : tidak terdapat autokorelasi

$H_1$ : terdapat autokorelasi

Dengan aturan keputusan adalah jika nilai  $p - value < \alpha$ , maka tolak  $H_0$ .

### 2.8.3 Uji *Heteroskedastisitas*

Menurut Ghazali (2021), uji ini dilakukan untuk menentukan apakah ada perbedaan dalam variasi residu antara observasi dalam model yang dihasilkan. Apabila variasi residual antara observasi konstan, disebut *homoskedastisitas*, sedangkan jika variasinya berbeda disebut *heteroskedastisitas*. Jika model mengalami masalah *heteroskedastisitas*, estimator tetap linear dan tidak bias namun varian minimum tidak lagi dimiliki (Widarjono, 2007). Konsekuensi yang dihadirkan jika estimator tidak mempunyai varian yang minimum dalam menggunakan adalah:

1. Menyebabkan ketidakpercayaan pada keakuratan perhitungan *standar error* dalam metode OLS jika variasi tidak minimum.
2. Menurut Widarjono (2007), dampak dari poin sebelumnya adalah uji hipotesis dan interval estimasi yang bergantung pada distribusi  $t$  atau  $F$  tidak dapat diandalkan lagi untuk hasil regresi.



Melalui uji Breusch-Pagan, *heteroskedastisitas* dapat dideteksi dengan statistik uji sebagai berikut:

$$\chi^2 = n * R^2_{new} \quad (2.29)$$

keterangan:

$n$  = jumlah sampel

$R^2_{new}$  = R-kuadrat dari model regresi baru yang menggunakan residual kuadrat sebagai nilai respon

Menggunakan hipotesis:

$H_0$ : tidak terdapat *heteroskedastisitas*

$H_1$ : terdapat *heteroskedastisitas*

Dengan aturan keputusan adalah tolak  $H_0$  jika nilai  $p - value < \alpha$

#### 2.8.4 Uji Multikolinearitas

Multikolinearitas merupakan hubungan dekat-linear atau *near-linear relationships* diantara variabel independen. Efek multikolinearitas dapat membuat perkiraan koefisien regresi yang tidak akurat, menyebabkan kesalahan standar atau *standar error* koefisien regresi menjadi lebih besar, membuat uji-t parsial koefisien regresi lebih rendah, memberikan nilai  $p-value$  yang tidak signifikan dan menurunkan kemampuan model untuk memprediksi (Khalaf, 2022). Uji ini digunakan dalam menentukan antara variabel independen dalam model terdapat korelasi yang signifikan. Ketika variabel independen berkorelasi, maka menunjukkan adanya multikolinearitas yang dapat memengaruhi interpretasi model. Mengetahui nilai *variance inflation factor* (VIF) ialah cara yang dilakukan untuk mengidentifikasi adanya multikolinearitas. Nilai VIF menunjukkan sejauh mana variabel independen ke- $j$  bergantung pada kumpulan variabel independen yang lain (Kusumawardhani, dkk., 2021). Persamaan VIF dinyatakan sebagai berikut:

$$VIF_j = \frac{1}{1-R_j^2} \quad (2.30)$$



keterangan:

$R_j^2$  = koefisien korelasi kuadrat antara variabel independen  $x_j$  dengan variabel independen lainnya

$j$  = jumlah sampel 1, 2, ..., n

Menggunakan hipotesis:

$H_0$ : tidak terdapat multikolinearitas

$H_1$ : terdapat multikolinearitas

Dengan aturan keputusan adalah tolak  $H_0$  jika nilai VIF > 10.

## 2.9 Uji Statistik Model

Dalam penelitian ini, uji statistik model yang akan dilakukan mencakup uji F statistik, uji t statistik, dan uji koefisien determinasi ( $R^2$ ).

### 2.9.1 Uji F

Uji ini dilakukan untuk mengevaluasi pengaruh variabel terikat terhadap variabel bebas secara bersamaan atau simultan pada model regresi melalui nilai F hitung yang didapat dibandingkan dengan nilai F tabelnya. Menggunakan hipotesis:

$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$  (variabel independen terhadap variabel dependen secara bersamaan tidak memiliki pengaruh yang signifikan)

$H_1$  : setidaknya terapat satu variabel yang memiliki pengaruh secara signifikan terhadap variabel dependen

Dengan statistik uji:

$$F_{hitung} = \frac{R^2/(n+K-1)}{(1-R^2)/(nT-n-K)} \quad (2.31)$$



keterangan:

$R^2$  = koefisien determinasi

$n$  = banyak individu

$T$  = jumlah waktu

$K$  = banyaknya variabel bebas.

Apabila nilai  $F_{hitung}$  melebihi nilai  $F_{tabel}$  yang dihitung pada tingkat signifikansi  $\alpha$  dengan derajat kebebasan  $(n + K - 1, nT - n - K)$ , maka dapat tolak  $H_0$ , menunjukkan bahwa terdapat pengaruh signifikan dari variabel bebas terhadap variabel terikat (Gujarati & Porter, 2009).

### 2.9.2 Uji t

Uji t, juga dikenal sebagai uji satu-satu, digunakan untuk menilai signifikansi variabel bebas masing-masing terhadap variabel terikatnya. Menggunakan hipotesis:

$H_0 : \beta_j = 0$  (variabel independen tidak memiliki pengaruh signifikan terhadap variabel dependen)

$H_1$  : paling tidak terdapat satu  $\beta_j \neq 0$

Dengan statistik uji:

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (2.32)$$

keterangan:

$\hat{\beta}_j$  = koefisien regresi (*slope*) variabel bebas ke  $i$

$SE(\hat{\beta}_j)$  = standar error penduga  $\hat{\beta}_i$

$i = 1, 2, 3, \dots, k$

Kriteria uji yang digunakan adalah jika  $|t_{hitung}| > t_{tabel(\alpha, n-1)}$ , maka tolak  $H_0$ ,



artinya variabel bebas ke- $j$  memiliki pengaruh yang signifikan terhadap variabel terikat (Wei, 2006).

### 2.9.3 Koefisien Determinasi ( $R^2$ )

Koefisien determinasi ( $R^2$ ) merupakan metrik penting dalam analisis dengan memberikan informasi mengenai seberapa besar variabel independen dapat menjelaskan variasi dalam variabel dependen (Gaur & Gaur, 2006). Nilai  $R^2$  diperoleh dari:

$$R^2 = \frac{JK_{regresi}}{JK_{total}} = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (2.33)$$

Jika suatu  $R^2$  sebesar 1 atau semakin tinggi nilainya, artinya kecocokan sempurna, namun jika  $R^2$  bernilai 0 atau semakin rendah artinya tidak terdapat keterkaitan atau relasi antar variabel bebas dengan variabel yang menjelaskan (Gujarati & Porter, 2009). Dimana batasnya adalah  $0 \leq R^2 \leq 1$ .

## 2.10 Indeks Pembangunan Manusia

Indeks Pembangunan Manusia menjadi alat dalam mengukur tingkat pencapaian pembangunan manusia di suatu daerah serta menjadi indikator yang sangat penting untuk mengevaluasi aspek lain dari pembangunan (BPS, 2020). Umur panjang dan hidup sehat, pengetahuan, dan standar hidup layak merupakan tiga dimensi dasar IPM.

### 1. Dimensi Kesehatan

Angka Harapan Hidup (AHH), adalah estimasi rata-rata lamanya waktu yang dapat dijalani seseorang selama hidupnya (dalam tahun). Ini merupakan indikator yang mencerminkan dimensi umur panjang dan hidup sehat. Untuk menghitung AHH, terlebih dahulu menghitung Angka Lahir Hidup (ALH) dan



Angka Masih Hidup (AMH). ALH adalah jumlah bayi yang lahir hidup dalam satu tahun kalender di suatu wilayah, sedangkan AMH adalah jumlah orang yang masih hidup pada akhir tahun kalender di wilayah yang sama (BPS, 2022).

$$I_{kesehatan} = \frac{AHH - AHH_{min}}{AHH_{maks} - AHH_{min}} \quad (2.34)$$

keterangan:

$$AHH = \frac{(ALH + AMH)}{2}$$

## 2. Dimensi Pendidikan

Harapan Lama Sekolah (HLS) dan Rata-Rata Lama Sekolah (RLS) mencerminkan akses pendidikan formal masyarakat. HLS mencerminkan peluang masyarakat untuk mengikuti pendidikan formal, sementara RLS menghitung masyarakat berusia 7 tahun ke atas dikarenakan program wajib belajar. Penghitungan melibatkan individu berusia diatas atau sama dengan 25 tahun.

$$I_{HLS} = \frac{HLS - HLS_{min}}{HLS_{maks} - HLS_{min}} \quad (2.35)$$

keterangan:

$$HLS = \sum_{i=7}^k \frac{E_i}{P_i}$$

$E_i$  = jumlah masyarakat yang masih sekolah berusi 7 tahun keatas

$P_i$  = jumlah masyarakat berusia 7 tahun ke atas

$$I_{RLS} = \frac{RLS - RLS_{min}}{RLS_{maks} - RLS_{min}} \quad (2.36)$$

keterangan:

$$RLS = \frac{1}{n} \sum_{i=1}^n x_i$$

$x_i$  = lama sekolah masyarakat ke- $i$  di suatu wilayah dan  $n$  jumlah

masyarakat ( $i = 1, 2, 3, \dots, n$ )

$$I_{pendidikan} = \frac{I_{HLS} + I_{RLS}}{2} \quad (2.37)$$

## 3. Dimensi Standar Hidup Layak

Dimensi ini digambarkan melalui pengeluaran riil per kapita yang telah ditetapkan (BPS, 2022).

$$I_{pengeluaran} = \frac{\ln(pengeluaran) - \ln(pengeluaran_{min})}{\ln(pengeluaran_{maks}) - \ln(pengeluaran_{min})} \quad (2.38)$$



#### 4. Menghitung IPM

IPM adalah rata-rata geometrik dari indeks pengeluaran, kesehatan dan pendidikan.

$$IPM = \sqrt[3]{I_{kesehatan} \times I_{pendidikan} \times I_{pengeluaran}} \quad (2.39)$$

Dikatakan sangat tinggi jika  $IPM \geq 80$ , tinggi jika  $70 \leq IPM < 80$ , sedang jika  $60 \leq IPM < 70$ , dan rendah jika  $IPM < 60$  (BPS, 2022).

### 2.11 Persentase Penduduk Miskin

Kemiskinan merupakan masalah yang kompleks dan memiliki dimensi multidimensional, sehingga menjadi prioritas dalam agenda pembangunan. Tingkat pengangguran, tingkat kesehatan, dan pendidikan semuanya terkait dengan tingkat kemiskinan di Indonesia. Biasanya, tingkat kemiskinan yang tinggi diikuti oleh rendahnya tingkat pendidikan dan kesehatan serta tingginya angka pengangguran (Auliana & Mirtawati, 2021). Persentase penduduk miskin nasional dihitung dengan:

$$\%PM_I = \frac{PM_I}{P_I} \times 100\% \quad (2.40)$$

keterangan:

$\%PM_I$  = persentase penduduk miskin

$PM_I$  = jumlah penduduk miskin

$P_I$  = jumlah penduduk



### **III. METODOLOGI PENELITIAN**

#### **3.1 Waktu dan Tempat Penelitian**

Penelitian ini dilakukan pada semester ganjil tahun ajaran 2023/2024 di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

#### **3.2 Data Penelitian**

Data pada penelitian ini berasal dari situs resmi Badan Pusat Statistik (BPS) Sumatera Utara yang terdiri atas variabel  $Y$  = Indeks Pembangunan Manusia (IPM),  $X_1$  = Angka Harapan Hidup,  $X_2$  = Rata-rata Lama Sekolah,  $X_3$  = Harapan Lama Sekolah,  $X_4$  = Pengeluaran Perkapita,  $X_5$  = Persentase Penduduk Miskin (%) di Provinsi Sumatera Utara pada tahun 2012-2020.

#### **3.3 Metode Penelitian**

Pada penelitian ini, metode *library research* atau studi literatur adalah metode yang digunakan. Metode ini melibatkan pengumpulan informasi serta data dalam berbagai sumber seperti jurnal, buku dan sumber lainnya. Langkah-langkah dalam penelitian ini meliputi:



1. Melakukan pengumpulan data
2. Mengestimasi parameter  $\beta$  dalam regresi data panel menggunakan metode *Ordinary Least Square* dengan langkah-langkah:
  - a. Meminimumkan fungsi total kuadrat *error*.
  - b. Melakukan turunan pertama pada fungsi *error* terhadap parameter koefisien regresi.
  - c. Menyamakan turunan pertama dari fungsi *error* dengan nol.
3. Melakukan analisis melalui tiga pendekatan model regresi data panel, yaitu *common effect*, *fixed effect*, dan *random effect*.
4. Penentuan model regresi data panel terbaik.
5. Penerapan model *common effect* pada data Indeks Pembangunan Manusia (IPM) di Provinsi Sumatera Utara pada tahun 2012-2020 menggunakan *software RStudio*.
6. Melakukan uji asumsi dan statistik pada model yang telah diperoleh menggunakan *software RStudio* dengan langkah-langkah:
  - a. Melakukan uji normalitas residual dengan uji *Kolmogorov-Smirnov*.
  - b. Melakukan uji heteroskedastisitas dengan uji Breusch-Pagan.
  - c. Melakukan uji multikolinearitas dengan melihat nilai *Variance Inflation Factor* (VIF).
  - d. Melakukan uji signifikansi parameter secara simultan (Uji F) dan parsial (Uji t).
  - e. Melakukan uji signifikansi model untuk memastikan kesesuaian regresi yang dapat tercermin melalui koefisien determinan ( $R^2$ ).
7. Membuat kesimpulan.



## V. KESIMPULAN

Berdasarkan hasil estimasi yang telah diperoleh, maka didapat *Common Effect Model* dari data Indeks Pembangunan Manusia (IPM) di Provinsi Sumatera Utara pada tahun 2012-2020 adalah:

$$\hat{Y}_{it} = 25,78 + 0,2178X_{1it} + 1,283X_{2it} + 0,6024X_{3it} + 0,001086 X_{4it} - 0,1043X_{5it}$$

Secara statistik, disimpulkan bahwa Angka Harapan Hidup, Rata-rata Lama Sekolah, Harapan Lama Sekolah, Pengeluaran Perkapita dan Persentase Penduduk Miskin secara signifikan mempengaruhi Indeks Pembangunan Manusia di Provinsi Sumatera Utara pada tahun 2012-2020. Variabel-variabel tersebut mampu menjelaskan 91,06% variasi dalam IPM di Provinsi Sumatera Utara, sementara 8,94% sisanya dijelaskan oleh faktor lain yang tidak ada dalam model.



## DAFTAR PUSTAKA

- 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, hlm. 254-266. Prosiding Seminar Nasional Matematika, Statistika dan Aplikasinya, Samarinda.
- Alviani, L.O. 2021. Penggunaan Regresi Data Panel Pada Analisis Indeks Pembangunan Manusia. *Journal Riset Matematika*. 1(02): 99-108.
- Aulina, N. & Mirtawati, M. 2021. Analisis Regresi Data Panel Pada Faktor-faktor yang Mempengaruhi Kemiskinan di Indonesia Tahun 2015–2019. *Kinerja*. 4(01): 78-90.
- Aziz, Abdul. 2010. *Ekonometrika Teori dan Praktik Eksperimen dengan MATLAB*. UIN-Maliki Press, Malang.
- Badan Pusat Statistik Indonesia. 2022. Indeks Pembangunan Manusia Badan Pusat Statistik 2022. <https://www.bps.go.id/>. Diakses pada 10 Oktober 2023.
- Badan Pusat Statistik Sumatera Utara. 2020. Indeks Pembangunan Manusia Badan Pusat Sumatera Utara 2020. <https://sumut.bps.go.id/>. Diakses pada 3 November 2023.
- Basuki, A.T. & Prawoto, N. 2017. *Analisis Regresi Dalam Penelitian Ekonomi Dan Bisnis*. Rajawali, Jakarta.
- Best, H. & Wolf, C. 2015. *Regression Analysis and Causal Inference*. Sage, Los Angeles.



- Chairunnisa, A.D.A. & Fauzan, A. 2023. Implementation of Panel Data Regression in the Analysis of Factors Affecting Poverty Levels in Bengkulu Province in 2017-2020: Implementation of Panel Data Regression. *EKSAKTA, Journal of Sciences and Data Analysis*. **4**(1): 40-45.
- Gaur, A.S. & Gaur, S.S. 2006. *Statistical Methods for Practice and Research: A Guide to Data Analysis Using SPSS*. 2<sup>nd</sup> Edition. Sage Publications, United Kingdom.
- Ghozali, I. 2021. *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 26*. Badan Penerbit Universitas Diponegoro, Semarang.
- Ghozi, S. & Hermansyah, H. 2018. Analisis Regresi Data Panel Profitabilitas Bank Pembangunan Daerah (BPD) di Indonesia. *Jurnal Matematika*. **8**(1): 1-12.
- Greene, W.H. 2011. *Econometric Analysis*. 7<sup>th</sup> Edition. Pearson, New York.
- Gujarati, D.N. & Porter, D.C. 2009. *Basic Econometrics*. 5<sup>th</sup> Edition. McGraw Hill Companies, New York.
- Hill, R.C, Griffiths, W.E., & Lim, G.C. 2011. *Principles of Econometrics*. 4<sup>th</sup> Edition. John Wiley and Sons, New York.
- Hsiao, C. 2014. *Analysis of Panel Data* (No. 54). 3<sup>rd</sup> Edition. Cambridge University Press, United States of America.
- Hutagalung, I.P. & Darnius, O. 2022. Analisis Regresi Data Panel dengan Pendekatan Common Effect Model (CEM), Fixed Effect Model (FEM) dan Random Effect Model (REM) (Studi Kasus: IPM Sumatera Utara Periode 2014–2020). *Jurnal Matematika dan Pendidikan Matematika FARABI*. **5**(2): 217-226.
- Khalaf, G. 2022. Improving the Ordinary Least Squares Estimator by Ridge Regression. *Open Access Library Journal*. **9**(5): 1-8.



- Khasanah, N. R., Ernitawati, Y., & Sunanti, S. 2021. Komite Audit dan Kepemilikan Institusional: Kinerja Keuangan Perusahaan Terindeks LQ45. *Journal of Accounting and Finance (JACFIN)*. 2(1): 28-44.
- Kusumawardhani, R., Rizqiena, Z. D., & Astuti, S. P. 2021. *Ekonometrika Suatu Pengantar*. CV Gerbang Media Aksara, Yogyakarta.
- Madany, N., Ruliana, & Rais, Z. 2022. Regresi Data Panel dan Aplikasinya dalam Kinerja Keuangan Terhadap Pertumbuhan Laba Perusahaan IDX LQ45 Bursa Efek Indonesia. *Journal of Statistics and Its Application on Teaching and Research*. 4(2): 79-94.
- Nurhalizah, & Sitompul, P. 2022. Analysis of Ordinary Least Square and Geographically Weighted Regression on the Human Development Index of North Sumatera 2021. *Formosa Journal of Applied Sciences*. 1(6): 981-1000.
- Nuryadi, N., Astuti, T.D., Utami, E.S., & Budiantara, M. 2018. *Dasar-dasar Statistik Penelitian*. Sibuku Media, Yogyakarta.
- Petronilla, O.U. 2022. On the Comparison of Ordinary Least Squares and Quantile Regression with Nigerian Financial Data on Trade Balance, Foreign Inflow and Imports. *Asian Journal of Probability and Statistics*. SAGE Publications, Croydon. 17(1): 43-52.
- Sintia, I., Esra, R., Pasarella, M.D., & Nohe, D.A. 2022. Perbandingan Tingkat Konsistensi Uji Distribusi Normalitas Pada Kasus Tingkat Pengangguran Di Jawa, hlm. 323-333. Prosiding Seminar Nasional Matematika, Statistika dan Aplikasinya, Samarinda.
- Sulistiyowati, W. & Astuti, C.C. 2017. *Statistika Dasar Konsep dan Aplikasinya*. UMSIDA Press, Sidoarjo.
- Sunyoto, D. 2020. *Analisis Regresi dan Uji Hipotesis*. CAPS, Yogyakarta.
- Wardhana, F.A. & Indawati, R. 2021. Panel Data Regression Analysis for Factors Affecting Mortality Rate in East Java 2013-2017. *The Indonesian Journal of Public Health*. 16(3): 437-448.



- Wasiaturrahma, W. & Rohmawati, H. 2021. Analysis of International Tourism Demand in Indonesia: An Ordinary Least Square (OLS) Approach. *Journal of Economic Education*. **10**(2): 158-172.
- Wei, W.W.S. 2006. *Time Series Analysis: Univariate and Multivariate Methods*. 2<sup>nd</sup> Edition. Addison-Wesley Publishing Company, New York.
- Widarjono, A. 2007. *Ekonometrika: Teori dan Aplikasi Untuk Ekonomi dan Bisnis*. Ekonisia, Yogyakarta.
- Winarno, W.W. 2007. *Analisis Ekonometrika dan Statistika dengan Eviews*. UPP STIM YKPN, Yogyakarta.