PEMODELAN REGRESI DATA PANEL DENGAN MENGGUNAKAN PENDEKATAN *RANDOM EFFECT MODEL*

(Skripsi)

Oleh

RAFADHIA ARDINA
NPM 1717031033

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2021

ABSTRACT

PANEL DATA REGRESSION MODELING USING RANDOM EFFECT MODEL APPROACH

By

RAFADHIA ARDINA

Panel data is a combination of cross section data and time series data in several sectors observed from an object of research during a certain period of time. One of the panel data regression models is the random effect model. This study aims to determine the Human Development Index model in Lampung Province with panel data regression using the Random Effect Model (REM) approach and to determine the variables that affect the Human Development Index in Lampung Province. The research was conducted by performing parameter estimation, selection test, classical assumption test and feasibility test on panel data regression model. The results of this study obtained a panel data regression model to model the Human Development Index in Lampung Province in 2016-2018 is the Random Effect Model (REM) where the variables of life expectancy, average length of schooling and adjusted per capita expenditure are able to explain the variable Human Development Index (HDI) is 98.56%, while other variables outside the model explain the remaining 1.44%. With the equation model the estimation results are $\hat{Y}_{it} = 8.383762 \ + 0.500353\, X_{1it} + 1.615153\, X_{2it} + \ 0.001290\, X_{3it}$.

Keywords : Panel Data, Human Development Index (IPM), Random Effect Model (REM)

ABSTRAK

PEMODELAN REGRESI DATA PANEL DENGAN MENGGUNAKAN PENDEKATAN *RANDOM EFFECT MODEL*

Oleh

RAFADHIA ARDINA

Data panel merupakan gabungan antara data *cross section* dan data *time series* pada beberapa sektor yang diamati dari suatu objek penelitian selama periode waktu tertentu. Salah satu model regresi data panel yaitu model *random effect*. Tujuan dari penelitian ini yaitu untuk menentukan model Indeks Pembangunan Manusia di Provinsi Lampung dengan regresi data panel menggunakan pendekatan *Random Effect Model* (REM) dan mengetahui variabel-variabel yang berpengaruh terhadap Indeks Pembangunan Manusia di Provinsi Lampung. Penelitian dilakukan dengan melakukan estimasi parameter, uji pemilihan, uji asumsi klasik dan uji kelayakan pada model regresi data panel. Hasil dari penelitian ini diperoleh model regresi data panel untuk memodelkan Indeks Pembangunan Manusia di Provinsi Lampung tahun 2016-2018 adalah *Random Effect Model* (REM) dimana variabel angka harapan hidup, rata-rata lama sekolah dan pengeluaran perkapita disesuaikan mampu menjelaskan variabel Indeks Pembangunan Manusia (IPM) sebesar 98.56%, sedangkan sisanya sebesar 1.44% dijelaskan oleh variabel lain di luar model. Dengan model persamaan hasil estimasi yaitu $\hat{Y}_{it} = 8.383762\ + 0.500353\ X_{1it} + 1.615153\ X_{2it} +\ 0.001290\ X_{3it}$.

Kata kunci : Data Panel, Indeks Pembangunan Manusia (IPM), *Random Effect Model* (REM)

PEMODELAN REGRESI DATA PANEL DENGAN MENGGUNAKAN PENDEKATAN *RANDOM EFFECT MODEL*

Oleh

RAFADHIA ARDINA

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
2021

Judul Skripsi : **PEMODELAN REGRESI DATA PANEL DENGAN MENGGUNAKAN PENDEKATAN *RANDOM EFFECT MODEL***

Nama Mahasiswa : **Rafadhia Ardina**

Nomor Pokok Mahasiswa : **1717031033**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI

1. Komisi Pembimbing

Drs. Nusyirwan, M.Si.	**Dr. Fitriani, S.Si., M.Sc.**
NIP. 19661010 199205 1 001	NIP. 19840627 200604 2 001

2. 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. Fitriani, S.Si., M.Sc.**	....
Penguji Bukan Pembimbing	: **Prof. Drs. Mustofa Usman, MA, Ph.D.**	....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Suripto Dwi Yuwono, S.Si., M.T.
NIP. 19740705 200003 1 001

Tanggal Lulus Ujian Skripsi : **8 Oktober 2021**

PERNYATAAN SKRIPSI MAHASISWA

Saya yang bertanda tangan di bawah ini :

Nama	: **Rafadhia Ardina**
Nomor Pokok Mahasiswa	: **1717031033**
Jurusan	: **Matematika**
Judul	: **PEMODELAN REGRESI DATA PANEL DENGAN MENGGUNAKAN PENDEKATAN *RANDOM EFFECT MODEL***

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri dan semua hasil yang tertuang dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 19 Oktober 2021
Penulis

Rafadhia Ardina
NPM. 1717031033

RIWAYAT HIDUP

Penulis dilahirkan di Pringsewu pada tanggal 2 Mei 1999. Penulis merupakan anak pertama dari tiga bersaudara, dari pasangan Bapak Herman Pratikno dan Ibu Reni Wahyuni.

Pendidikan Taman Kanak-Kanak (TK) diselesaikan penulis di TK Pertiwi Gedong Tataan pada tahun 2005, Sekolah Dasar (SD) diselesaikan di SDS Adhyaksa 1 Kota Jambi pada tahun 2011, Sekolah Menengah Pertama (SMP) diselesaikan di SMP Negeri 7 Kota Jambi pada tahun 2014 dan Sekolah Menengah Atas (SMA) diselesaikan di SMA Negeri 9 Bandar Lampung pada tahun 2017.

Pada pertengahan tahun 2017, penulis terdaftar sebagai mahasiswi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SNMPTN. Sebagai bentuk pengabdian mahasiswa dan menjalankan Tri Dharma Perguruan Tinggi, pada bulan Januari hingga Februari tahun 2020 penulis melaksanakan Kuliah Kerja Nyata (KKN) di Desa Kota Agung, Kecamatan Sungkai Selatan, Kabupaten Lampung Utara. Pada bulan Juli hingga Agustus tahun 2020, penulis melaksanakan Kerja Praktik (KP) di PT. Lotte Shopping Indonesia, Lotte Grosir Lampung sebagai bentuk penerapan ilmu yang telah diperoleh selama kuliah.

KATA INSPIRASI

"Barang siapa bertakwa kepada Allah niscaya Dia akan membukakan jalan keluar baginya, dan Dia memberinya rezeki dari arah yang tidak disangka-sangkanya. Dan barang siapa bertawakal kepada Allah, niscaya Allah akan mencukupkan (keperluan)nya."
(Q.S. At-Talaq: 2-3)

"Maka sesungguhnya bersama kesulitan itu ada kemudahan."
(Q.S. Al-Insyirah: 5)

"Kelemahan terbesarmu adalah ketika kamu menyerah dan kehebatan terbesarmu adalah ketika kamu mencoba sekali lagi."

"When you want to give up, look at back and then see how far you have climbed to reach your goals."

"We will never know the real answer before we try."

PERSEMBAHAN

Atas rahmat dan karunia-Nya, dengan segala kerendahan hati
kupersembahkan karya ini untuk:

Bapak dan Ibu tercinta
Yang tak kenal lelah dalam merawat, mendidik dan membesarkanku.
Yang selalu memberikan doa dan dukungan di setiap langkahku.
Terima kasih atas segala hal yang telah diberikan untukku.

Adik-adikku tersayang
Yang telah membantu dan memberiku semangat.
Terima kasih sudah menjadi tempatku untuk berbagi cerita.

Dosen Pembimbing dan Penguji
Terima kasih telah senantiasa meluangkan waktunya
untuk memberikan bimbingan dan ilmu pengetahuan kepadaku.

Almamater Tercinta, Universitas Lampung

SANWACANA

Puji syukur atas kehadirat Allah SWT yang telah melimpahkan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "**Pemodelan Regresi Data Panel Dengan Menggunakan Pendekatan *Random Effect Model***". Penulis menyadari bahwa skripsi ini tidak akan terwujud tanpa adanya bimbingan, saran dan bantuan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Bapak Drs. Nusyirwan, M.Si., selaku Dosen Pembimbing I yang telah memberikan bimbingan, arahan dan saran dalam menyelesaikan skripsi ini.
2. Ibu Dr. Fitriani, S.Si., M.Sc., selaku Dosen Pembimbing II yang telah memberikan bimbingan, saran dan motivasi dalam menyelesaikan skripsi ini.
3. Bapak Prof. Drs. Mustofa Usman, MA, Ph.D., selaku Dosen Penguji yang telah memberikan saran dan masukan kepada penulis dalam menyelesaikan skripsi ini.
4. Bapak Agus Sutrisno, S.Si., M.Si., selaku Dosen Pembimbing Akademik yang telah memberikan bimbingan dan arahan kepada penulis selama kuliah.
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. Suripto Dwi Yuwono, S.Si., M.T., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Seluruh Dosen, Staf dan Karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung yang telah memberikan ilmu pengetahuan yang tak ternilai dan bantuan kepada penulis.
8. Bapak, Ibu, Shafa dan Danang yang selalu mendoakan, memberikan semangat dan memotivasi penulis dalam menyelesaikan perkuliahan.
9. Nabilla, Beta, Umroh, Arina dan Viona yang telah memberikan cerita, pengalaman dan motivasi kepada penulis selama kuliah.
10. Yulica, Chaterina, Stefani, Dini, Eka Anisa, Felicia yang telah memberikan canda tawa, dukungan dan semangat kepada penulis selama ini.
11. Teman-teman satu bimbingan yang sudah berjuang bersama dan saling memotivasi satu sama lain.
12. Teman-teman Mahasiswa Jurusan Matematika FMIPA Unila Angkatan 2017 dan semua pihak yang telah membantu yang tidak mungkin disebutkan satu persatu.

Akhir kata, penulis menyadari bahwa penulisan dalam skripsi ini masih banyak kekurangan. Oleh karenanya, penulis mengharapkan kritik dan saran yang membangun kedepannya dan semoga skripsi ini dapat bermanfaat bagi siapapun.

Bandar Lampung, 19 Oktober 2021
Penulis,

Rafadhia Ardina

DAFTAR ISI

Halaman

DAFTAR TABEL

DAFTAR GAMBAR

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis regresi merupakan suatu metode analisis statistik yang banyak digunakan dalam berbagai bidang kehidupan. Analisis regresi digunakan untuk mengukur pengaruh antara variabel bebas terhadap variabel terikat. Regresi data panel adalah analisis regresi yang digunakan dalam data panel. Data panel merupakan gabungan dari data *cross section* dan data *time series*. Apabila informasi dari kedua data tersebut tersedia, maka dapat dilakukan analisis dengan menggunakan data panel.

Model regresi yang menggunakan data panel disebut dengan model regresi data panel. Terdapat tiga pendekatan yang biasa digunakan untuk mengestimasi parameter model regresi pada data panel yaitu *Common Effect Model* (CEM) menggunakan metode *Ordinary Least Square* (OLS), *Fixed Effect Model* (FEM) dengan menggunakan metode *Least Square Dummy Variable* (LSDV) dan *Random Effect Model* (REM) dengan metode *Generalized Least Square* (GLS).

Menurut Gujarati (2012), walaupun model *fixed effect* dapat diestimasi secara langsung dengan menggunakan metode LSDV, namun model yang terbentuk akan kehilangan sejumlah derajat bebas seiring dengan banyaknya unit *cross section* yang akan digunakan. Semakin kecil derajat bebas, maka akan berpengaruh pada Uji F yang cenderung bernilai kecil, sehingga peluang untuk menolak H_0 pun semakin kecil. Pada model *random effect*, tidak akan terjadi resiko kehilangan derajat bebas karena tidak menggunakan variabel *dummy*.

Manusia merupakan kekayaan bangsa yang sesungguhnya. Indeks Pembangunan Manusia (IPM) merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia. IPM menjelaskan mengenai bagaimana penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan dan lain sebagainya. IPM dibentuk oleh tiga dimensi dasar yaitu umur panjang dan hidup sehat (*a long and healthy life*), pengetahuan (*knowledge*) dan standar hidup layak (*decent standard of living*).

Berdasarkan uraian yang telah dipaparkan tersebut, pada penelitian ini penulis akan menentukan model regresi data panel pada data Indeks Pembangunan Manusia di Provinsi Lampung tahun 2016-2018 berdasarkan variabel-variabel yang mempengaruhinya dengan menggunakan pendekatan *Random Effect Model* (REM).

1.2 Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, tujuan dari penelitian ini adalah sebagai berikut:

1. menentukan model Indeks Pembangunan Manusia di Provinsi Lampung dengan regresi data panel menggunakan pendekatan *Random Effect Model* (REM),
2. mengetahui variabel-variabel yang berpengaruh terhadap Indeks Pembangunan Manusia (IPM) di Provinsi Lampung.

1.3 Manfaat Penelitian

Adapun manfaat dari penelitian ini yaitu:

1. menambah wawasan dan pengetahuan mengenai analisis regresi data panel,
2. sebagai referensi bagi penelitian selanjutnya dalam mengolah data dengan menggunakan data panel.

II. TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan suatu metode analisis statistik yang dapat digunakan untuk memeriksa dan memodelkan hubungan antar variabel (Montgomery & Peck, 1992). Dalam analisis regresi, dikenal dua jenis variabel yaitu variabel bebas dan variabel terikat. Variabel yang mempengaruhi atau menjadi sebab dari perubahan variabel terikat disebut dengan variabel bebas, sedangkan variabel terikat merupakan variabel yang dipengaruhi atau merupakan akibat dari variabel bebas.

Analisis regresi dibagi menjadi dua yaitu analisis regresi linear dan analisis regresi non linear. Analisis regresi linear terdiri dari analisis regresi linear sederhana dan analisis regresi linear berganda. Regresi linear sederhana ialah hubungan antara satu variabel bebas dengan variabel terikat. Model regresi linear sederhana secara matematis dapat dituliskan sebagai berikut (Draper & Smith, 1998):

$$Y_i = \beta_0 + \beta_1 X_1 + \varepsilon_i \tag{2.1}$$

dengan Y merupakan variabel terikat (*dependent variable*), X merupakan variabel bebas (*independent variable*), β_0 merupakan intersep pada model regresi, β_1 merupakan koefisien kemiringan (*slope*) pada model regresi, ε merupakan *error* dan indeks i menunjukkan pengamatan ke-i.

Sedangkan regresi linear berganda merupakan hubungan antara dua atau lebih variabel bebas dengan variabel terikat. Model regresi linear berganda dengan k variabel bebas dapat ditulis sebagai berikut (Montgomery & Peck, 1992):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ \dots \ + \beta_k X_{ik} + \varepsilon_i \ ; \quad i = 1,2,\dots,n \tag{2.2}$$

dengan:

Y_i	= nilai variabel terikat pada pengamatan ke-i
X_{ij}	= nilai variabel bebas ke-j ; $j = 1,2,\dots,k$ pada pengamatan ke-i
β_0	= intersep pada model regresi
$\beta_1, \beta_2, \dots, \beta_k$	= koefisien *slope*
ε_i	= nilai galat atau *error* ; $\varepsilon_i \sim N(0, \sigma^2)$
k	= banyaknya variabel bebas
n	= banyaknya data pengamatan

Oleh karena i menunjukkan pengamatan, maka n persamaan:

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

dalam bentuk matriks, Persamaan (2.3) menjadi:

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

Persamaan (2.4) dapat ditulis dalam matriks:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}. \tag{2.5}$$

Menurut Gujarati (2003), asumsi-asumsi pada model regresi liner berganda adalah sebagai berikut:

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

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

Persamaan (2.6) jika dituliskan dalam bentuk matriks adalah sebagai berikut:

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

4. Tidak terdapat korelasi antara *error* (ε_i) dan *error* lainnya $\left(\varepsilon_j\right)$.

$$E\left(\varepsilon_i \varepsilon_j\right) = 0\ ;\quad i \neq j \tag{2.8}$$

5. Varian dari setiap *error* adalah sama (homoskedastisitas).

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

Persamaan (2.9) jika dituliskan dalam bentuk matriks adalah sebagai berikut:

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

$$Var(\boldsymbol{\varepsilon}) = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_\varepsilon^2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_\varepsilon^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \sigma_\varepsilon^2 \end{bmatrix}$$

$$= \sigma_\varepsilon^2 \boldsymbol{I} \tag{2.10}$$

6. *Error* berdistribusi normal.

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

2.2 Analisis Regresi Data Panel

Analisis regresi data panel merupakan suatu teknik yang dapat digunakan untuk memodelkan pengaruh antara variabel bebas terhadap variabel terikat pada beberapa sektor yang diamati dari suatu objek penelitian selama periode waktu tertentu. Menurut Gujarati (2012), data panel merupakan gabungan antara data *cross section* dan data *time series*. Jika setiap unit *cross section* memiliki jumlah pengamatan *time series* yang sama, maka data panel disebut seimbang (*balanced panel data*). Sebaliknya, jika setiap unit *cross section* memiliki jumlah pengamatan *time series* yang berbeda, maka data panel disebut tidak seimbang (*unbalanced data panel*). Bentuk umum regresi data panel adalah:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + u_{it} \; ; \quad i = 1,2, \dots , N \; ; \quad t = 1,2, \dots , T \tag{2.12}$$

dengan:

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

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

β_0 = intersep

β_1 = *slope* untuk semua unit

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

N = banyaknya unit *cross section*

T = banyaknya unit *time series*.

Penggunaan regresi data panel akan menghasilkan intersep dan *slope* yang berbeda-beda untuk setiap unit *cross section* dan *time series*, sehingga dalam mengestimasi model regresi bergantung pada asumsi mengenai intersep, *slope* dan *error*. Menurut Hsiao (2003), terdapat beberapa kemungkinan asumsi yang akan muncul yaitu:

1. Intersep dan *slope* adalah konstan antar unit *cross section* dan *time series*. Modelnya dinyatakan sebagai berikut:
$$Y_{it} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{kit} + u_{it} \tag{2.13}$$
2. *Slope* adalah konstan, tetapi intersep berbeda antar unit *cross section*. Modelnya dinyatakan sebagai berikut:
$$Y_{it} = \beta_{0i} + \sum_{k=1}^{K} \beta_k X_{kit} + u_{it} \tag{2.14}$$
3. *Slope* adalah konstan, tetapi intersep berbeda antar unit *cross section* dan *time series*. Modelnya dinyatakan sebagai berikut:
$$Y_{it} = \beta_{0it} + \sum_{k=1}^{K} \beta_k X_{kit} + u_{it} \tag{2.15}$$
4. Intersep dan *slope* berbeda antar unit *cross section*. Modelnya dinyatakan sebagai berikut:
$$Y_{it} = \beta_{0i} + \sum_{k=1}^{K} \beta_{ki} X_{kit} + u_{it} \tag{2.16}$$
5. Intersep dan *slope* berbeda antar unit *cross section* dan *time series*. Modelnya dinyatakan sebagai berikut:
$$Y_{it} = \beta_{0it} + \sum_{k=1}^{K} \beta_{kit} X_{kit} + u_{it}, \tag{2.17}$$

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

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

K $=$ jumlah parameter dalam regresi yang akan diduga.

Keuntungan penggunaan metode regresi data panel diantaranya adalah sebagai berikut (Hsiao, 1992):

1. data panel yang merupakan gabungan dari data *cross section* dan data *time series* mampu menyediakan data yang lebih banyak, sehingga akan menghasilkan *degree of freedom* yang lebih besar,
2. menggabungkan informasi dari data *cross section* dan data *time series* dapat mengatasi masalah yang timbul ketika terdapat masalah penghilangan variabel.

2.3 Estimasi Parameter Model Regresi Data Panel

Untuk mengestimasi parameter model regresi pada data panel, terdapat tiga teknik yang dapat dilakukan yakni *Common Effect Model* (CEM), *Fixed Effect Model* (FEM) dan *Random Effect Model* (REM).

2.3.1 *Common Effect Model* (CEM)

Common Effect Model (CEM) merupakan teknik yang paling sederhana yang dapat dilakukan untuk mengestimasi parameter model data panel, yaitu dengan cara mengkombinasikan data *cross section* dan data *time series* sebagai satu kesatuan tanpa mempertimbangkan adanya perbedaan antar unit *cross section* dan *time series*. Model tersebut mengasumsikan bahwa intersep dan *slope* adalah konstan untuk semua unit *cross section* dan *time series* (Gujarati, 2012). Persamaan model *common effect* dapat ditulis sebagai berikut:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \ \dots + \beta_k X_{it} + u_{it} \tag{2.18}$$

dengan:

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

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

β_0 = intersep pada model regresi

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

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

2.3.1.1 *Ordinary Least Square* (OLS)

Metode estimasi parameter yang digunakan pada model *common effect* sama halnya dengan metode regresi linear biasa yaitu dengan metode *Ordinary Least Square* (OLS) atau juga dikenal dengan Metode Kuadrat Terkecil (MKT). Berikut adalah prinsip dasar MKT:

$$u = \boldsymbol{Y} - \boldsymbol{X\beta}, \tag{2.19}$$

sehingga didapatkan jumlah kuadrat galat sebagai berikut:

$$\begin{aligned}
\sum_{i=1}^{n} u_i^2 &= u'u \\
&= (\boldsymbol{Y} - \boldsymbol{X\beta})'(\boldsymbol{Y} - \boldsymbol{X\beta}) \\
&= (\boldsymbol{Y}' - \boldsymbol{X}'\boldsymbol{\beta}')(\boldsymbol{Y} - \boldsymbol{X\beta}) \\
&= \boldsymbol{Y}'\boldsymbol{Y} - \boldsymbol{\beta X Y}' - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{Y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X\beta} \\
&= \boldsymbol{Y}'\boldsymbol{Y} - \boldsymbol{2\beta}'\boldsymbol{X}'\boldsymbol{Y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X\beta}.
\end{aligned} \tag{2.20}$$

Apabila matriks $(\boldsymbol{X\beta})' = \boldsymbol{\beta}'\boldsymbol{X}'$, maka skalar $\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{Y} = \boldsymbol{Y}'\boldsymbol{X\beta}$. Untuk mendapatkan penduga parameter $\boldsymbol{\beta}$, maka hasil turunan disamakan dengan nol sehingga diperoleh persamaan sebagai berikut:

$$\frac{\partial(u'u)}{\partial \boldsymbol{\beta}} = \frac{\partial(\boldsymbol{Y}'\boldsymbol{Y} - \boldsymbol{2\beta}'\boldsymbol{X}'\boldsymbol{Y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X\beta})}{\partial \boldsymbol{\beta}} = 0$$

$$-\mathbf{2X'Y} + \mathbf{2X'X}\hat{\boldsymbol{\beta}} = 0$$
$$\mathbf{2X'X}\hat{\boldsymbol{\beta}} = \mathbf{2X'Y}$$
$$\mathbf{X'X}\hat{\boldsymbol{\beta}} = \mathbf{X'Y}$$
$$(\mathbf{X'X})^{-\mathbf{1}}(\mathbf{X'X})\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'Y}$$
$$\mathbf{I}\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'Y}$$
$$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'Y} \qquad (2.21)$$

Jika diasumsikan bahwa β_0 dan β_1 akan sama (konstan) untuk setiap data *time series* dan *cross section*, maka β_0 dan β_1 dapat diestimasi dengan menggunakan metode OLS. Sehingga model *common effect* dengan $n \times T$ pengamatan dapat dinyatakan sebagai berikut:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + u_{it}\ ; \quad i = 1,2,\dots,n\ ; \quad t = 1,2,\dots,T. \qquad (2.22)$$

2.3.2 *Fixed Effect Model* (FEM)

Salah satu cara untuk memperlihatkan perbedaan unit *cross section* adalah dengan mengijinkan nilai intersep yang berbeda-beda untuk setiap unit *cross section* tetapi masih mengasumsikan *slope* konstan (Gujarati, 2003). Pendekatan yang dipakai pada model *fixed effect* mengasumsikan bahwa intersep dari setiap unit *cross section* adalah berbeda, sedangkan *slope* antar unit *cross section* adalah konstan. Persamaan model *fixed effect* dapat ditulis sebagai berikut:

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_k X_{kit} + u_{it}. \qquad (2.23)$$

2.3.2.1 *Least Square Dummy Variable* (LSDV)

Menurut Greene (2007), estimasi parameter pada model *fixed effect* dilakukan dengan menggunakan metode *Least Square Dummy Variable* (LSDV). LSDV merupakan metode yang digunakan untuk mengestimasi parameter regresi linear dengan menggunakan Metode Kuadrat Terkecil (MKT) pada model yang melibatkan variabel *dummy* sebagai salah satu variabel bebasnya. Variabel *dummy* merupakan sebuah variabel yang hanya memiliki dua kemungkinan. Analisis pada variabel *dummy* dilakukan dengan memberi kode 1 pada salah satu kategori, sedangkan kategori lainnya diberi kode 0. Model *fixed effect* menggunakan variabel *dummy* untuk menjelaskan adanya perbedaan intersep antar unit *cross section*. Persamaan model *fixed effect* dengan variabel *dummy* adalah sebagai berikut:

$$Y_{it} = \beta_0 + \sum_{i=2}^{N} \beta_{0i} D_{ki} + \sum_{k=1}^{K} \beta_k X_{kit} + u_{it} \tag{2.24}$$

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

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

Subskrip $0i$ pada konstanta β_{oi} menunjukkan bahwa i merupakan objeknya. Dengan demikian dapat diartikan bahwa setiap objek memiliki konstanta yang berbeda. Variabel *dummy* D_{1i} berarti bahwa 1 untuk objek pertama, sedangkan 0 untuk objek lainnya. Variabel *dummy* yang dibentuk sebanyak $N - 1$, dengan β_0 sebagai intersep untuk unit *cross section* yang pertama (Singh & Sachdeva, 2020).

2.3.3 *Random Effect Model* (REM)

Pada model *random effect*, unit *cross section* yang digunakan tidak ditentukan terlebih dahulu melainkan dipilih secara acak dari suatu populasi (Greene, 2007). Pendekatan yang digunakan pada model *random effect* mengasumsikan bahwa setiap unit *cross section* memiliki perbedaan intersep. Namun demikian, diasumsikan bahwa β_{0i} merupakan variabel acak dengan *mean* β_0. Oleh karena itu, intersep dapat ditulis $\beta_{0i} = \beta_0 + \varepsilon_i$ dengan ε_i merupakan *error random* yang mempunyai *mean* nol dan varian σ_ε^2. Selain itu, teknik tersebut juga memperhitungkan bahwa *error* mungkin saling berkorelasi antar unit *cross section* dan *time series*. Model *random effect* digunakan untuk mengatasi kelemahan model *fixed effect* yang menggunakan variabel *dummy* (Widarjono, 2009). Persamaan model *random effect* dapat ditulis sebagai berikut:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \ ... + \beta_k X_{kit} + w_{it} \tag{2.25}$$

dengan:

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

β_0 = intersep

$\beta_1, \dots, \beta_k$ = koefisien *slope*

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

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

i = banyaknya unit *cross section* sebanyak N

t = banyaknya periode waktu (unit *time series*) sebanyak T,

dengan $w_{it} = \varepsilon_i + u_{it}$. Suku *error* gabungan w_{it} memuat dua komponen *error* yaitu ε_i sebagai komponen *error* pada unit *cross section* dan u_{it} yang merupakan kombinasi komponen *error* pada unit *cross section* dan *time series*. Karena inilah, REM juga disebut *Error Components Model* (ECM). Beberapa asumsi yang berlaku pada REM adalah:

$$\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$$

$$u_{it} \sim N(0, \sigma_u^2)$$
$$E(\varepsilon_i u_{it}) = 0 \quad (2.26)$$
$$E\left(\varepsilon_i \varepsilon_j\right) = 0\,; \quad (i \neq j)$$
$$E(u_{it} u_{is}) = E\left(u_{it} u_{jt}\right) = E(u_{it} u_{js}) = 0; \quad (i \neq j; t \neq s).$$

2.3.3.1 *Generalized Least Square* (GLS)

Estimasi parameter pada model *random effect* dilakukan menggunakan metode *Generalized Least Square* (GLS). Diketahui:

$$\boldsymbol{W} = \begin{bmatrix} \Omega_1 & 0 & \cdots & 0 \\ 0 & \Omega_2 & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & \Omega_n \end{bmatrix} \quad (2.27)$$

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

dengan $\sigma_\varepsilon^2 + \sigma_u^2$ sebagai nilai diagonalnya, sedangkan σ_ε^2 sebagai nilai lainnya.

Diberikan model regresi linear:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{u}\,; \quad E(\boldsymbol{uu}') = \boldsymbol{W} \quad (2.29)$$

Persamaan (2.29) dikalikan dengan matriks $\boldsymbol{T}$ sehingga diperoleh:

$$\boldsymbol{TY} = \boldsymbol{TX\beta} + \boldsymbol{Tu}. \quad (2.30)$$

dengan $E[(\boldsymbol{Tu})(\boldsymbol{Tu})'] = \boldsymbol{I}$ atau secara ekuivalen $\boldsymbol{TWT}' = \boldsymbol{I}$. $\boldsymbol{T}'(\boldsymbol{TWT}')\boldsymbol{T} = \boldsymbol{T}'\boldsymbol{T}$ sehingga $\boldsymbol{T}'\boldsymbol{T} = \boldsymbol{W}^{-1}$. Dari Persamaan (2.30) dapat dibentuk persamaan baru untuk galat yaitu:

$$\boldsymbol{Tu} = \boldsymbol{TY} - \boldsymbol{TX\beta} \quad (2.31)$$

Kemudian turunan pertama dari jumlah kuadrat $\boldsymbol{Tu}$ disamakan dengan nol sehingga diperoleh:

$$\begin{aligned}\widehat{\boldsymbol{\beta}} &= [(\boldsymbol{TX})'(\boldsymbol{TX})]^{-\mathbf{1}}(\boldsymbol{TX})'(\boldsymbol{TY}) \\ &= (\boldsymbol{X}'\boldsymbol{T}'\boldsymbol{TX})^{-\mathbf{1}}(\boldsymbol{X}'\boldsymbol{T}'\boldsymbol{TY}) \\ &= (\boldsymbol{X}'\boldsymbol{W}^{-\mathbf{1}}\boldsymbol{X})^{-\mathbf{1}}(\boldsymbol{X}'\boldsymbol{W}^{-\mathbf{1}}\boldsymbol{Y}). \qquad (2.32)\end{aligned}$$

2.4 Pemilihan Model Regresi Data Panel

Pengujian yang dapat dilakukan untuk memilih model yang paling tepat digunakan dalam mengestimasi data panel antara lain Uji Chow, Uji Hausman dan Uji *Lagrange Multiplier*.

2.4.1 Uji Chow

Uji Chow merupakan pengujian yang dilakukan untuk menentukan apakah model *common effect* atau model *fixed effect* yang paling tepat digunakan dalam mengestimasi data panel (Widarjono, 2009). Hipotesis yang digunakan pada Uji Chow yaitu:

H_0 : model yang digunakan adalah model *common effect*

H_1 : model yang digunakan adalah model *fixed effect*

Statistik uji yang digunakan adalah uji F, yaitu:

$$F_{hitung} = \frac{(SSE_1 - SSE_2)/(N-1)}{SSE_2/(NT-N-k)} \qquad (2.33)$$

dengan:

N = banyaknya unit *cross section*

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

k $=$ banyaknya variabel bebas

SSE_1 $=$ *Sum Square Error* yang berasal dari model *common effect*

SSE_2 $=$ *Sum Square Error* yang berasal dari model *fixed effect*.

Adapun kriteria penolakan dari hipotesis tersebut yaitu jika nilai F_{hitung} lebih besar dari nilai $F_{(N-1,NT-N-k)}$ atau nilai $p-value$ lebih kecil dari nilai α, maka hipotesis nol ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa model yang lebih tepat untuk digunakan adalah model *fixed effect*. Sebaliknya, jika nilai F_{hitung} lebih kecil dari nilai $F_{(N-1,NT-N-k)}$ atau nilai $p-value$ lebih besar dari nilai α, maka hipotesis nol tidak ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa model *common effect* yang lebih tepat untuk digunakan.

2.4.2 Uji Hausman

Menurut Gujarati (2012), Uji Hausman adalah pengujian yang dilakukan untuk mengetahui apakah model *random effect* atau model *fixed effect* yang paling tepat digunakan untuk mengestimasi data panel. Adapun hipotesis yang digunakan dalam Uji Hausman adalah:

H_0 : model yang digunakan adalah model *random effect*

H_1 : model yang digunakan adalah model *fixed effect*

Statistik uji yang digunakan mengikuti *chi-squared's distribution* berdasarkan kriteria Wald yaitu:

$$W = (\hat{\beta}_{MFE} - \hat{\beta}_{MRE})'\left[var\left(\hat{\beta}_{MFE} - \hat{\beta}_{MRE}\right)\right]^{-1}(\hat{\beta}_{MFE} - \hat{\beta}_{MRE}) \qquad (2.34)$$

dengan:

$\hat{\beta}_{MFE}$ $=$ vektor estimasi kemiringan (*slope*) model *fixed effect*

$\hat{\beta}_{MRE}$ $=$ vektor estimasi kemiringan (*slope*) model *random effect*

Kriteria penolakan pada hipotesis tersebut yaitu jika nilai W lebih kecil dari nilai $\chi^2_{(\alpha,k)}$ atau nilai $p - value$ lebih besar dari nilai α, maka hipotesis nol tidak ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa model *random effect* yang lebih tepat untuk digunakan. Begitupun sebaliknya, jika nilai W lebih besar dari nilai $\chi^2_{(\alpha,k)}$ atau nilai $p - value$ lebih kecil dari nilai α, maka hipotesis nol ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa model yang lebih tepat untuk digunakan yaitu model *fixed effect*.

2.4.3 Uji *Lagrange Multiplier*

Uji *Lagrange Multiplier* (LM) merupakan uji yang dilakukan untuk melihat apakah model *common effect* atau model *random effect* yang paling tepat digunakan dalam mengestimasi data panel (Basuki & Prawoto, 2016). Hipotesis yang digunakan untuk Uji *Lagrange Multiplier* adalah sebagai berikut:

H_0 : model yang digunakan adalah model *common effect*

H_1 : model yang digunakan adalah model *random effect*

Statistik uji yang digunakan adalah:

$$LM = \frac{NT}{2(T-1)}\left[\frac{\sum_{i=1}^{N}(T\bar{\hat{u}}_i)^2}{\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{u}_{it}^2} - 1\right] \tag{2.35}$$

dengan:

$N =$ jumlah unit *cross section*

$T =$ jumlah periode waktu (unit *time series*)

$u =$ nilai residual model *common effect*.

Adapun kriteria penolakan dari hipotesis tersebut yakni jika nilai LM lebih besar dari nilai $\chi^2_{(\alpha,1)}$ atau nilai $p - value$ lebih kecil dari nilai α, maka hipotesis nol ditolak.

Akibatnya, cukup bukti untuk menyatakan bahwa model yang paling tepat untuk digunakan yaitu model *random effect*. Demikian pula, jika nilai LM lebih kecil dari nilai $\chi^2_{(\alpha,1)}$ atau nilai $p - value$ lebih besar dari nilai α, maka hipotesis nol tidak ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa model *common effect* yang lebih tepat untuk digunakan.

2.5 Pengujian Asumsi Klasik pada Model Regresi Data Panel

Regresi data panel memberikan alternatif model berupa model *common effect*, model *fixed effect* dan model *random effect*. Uji asumsi klasik yang meliputi uji normalitas, multikolinearitas, heteroskedastisitas dan autokorelasi dilakukan dengan tujuan untuk mengetahui apakah model sudah memenuhi asumsi-asumsi tersebut atau belum.

2.5.1 Uji Normalitas

Uji normalitas bertujuan untuk mengetahui apakah sebaran dari data yang telah diperoleh mengikuti distribusi normal atau diambil dari populasi yang berdistribusi normal (Ghozali, 2016). Untuk membuktikan apakah suatu data berdistribusi normal atau tidak, terdapat banyak cara yang dapat dilakukan.

Pengujian normalitas pada data dapat dilakukan baik secara visual maupun dengan menggunakan analisis statistik. Secara visual, pengujian dapat dilakukan dengan cara menampilkan *histogram* dan *boxplot*. Sementara itu, dengan analisis statistik dapat menggunakan Uji Jarque Bera (JB) dengan rumus sebagai berikut:

$$JB = \frac{n}{6}\left[S^2 + \frac{(Kr-3)^2}{4}\right] \tag{2.36}$$

$$S = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}^2)\right]^{\frac{3}{2}}} \tag{2.37}$$

$$Kr = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}^2)\right]^{2}} \tag{2.38}$$

dengan:

n = banyaknya data pengamatan

S = *skewness*

Kr = *kurtosis*.

Hipotesis yang digunakan pada uji normalitas yaitu:

H_0 : residual berdistribusi normal

H_1 : residual tidak berdistribusi normal

Kriteria penolakan dari uji hipotesis tersebut yaitu jika nilai JB lebih kecil dari nilai $\chi^2_{(\alpha,2)}$ atau nilai $p - value$ lebih besar dari nilai α, maka hipotesis nol tidak ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa residual berdistribusi normal. Sebaliknya, jika nilai JB lebih besar dari nilai $\chi^2_{(\alpha,2)}$ atau nilai $p - value$ lebih kecil dari nilai α, maka hipotesis nol ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa residual tidak berdistribusi normal.

2.5.2 Uji Multikolinearitas

Tujuan dilakukannya uji multikolinearitas yaitu untuk menguji apakah ditemukan adanya korelasi (hubungan yang kuat) antar variabel bebas pada suatu model regresi. Jika pada model regresi menggunakan lebih dari satu variabel bebas, maka perlu dilakukan uji multikolineritas. Model regresi yang baik adalah model yang tidak terdapat korelasi yang kuat antar variabel bebas. Adanya korelasi yang kuat antar variabel bebas sangatlah tidak dianjurkan, karena akan berdampak pada keakuratan pendugaan parameter (koefisien regresi) dalam memperkirakan nilai yang sebenarnya.

Untuk mengetahui ada atau tidaknya multikolinearitas, dapat dilakukan dengan cara:

1. Memeriksa nilai *Variance Inflation Factor* (VIF) dari masing-masing variabel bebas. Jika terdapat terdapat nilai $VIF > 10$, maka dapat diindikasikan adanya multikolinearitas. Rumus VIF dinyatakan sebagai berikut:

$$VIF = \frac{1}{1-R_j^2} \; ; \quad j = 1,2,\dots,k \tag{2.39}$$

   dengan R_j^2 merupakan koefisien determinasi yang diperoleh dari variabel bebas X_j yang diregresikan terhadap variabel bebas lainnya. Apabila variabel bebas X_j tidak berkorelasi dengan variabel bebas lainnya, maka nilai R_j^2 akan kecil dan nilai VIF akan mendekati 1. Begitupun sebaliknya, jika variabel bebas X_j berkorelasi dengan variabel bebas lainnya, maka nilai R_j^2 akan mendekati 1 dan nilai VIF akan menjadi besar (Montgomery & Peck, 1992).
2. Melihat kekuatan korelasi antar variabel bebas dengan menggunakan matriks korelasi. Jika terdapat korelasi antar variabel bebas yang nilainya lebih besar dari 0.08, maka dapat diindikasikan adanya multikolinearitas (Ghozali, 2016).

2.5.3 Uji Heteroskedastisitas

Uji heteroskedastisitas dilakukan untuk melihat apakah residual dari model yang terbentuk memiliki varians yang konstan atau tidak (Kuncoro, 2011). Uji heteroskedastisitas sangatlah penting dilakukan pada model yang terbentuk. Model yang baik adalah model yang memiliki varians dari setiap residualnya konstan. Heteroskedastisitas adalah keadaan dimana asumsi tersebut tidak tercapai. Dengan adanya heteroskedastisitas akan mengakibatkan pendugaan parameternya tidak efisien (Nachrowi & Usman, 2006). Heteroskedastisitas dapat dideteksi dengan menggunakan Uji Breusch-Pagan (BP). Pengujian dilakukan dengan meregresikan residu kuadrat (sebagai variabel terikat) dengan variabel bebas model asli. Misalkan diberikan model regresi linear dengan variabel bebas sebanyak i:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{2.40}$$

dengan ε_i sebagai residu. Pada Metode Kuadrat Terkecil, *error* memiliki rata-rata 0 dan dengan asumsi ragamnya tidak bergantung pada variabel bebas. Estimasi ragam dapat diperoleh dari rata-rata nilai kuadrat *error*. Apabila asumsi tidak dianggap benar, maka ragam dari model regresi terkait secara linear dengan variabel bebasnya. Model tersebut dapat diperiksa dengan cara meregresikan kuadrat residu dengan variabel bebasnya, sehingga diperoleh persamaan:

$$\varepsilon_i^2 = \alpha_0 + \alpha_1 X_i + v_i \tag{2.41}$$

Adapun hipotesis yang digunakan adalah:

H_0 : tidak terdapat heteroskedastisitas pada data

H_1 : terdapat heteroskedastisitas pada data

Adapun kriteria penolakan dari hipotesis tersebut yaitu jika nilai BP lebih kecil dari nilai $\chi^2_{(\alpha,k)}$ atau nilai $p - value$ lebih besar dari nilai α, maka hipotesis nol tidak ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa tidak terdapat heteroskedastisitas pada data. Sebaliknya, jika nilai BP lebih besar dari nilai $\chi^2_{(\alpha,k)}$ atau nilai $p - value$ lebih kecil dari nilai α, maka hipotesis nol ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa terdapat heteroskedastisitas pada data.

2.5.4 Uji Autokorelasi

Uji autokorelasi dilakukan untuk mengetahui adakah korelasi yang terjadi antara residual pada suatu pengamatan dengan pengamatan yang lainnya (Ghozali, 2016). Untuk mengetahui ada atau tidaknya autokorelasi dapat dilakukan dengan menggunakan Uji Durbin Watson. Hipotesis yang digunakan adalah:

H_0 : tidak terdapat autokorelasi pada data

H_1 : terdapat autokorelasi pada data

Statistik uji yang digunakan yaitu:

$$d = \frac{\sum_{t=2}^{n}(\hat{e}_t - \hat{e}_{t-1})^2}{\sum_{t=1}^{n}\hat{e}_t^{\,2}} \quad (2.42)$$

dengan:

d = nilai Durbin Watson *Statistic*

e_t = nilai residual periode t

e_{t-1} = nilai residual periode $t-1$ (sebelumnya).

Tabel 1. Uji Statistik Durbin Watson

Hipotesis Nol (H_0)	Keputusan	Nilai Statistik d
Terdapat autokorelasi positif	Hipotesis nol (H_0) ditolak	$0 < d < d_L$
Tidak terdapat autokorelasi positif	Tidak ada keputusan	$d_L \leq d \leq d_U$
Terdapat autokorelasi negatif	Hipotesis nol (H_0) ditolak	$4 - d_L \leq d \leq 4$
Tidak terdapat autokorelasi negatif	Tidak ada keputusan	$4 - d_U \leq d \leq 4 - d_L$

Tabel 1. Lanjutan

Hipotesis Nol (H_0)	Keputusan	Nilai Statistik d
Tidak terdapat autokorelasi	Hipotesis nol (H_0) tidak ditolak	$d_U \leq d \leq 4 - d_U$

dengan:

$d_L =$ nilai Durbin *Lower*,

$d_U =$ nilai Durbin *Upper*.

Adapun kriteria penolakan pada hipotesis tersebut yaitu jika nilai $p - value$ lebih besar dari nilai α, maka hipotesis nol tidak ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa tidak terdapat autokorelasi pada data. Begitu juga sebaliknya, jika nilai $p - value$ lebih kecil dari nilai α, maka hipotesis nol ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa terdapat autokorelasi pada data.

2.6 Uji Kelayakan (*Goodness of Fit*) Model Regresi Data Panel

Terdapat dua uji yang digunakan untuk menguji signifikansi koefisien regresi yaitu Uji F (Uji Simultan) dan Uji t (Uji Parsial).

2.6.1 Uji F (Uji Simultan)

Uji statistik F pada dasarnya dilakukan untuk menunjukkan apakah seluruh variabel bebas yang dimasukkan ke dalam model regresi mempunyai pengaruh secara simultan terhadap variabel terikat (Ghozali, 2016). Uji F digunakan untuk menguji koefisien kemiringan (*slope*) pada model regresi secara bersamaan. Dengan demikian, hipotesisnya adalah sebagai berikut:

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

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

Adapun statistik uji yang digunakan adalah:

$$F_{hitung} = \frac{R^2/(N+k-1)}{(1-R^2)/(NT-N-k)} \quad (2.43)$$

dengan:

N = banyaknya unit *cross section*

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

k = banyaknya variabel bebas

Kriteria penolakan dari hipotesis tersebut yaitu jika nilai F_{hitung} lebih besar dari nilai $F_{(N+k-1,NT-N-k)}$ atau nilai $p-value$ lebih kecil dari nilai α, maka hipotesis nol ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa secara simultan variabel bebas memberikan pengaruh signifikan terhadap variabel terikat. Sebaliknya, jika nilai F_{hitung} lebih kecil dari nilai $F_{(N+k-1,NT-N-k)}$ atau nilai $p-value$ lebih besar dari nilai α, maka hipotesis nol tidak ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa secara simultan variabel bebas tidak memberikan pengaruh signifikan terhadap variabel terikat.

2.6.2 Uji t (Uji Parsial)

Pada dasarnya, uji statistik t digunakan untuk mengetahui apakah suatu variabel bebas memberikan pengaruh signifikan secara parsial atau individual terhadap variabel terikat (Gujarati, 2012). Hipotesis dalam uji ini adalah sebagai berikut:

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

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

Uji t didefinisikan sebagai berikut:

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \tag{2.44}$$

dengan:

$\hat{\beta}_j$ = penduga parameter ke-j yang dihipotesiskan ($j = 1,2,\ldots,k$)

$SE(\hat{\beta}_j)$ = *Standard Error* $\hat{\beta}_j$

Adapun kriteria penolakan dari hipotesis tersebut yaitu jika nilai $|t_{hitung}|$ lebih besar dari nilai $t_{(\frac{\alpha}{2},NT-N-k)}$ atau nilai $p-value$ lebih kecil dari nilai α, maka hipotesis nol ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa secara individual variabel bebas memberikan pengaruh signifikan terhadap variabel terikat. Sebaliknya, jika nilai $|t_{hitung}|$ lebih kecil dari nilai $t_{(\frac{\alpha}{2},NT-N-k)}$ atau nilai $p-value$ lebih besar dari nilai α, maka hipotesis nol tidak ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa secara simultan variabel bebas tidak memberikan pengaruh signifikan terhadap variabel terikat.

2.7 Koefisien Determinasi

Nilai koefisien determinasi (R^2) digunakan untuk mengetahui seberapa besar pengaruh variabel bebas (X) terhadap variabel terikat (Y) atau seberapa besar variasi dari variabel terikat dapat dijelaskan oleh variabel bebas. Apabila nilai koefisien determinasi mendekati angka 1 (satu), berarti kemampuan variabel bebas dalam menerangkan variabel terikat cukup baik. Sebaliknya, jika nilai koefisien determinasi mendekati angka 0 (nol), berarti kemampuan variabel bebas dalam menerangkan variabel terikat cukup terbatas (Kuncoro, 2011). Dengan demikian, baik atau buruknya suatu model regresi ditentukan oleh nilai koefisien determinasi yang terletak antara 0 dan 1.

$$R^2 = 1 - \frac{\sum e_i^2/(n-k)}{\sum(Y_i-\bar{Y})^2/(n-1)} \tag{2.45}$$

dengan:

$k =$ jumlah parameter (termasuk intersep)

$n =$ jumlah observasi.

2.8 Indeks Pembangunan Manusia (IPM)

Indeks Pembangunan Manusia (IPM) merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk). IPM dapat menentukan peringkat atau level pembangunan suatu wilayah/negara (Badan Pusat Statistik Provinsi Lampung, 2020).

Indeks Pembangunan Manusia mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup. Sebagai ukuran kualitas hidup, IPM dibangun melalui pendekatan tiga dimensi dasar. Dimensi tersebut mencakup umur panjang dan sehat, pengetahuan dan kehidupan yang layak. Ketiga dimensi tersebut memiliki pengertian sangat luas karena terkait banyak faktor.

a. Angka Harapan Hidup

   Angka Harapan Hidup (AHH) pada waktu lahir merupakan rata-rata perkiraan banyaknya tahun yang dapat ditempuh oleh seseorang selama hidup. Angka harapan hidup merupakan alat untuk mengevaluasi kinerja pemerintah dalam meningkatkan kesejahteraan penduduk pada umumnya dan meningkatkan derajat kesehatan pada khususnya. Angka harapan hidup yang rendah di suatu daerah harus diikuti dengan program pembangunan kesehatan dan program sosial lainnya termasuk kesehatan lingkungan, kecukupan gizi dan kalori serta program pemberantasan kemiskinan.

b. Rata-rata Lama Sekolah

   Rata-rata Lama Sekolah (RLS) / *Mean Years School* (MYS) didefinisikan sebagai jumlah tahun yang digunakan oleh penduduk dalam menjalani pendidikan formal. Rata-rata lama sekolah dapat digunakan untuk mengetahui kualitas pendidikan masyarakat dalam suatu wilayah. Penduduk yang tamat SD diperhitungkan lama sekolah selama 6 tahun, tamat SMP diperhitungkan lama sekolah selama 9 tahun dan tamat SMA diperhitungkan lama sekolah selama 12 tahun tanpa memperhitungkan apakah pernah tinggal kelas atau tidak.

c. Pengeluaran Perkapita Disesuaikan

   Pengeluaran perkapita adalah biaya yang dikeluarkan untuk konsumsi semua anggota rumah tangga selama sebulan dibagi dengan banyaknya anggota rumah tangga. Data pengeluaran dapat mengungkap tentang pola konsumsi rumah tangga secara umum menggunakan indikator proporsi pengeluaran untuk makanan dan non makanan. Komposisi pengeluaran rumah tangga dapat dijadikan ukuran untuk menilai tingkat kesejahteraan ekonomi penduduk, makin rendah persentase pengeluaran untuk makanan terhadap total pengeluaran makin membaik tingkat kesejahteraan.

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester genap tahun akademik 2020/2021 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan pada penelitian ini merupakan data sekunder yaitu data Indeks Pembangunan Manusia di Provinsi Lampung pada tahun 2016-2018 yang diperoleh dari *website* BPS Provinsi Lampung. Dengan variabel angka harapan hidup (X1), rata-rata lama sekolah (X2) dan pengeluaran perkapita disesuaikan (X3) sebagai variabel bebas, sedangkan variabel IPM (Y) sebagai variabel terikat. Pengolahan data pada penelitian ini dibantu menggunakan *software* R.

Tabel 2. Data Penelitian

No.	Kabupaten/Kota	Tahun	Y	X1	X2	X3
1.	Tulang Bawang	2016	66.74	69.28	7.12	10034
		2017	67.07	69.41	7.15	10098
		2018	67.7	69.59	7.22	10553

Tabel 2. Lanjutan

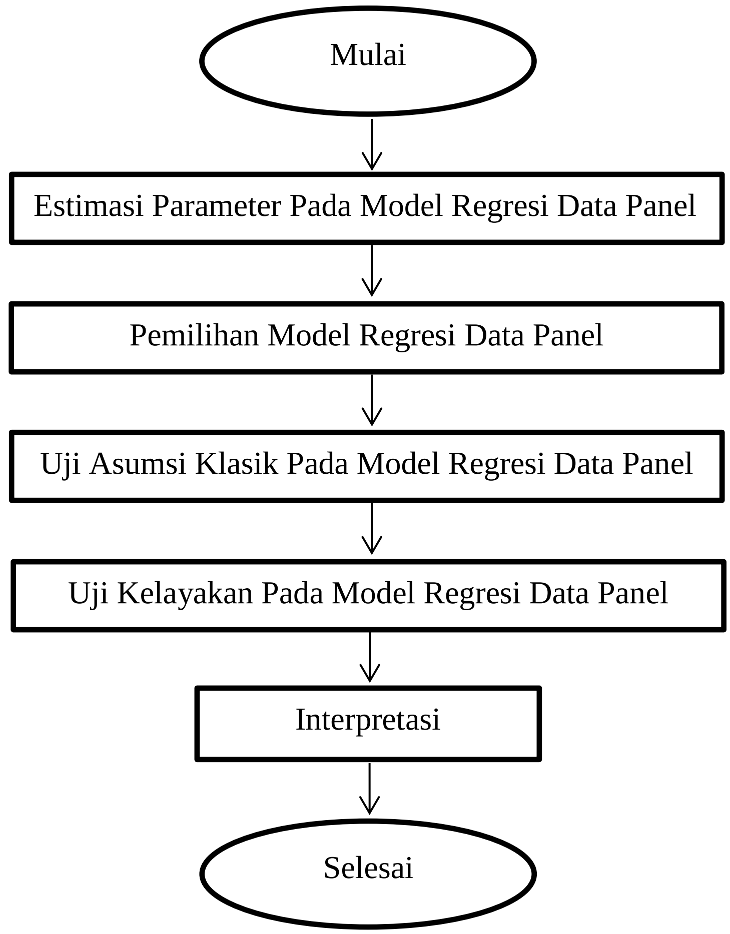
No.	Kabupaten/Kota	Tahun	Y	X1	X2	X3
2.	Bandar Lampung	2016	75.34	70.75	10.88	11266
		2017	75.98	70.84	10.89	11699
		2018	76.63	71.01	10.9	11952
3.	Tanggamus	2016	64.41	67.61	6.87	8483
		2017	64.94	67.8	6.88	8661
		2018	65.67	68.04	6.96	9107
4.	Pringsewu	2016	68.26	68.88	7.84	9533
		2017	68.61	69.14	7.85	9731
		2018	69.42	69.44	8.01	10190
5.	Metro	2016	75.45	71.05	10.56	11007
		2017	75.87	71.13	10.57	11397
		2018	76.22	71.29	10.61	11636
6.	Mesuji	2016	60.72	67.32	6.13	7099
		2017	61.87	67.49	6.39	7319
		2018	62.88	67.71	6.6	7774
7.	Pesisir Barat	2016	61.5	62.29	7.48	7616
		2017	62.2	62.54	7.58	7890
		2018	62.96	62.85	7.59	8355
8.	Lampung Timur	2016	67.88	69.92	7.55	9416
		2017	68.05	70.11	7.56	9453
		2018	69.04	70.31	7.57	9908
9.	Pesawaran	2016	63.47	68.05	7.24	7055
		2017	64.43	68.29	7.45	7449
		2018	64.97	68.53	7.47	7724
10.	Lampung Tengah	2016	68.33	69.15	7.37	10674
		2017	68.95	69.28	7.38	10820
		2018	69.73	69.46	7.51	11052

Karena pada penelitian ini akan ditentukan model regresi data panel dengan menggunakan pendekatan *random effect model*, maka dipilih secara acak 10 dari 15 kabupaten/kota di Provinsi Lampung.

3.3 Metode Penelitian

Langkah-langkah dalam penelitian ini adalah sebagai berikut:

1. Melakukan estimasi parameter pada model regresi data panel dengan pendekatan *Common Effect* Model (CEM) menggunakan metode *Ordinary Least Square* (OLS), *Fixed Effect Model* (FEM) menggunakan metode *Least Square Dummy Variable* (LSDV) dan *Random Effect Model* (REM) menggunakan metode *Generalized Least Square* (GLS).
2. Melakukan pemilihan model regresi data panel dengan menggunakan Uji Chow dan Uji Hausman.
3. Melakukan uji asumsi klasik pada model regresi data panel yang meliputi:
    a. Uji normalitas dengan menggunakan Uji Jarque Bera.
    b. Uji multikolineritas dengan melihat nilai VIF dari setiap variabel bebas.
    c. Uji heteroskedastisitas dengan menggunakan Uji Breusch Pagan.
    d. Uji autokorelasi dengan menggunakan Uji Durbin Watson.
4. Melakukan uji kelayakan (*goodness of fit*) pada model regresi data panel yaitu dengan Uji F (Uji Simultan), Uji t (Uji Parsial) dan koefisien determinasi.
5. Interpretasi pada model regresi data panel.



Gambar 1. Diagram Tahapan Penelitian

V. KESIMPULAN

Berdasarkan uraian hasil dan pembahasan pada penelitian ini, diperoleh kesimpulan bahwa model regresi data panel yang lebih sesuai untuk memodelkan Indeks Pembangunan Manusia (IPM) di Provinsi Lampung dari tahun 2016-2018 adalah *Random Effect Model* (REM), dengan model persamaan hasil estimasi adalah $\hat{Y}_{it} = 8.383762 + 0.500353\,X_{1it} + 1.615153\,X_{2it} + 0.001290\,X_{3it}$. Dengan diperoleh nilai *R-Squared* sebesar 0.9856, artinya variabel angka harapan hidup, rata-rata lama sekolah dan pengeluaran perkapita disesuaikan mampu menjelaskan variabel Indeks Pembangunan Manusia (IPM) sebesar 98.56%, sedangkan sisanya yakni sebesar 1.44% dijelaskan oleh variabel lain di luar model.

DAFTAR PUSTAKA

Badan Pusat Statistik Provinsi Lampung. 2020. Indeks Pembangunan Manusia. https://lampung.bps.go.id/subject/26/indeks-pembangunan-manusia.html#subjekViewTab1. Diakses pada 9 Desember 2020.

Basuki, A. T. & Prawoto, N. 2016. *Analisis Regresi Dalam Penelitian Ekonomi & Bisnis*. Rajawali Pers, Jakarta.

Draper, N. R. & Smith, H. 1998. *Applied Regression Analysis*. 3th Edition. John Wiley & Sons, New York.

Ghozali, I. 2016. *Aplikasi Analisis Multivariate dengan Program IBM SPSS 23*. Ed. ke-8. Badan Penerbit Universitas Diponegoro, Semarang.

Gujarati, D. N. 2003. *Basic Econometric*. 4th Edition. The McGraw-Hill Companies, New York.

Gujarati, D. N. 2012. *Dasar-Dasar Ekonometrika*. Ed. ke-5. Terjemahan R. C. Mangunsong. Salemba Empat, Jakarta.

Greene, W. H. 2007. *Econometric Analysis*. 6th Edition. Prentice Hall International, New Jersey.

Hsiao, C. 1992. *Panel Analysis for Metric Data*. Departement of Economics University of Southern California, Los Angeles.

Hsiao, C. 2003. *Analysis of Panel Data*. 2nd Edition. Cambridge University Press, California.

Kuncoro, M. 2011. *Metode Kuantitatif Teori dan Aplikasi Untuk Bisnis & Ekonomi*. Ed. ke-4. UPP STIM YKPN, Yogyakarta.

Montgomery, D. C. & Peck, E. A. 1992. *Introduction to Linear Regression Analysis*. 2nd Edition. John Wiley & Sons, Toronto.

Nachrowi, N. D. & Usman, H. 2006. *Pendekatan Populer dan Praktis Ekonometrika Untuk Analisis Ekonomi dan Keuangan*. Fakultas Ekonomi Universitas Indonesia, Jakarta.

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.

Widarjono. A. 2009. *Ekonometrika Pengantar dan Aplikasinya*. Ed. ke-3. Ekonisia Fakultas Ekonomi Universitas Islam Indonesia, Yogyakarta.