

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **A. Tujuan Penelitian**

Berdasarkan permasalahan yang telah peneliti rumuskan, penelitian ini bertujuan untuk memperoleh ilmu pengetahuan yang valid dan reliable berdasarkan data yang telah diperoleh mengenai:

1. Mengetahui bagaimana pengaruh pendidikan terhadap ketimpangan pembangunan antar wilayah di Indonesia
2. Mengetahui bagaimana pengaruh kesehatan terhadap ketimpangan pembangunan antar wilayah di Indonesia
3. Mengetahui bagaimana pengaruh infrastruktur terhadap ketimpangan pembangunan antar wilayah di Indonesia

#### **B. Objek dan Ruang Lingkup Penelitian**

Objek dalam penelitian ini adalah ketimpangan pembangunan wilayah dan berbagai faktor yang mempengaruhi ketimpangan pembangunan wilayah di Indonesia bagian Tengah yang berjumlah 34 provinsi di Indonesia.

Ruang lingkup penelitian ini berkisar antara tahun 2017-2019 yang menggunakan data indeks Williamson untuk variabel ketimpangan pembangunan.

Sedangkan untuk variabel bebas pendidikan (X1) menggunakan data rata-rata lama sekolah dan (X2) angka buta huruf. Variabel kesehatan (X3) menggunakan data angka harapan hidup. Dan untuk variabel infrastruktur (X4) menggunakan data rasio kerapatan jalan.

### **C. Metode Penelitian**

Metode penelitian yang peneliti gunakan adalah menggunakan metode penelitian kuantitatif dengan pendekatan deskriptif. Menurut Sugiyono (2014) metode deskriptif merupakan metode yang digunakan untuk mengetahui nilai variabel-variabel yang diteliti dengan mandiri atau tanpa menghubungkan variabel satu dengan variabel lainnya.

Penelitian ini memiliki lima variabel yang dijadikan objek, antara lain Ketimpangan Pembangunan Wilayah yang merupakan variabel dependen (Y) dan variabel independennya antara lain rata-rata lama sekolah (X1), angka buta huruf (X2), angka harapan hidup (X3), dan kerapatan jalan (X4)

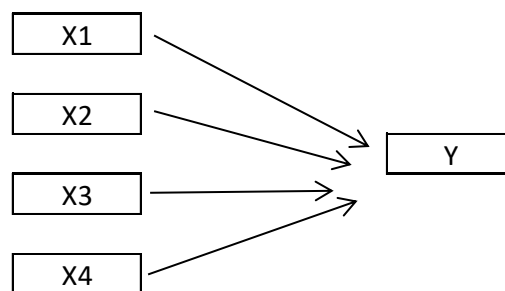
### **D. Konstelasi antar Hubungan**

Mengacu pada penjelasan di atas maka peneliti dapat mengajukan bahwa:

1. Variabel pendidikan melalui rata-rata lama sekolah (X1) memberikan pengaruh negatif dan signifikan terhadap Ketimpangan Pembangunan (Y)

2. Variabel pendidikan melalui angka buta huruf (X2) memberikan pengaruh positif dan signifikan terhadap Ketimpangan Pembangunan (Y)
3. Variabel kesehatan melalui angka harapan hidup (X3) memberikan pengaruh negatif dan signifikan terhadap Ketimpangan Pembangunan (Y)
4. Variabel infrastruktur melalui kerapatan Jalan (X4) memberikan pengaruh negatif dan signifikan terhadap Ketimpangan Pembangunan (Y)

Pengaruh antara variabel X dan Y dalam penelitian ini terbambar dalam konstelasi di bawah ini



Gambar 3.1  
Konstelasi Penelitian

Keterangan:

X1 = Rata-rata lama sekolah

X2 = Angka buta huruf

X2 = Angka Harapan Hidup

X3 = Kerapatan Jalan

Y = Ketimpangan Pembangunan Antar Wilayah

### **E. Jenis dan Sumber Data**

Data yang digunakan dalam penelitian ini adalah data sekunder, baik untuk variabel terikat yaitu ketimpangan pembangunan dan variabel bebas yaitu rata-rata lama sekolah, angka buta huruf, angka harapan hidup dan kerapatan jalan. Dalam melakukan penelitian ini peneliti menggunakan data sekunder. Untuk memperoleh data antara lain pertumbuhan ekonomi, jumlah penduduk, angka harapan hidup, dan rata-rata lama, angka buta huruf sekolah, panjang jalan dan luas wilayah tahun 2016-2019 yang terdiri dari 34 provinsi pada wilayah Indonesia peneliti mendapatkannya dari Badan Pusat Statistik, baik BPS Indonesia maupun BPS masing-masing provinsi.

### **F. Operasional Variabel Data**

Definisi operasional memiliki tujuan untuk mencegah adanya *missundertanding* penafsiran yang berhubungan dengan berbagai istilah dalam penelitian ini. Berdasarkan judul penelitian yakni “Pengaruh Pendidikan, Kesehatan, dan Infrastruktur terhadap Ketimpangan Pembangunan Antar Wilayah di Indonesia” maka berikut definisi operasionalnya, antara lain:

## **1. Ketimpangan Pembangunan**

### **a. Definisi Konseptual**

Ketimpangan pembangunan wilayah merupakan merupakan keadaan umum yang terjadi di setiap wilayah di berbagai negara. Badan Pusat Statistik (BPS) mengungkapkan bahwa ketimpangan pembangunan lebih mengarah pada perbedaan yang terjadi antar daerah (Badan Pusat Statistik, 2017). Menurut Sirojuzilam (2019) ketimpangan pembangunan wilayah dapat terjadi apabila adanya perbedaan pertumbuhan ekonomi antar wilayah. Perbedaan pertumbuhan ekonomi yang dipengaruhi berbagai faktor seperti investasi, pengeluaran pemerintah, pendidikan, transportasi, aglomerasi industri, dan budaya dapat menyebabkan adanya kesenjangan antar wilayah.

Bedasarkan pemaparan di atas dapat diambil kesimpulan bahwa ketimpangan pembangunan wilayah merupakan perbedaan yang terjadi antar berbagai wilayah yang disebabkan oleh berbagai macam faktor baik ekonomi, pendidikan, dan budaya.

### **b. Definisi Oprasional**

Ketimpangan pembangunan wilayah diukur menggunakan indeks Williamson. Dalam menghitung indeks Williamson dibutuhkan data PDRB (juta rupiah) dan jumlah penduduk (jiwa) per daerah sebagai dasarnya. Data di peroleh dari Badan Pusat Statistik setiap wilayahnya pada tahun 2016-2019.

## **2. Rata-Rata Lama Sekolah**

### **a. Definisi Konseptual**

Menurut BPS rata-rata lama sekolah merupakan rata-rata jumlah tahun yang ditempuh penduduk yang berusia 15 tahun keatas dalam suatu wilayah dalam menjalani pendidikan formal.

b. Definisi Operasional

Rata-rata lama sekolah dihitung menggunakan data lama sekolah penduduk ke-I yang berusia 15 tahun dan jumlah penduduk dengan usia 15 tahun ke atas. Data rata-rata lama sekolah disajikan dalam bentuk tahun. Dalam perhitungan rata-rata lama sekolah penduduk yang tamat SD dihitung lama sekolah selama 6 tahun, penduduk yang tamat SMP dihitung lama sekolah selama 9 tahun, dan penduduk yang tamat SMA akan diperhitungkan lama sekolah selama 12 tahun tanpa memperhitungkan tinggal kelas atau tidak.

### **3. Angka Buta Huruf**

a. Definisi Konseptual

Menurut BPS angka buta huruf adalah proporsi penduduk berusia 15 tahun ke atas yang tidak dapat membaca dan menulis kalimat sederhana dalam huruf latin, huruf arab, dan huruf lainnya ( seperti huruf jawa, kanji, dll) terhadap penduduk berusia 15 tahun ke atas.

b. Definisi Oprasional

Angka buta huruf dihitung berdasarkan jumlah penduduk berusia 15 tahun ke atas yang tidak bisa membaca dan menulis dan dibandingkan dengan jumlah penduduk berusia 15 tahun ke atas. Data angka buta huruf berbentuk persentase

yang berkisar antara 0-100. Semakin mendekati angka 100 tingkat buta huruf pada suatu wilayah semakin banyak.

#### **4. Angka Harapan Hidup**

##### **a. Definisi Konseptual**

Menurut BPS angka harapan hidup merupakan rata-rata tahun hidup yang masih akan dijalani oleh seseorang yang telah berhasil mencapai umur X, pada suatu tahun tertentu, dalam mortalitas yang berlaku di lingkungan masyarakatnya.

##### **a. Definisi Operasional**

Angka harapan hidup merupakan indikator tingkat kesehatan yang datanya disajikan dalam bentuk tahun. Dalam menghitung angka harapan hidup diperoleh dari angka kematian menurut umur yang datanya diperoleh dari catatan registrasi kematian secara bertahun-tahun.

#### **5. Kerapatan Jalan**

##### **a. Definisi Konseptual**

Kerapatan jalan merupakan ukuran rasio antara panjang jalan di sebuah wilayah dan luas wilayah. Kerapatan jalan yang tinggi di sebuah wilayah menandakan wilayah tersebut semakin lancar arus distribusi barang maupun jasa.

##### **b. Definisi Operasional**

Dalam mengukur kerapatan jalan diperlukan data jumlah panjang jalan di sebuah wilayah dan luas wilayah tersebut. Data kerapatan jalan berbentuk rasio. Semakin besar nilai rasio kerapatan jalan maka semakin baik infrastruktur jalan tersebut.

## G. Teknik Pengumpulan Data

Dalam proses pengumpulan data pada penelitian ini menggunakan teknik studi dokumenter. Dokumen ialah fakta-fakta dan berbagai data yang sengaja di simpan sebagai bentuk dokumentasi atau bukti dari sebuah kegiatan yang bersifat tertulis. Data-data yang tersimpan bisa berupa surat, catatan , dan lain sebagainya. Dokumen yang digunakan dalam penelitian ini diperoleh dari Badan Pusat Statistik (BPS) yang mencakup sebanyak 34 provinsi di seluruh wilayah Republik Indonesia.

## H. Teknik Analisis Data

### 1. Model Data Panel Analisis Linear Berganda

Teknik analisis data yang digunakan pada penelitian ini adalah analisis regresi berganda menggunakan jenis data panel dengan bantuan program *E-Views 10*. Menurut Gujarati (2004) data panel merupakan sebuah data panel yang menggabungkan antara data *time series* dan *cross-section*.

Adapun model regresi data panel untuk mengamati hubungan kuantitatif antara variabel bebas yakni rata-rata lama sekolah, angka buta huruf, angka melek huruf, dan kerapatan jalan terhadap variabel terikat yaitu ketimpangan pembangunan antar wilayah. Dengan persamaan sebagai berikut:

$$Y_{it} = \alpha + \beta_1 X_{it} + \beta_2 X_{it} + \beta_3 X_{it} + \dots + \beta_3 X_{it} + e_{it}$$



Keterangan:

$Y_{it}$  = Variabel terikat (dependent)

$X_{it}$  = Variabel bebas (independent)

$\alpha$  = Kostanta

$\beta$  = Koefisien regresi

e = Error

Ghozali & Ratmono (2017) mengungkapkan bahwa penelitian dengan menggunakan data panel mempunyai beberapa keuntungan dibandingkan data *cross section* atau *time series*. Di bawah ini merupakan beberapa keuntungannya antara lain:

- a. Penelitian yang menggunakan data panel dapat meningkatkan derajat kebebasan karena dibutuhkan data yang bervariasi sehingga dapat mengurangi kolineritas antar variabel dependent dan variabel independen.
- b. Dibandingkan dengan data *time series* dan *cross section*, data panel dapat memberikan banyak informasi.
- c. Data panel lebih baik dalam melakukan analisis data sampai peramalan dan pengambilan kesimpulan (inferensi) dibandingkan dengan data *cross section*.

## 2. Metode Estimasi Data Panel

Dalam menggunakan data panel, terdapat tiga metode estimasi regresi yang dapat digunakan, yaitu:

### a. *Common Effect Model* (CEM)

Metode CEM merupakan pendekatan model data panel yang paling sederhana. Pada model ini hanya mengkombinasikan data *time series* dan *cross section*. Model ini tidak melihat baik dimensi waktu ataupun individu, sehingga dapat diasumsikan bahwa perilaku data perusahaan sama dalam berbagai kurun waktu dan individu tersebut. Metode *common effect* yang mengkombinasikan data *time series* dan *cross section* dapat memakai Ordinary Least Square (OLS) dalam pendekatannya.

### b. *Fixed Effect Model* (FEM)

Model ini mengasumsikan bahwa perbedaan antar individu dapat diakomodasi dari perbedaan intersepnya. Dalam melakukan estimasi data panel, model Fixed Effects memberikan variable dummy untuk mengetahui perbedaan intersepnya. Model estimasi ini sering juga disebut dengan teknik Least Squares Dummy Variable (LSDV)..

c. *Random Effect Model (REM)*

Model ini akan mengestimasi data panel dimana variabel gangguan mungkin saling berhubungan antar waktu dan antar individu. Pada model Random Effect perbedaan intersep diakomodasi oleh error terms masing-masing objek penelitian. Keuntungan dalam menggunakan model Random Effect yaitu dapat menghilangkan heteroskedastisitas. Model ini juga disebut dengan Error Component Model (ECM) atau teknik Generalized Least Square (GLS).

Untuk menentukan model manakah yang paling tepat di antara ketiga model di atas adalah dengan melakukan beberapa uji, antara lain:

a. Uji Chow

Uji chow atau F statistic merupakan proses uji yang perlu dilakukan untuk mengetahui model manakah yang lebih tepat digunakan di antara *common effect* dan *fixed effect* (Sofyan, 2011). Adapun rumus yang uji chow ini adalah:

$$\text{CHOW} = \frac{\frac{(RRSS - URSS)}{N-1}}{\frac{URSS}{(NT - N - K)}}$$

Keterangan:

RRSS = restricted residual sum square

URSS = unrestricted residual sum square

N = jumlah data cross section

T = jumlah data time series

H0 : model menggunakan Common Effect

H1 : model menggunakan Fixed Effect

Apabila hasil dari uji chaw menunjukkan nilai F hitung lebih besar dari F tabel maka H0 ditolak. Maka, teknik regresi data panel lebih baik menggunakan model *fixed effect* dibandingkan dengan model *common effect*.

b. Uji Hausman

Uji hausman adalah uji yang digunakan untuk membandingkan manakah diantara *fixed effect* dan *random effect* yang lebih baik digunakan dalam penelitian yang di uji. Apabila hasil dari statistik hausman menyatakan nilai yang lebih besar dibandingkan dengan tabel (nilai kritis chi-square) artinya H0 ditolak dan lebih baik menggunakan *fixed effect*. Sebaliknya apabila nilai statistic hausman lebih kecil daripada nilai kritis maka lebih baik menggunakan model *random effect*.

c. Uji Breusch-Pagan

Uji Breusch-Pagan atau uji *lagrange multiplier* (LM) merupakan uji yang digunakan untuk menentukan manakah di antara *common effect* atau *random effect* yang lebih baik digunakan. *Random effect* dapat menjadi model yang tepat untuk digunakan apabila nilai dari *chi square* lebih kecil dibandingkan nilai LM hitung atau dengan kata lain H0 ditolak.

### 3. Uji Asumsi Klasik

Dalam penggunaan analisis regresi linear berganda terdapat persyaratan statistic yang harus dipenuhi, yakni dengan uji asumsi klasik dengan basis ordinary least square (OLS). Di bawah ini merupakan uji asumsi klasik antara lain(I Ghozali & Ratmono, 2013):

#### a. Uji Normalitas

Uji Normalitas digunakan untuk melihat apakah nilai residual terdistribusi normal atau tidak. Model regresi yang baik adalah memiliki nilai residual yang terdistribusi normal. Jadi uji normalitas bukan dilakukan pada masing-masing variabel tetapi pada nilai residualnya. Sering terjadi kesalahan yang jamak yaitu bahwa uji normalitas dilakukan pada masing-masing variabel. Hal ini tidak dilarang tetapi model regresi memerlukan normalitas pada nilai residualnya bukan pada masing-masing variabel penelitian. Jika residual tidak normal tetapi dekat dengan nilai kritis seperti signifikansi Kolmogorov Smirnov sebesar 0,049 maka dapat dicoba dengan metode lain yang mungkin memberikan justifikasi normal.

#### b. Uji Multikolinearitas

Uji Multikolinearitas adalah sebuah uji untuk melihat ada atau tidaknya korelasi yang tinggi antara variabel-variabel bebas dalam suatu model regresi linear berganda. Jika ada korelasi yang tinggi di antara variabel-variabel bebasnya, maka hubungan antara variabel bebas terhadap variabel terikatnya menjadi terganggu.

### c. Uji Heterokedastisitas

Uji heterokedastisitas adalah uji untuk melihat apakah terdapat ketidaksamaan varians dari residual satu ke pengamatan ke pengamatan yang lain. Model regresi yang memenuhi persyaratan adalah di mana terdapat kesamaan varians dari residual satu pengamatan ke pengamatan yang lain tetap atau disebut homoskedastisitas. Deteksi heteroskedastisitas dapat dilakukan dengan metode scatter plot dengan memplotkan nilai ZPRED (nilai prediksi) dengan SRESID (nilai residualnya). Model yang baik didapatkan jika tidak terdapat pola tertentu pada grafik, seperti mengumpul di tengah, menyempit kemudian melebar atau sebaliknya melebar kemudian menyempit. Uji statistik yang dapat digunakan adalah uji Glejser, uji Park atau uji White.

### d. Uji Autokorelasi

Uji autokorelasi adalah untuk melihat apakah terjadi korelasi antara suatu periode  $t$  dengan periode sebelumnya ( $t - 1$ ). Secara sederhana adalah bahwa analisis regresi adalah untuk melihat pengaruh antara variabel bebas terhadap variabel terikat, jadi tidak boleh ada korelasi antara observasi dengan data observasi sebelumnya. Uji autokorelasi hanya dilakukan pada data time series (runtut waktu) dan tidak perlu dilakukan pada data cross section seperti pada kuesioner di mana pengukuran semua variabel dilakukan secara serempak pada saat yang bersamaan. Beberapa uji statistik yang sering dipergunakan adalah uji Durbin-Watson, uji dengan Run Test dan jika data observasi di atas 100 data sebaiknya menggunakan uji Lagrange Multiplier. Beberapa cara untuk

menanggulangi masalah autokorelasi adalah dengan mentransformasikan data atau bisa juga dengan mengubah model regresi ke dalam bentuk persamaan beda umum (generalized difference equation). Selain itu juga dapat dilakukan dengan memasukkan variabel lag dari variabel terikatnya menjadi salah satu variabel bebas, sehingga data observasi menjadi berkurang 1.

#### 4. Uji Hipotesis

##### a. Uji Statistik t

Uji statistic t digunakan untuk mengetahui seberapa besar pengaruh variabel-variabel X (independen) pada variabel Y (dependen). Untuk melihat hipotesis dalam penelitian yang diteliti diterima atau di tolak peneliti harus melakukan perbandingan antara besaran nilai t tabel dengan t hitung. Apabila peneliti memperoleh nilai probabilitas  $< \alpha$  maka dapat disimpulkan bahwa variabel bebas (X) mempengaruhi variabel terikat (Y). Sebaliknya, apabila nilai probabilitas  $> \alpha$  maka bahwa variabel bebas (X) tidak mempengaruhi variabel terikat (Y).

Besaran signifikansi  $\alpha = 5\%$  atau dengan derajat keyakinan 95% dengan ketentuan apabila jika nilai probability t-statistik  $< 5\%$  maka H0 ditolak dan H1 diterima, lalu apabila nilai probability t-statistik  $> 5\%$  maka H0 diterima dan H1 ditolak.

### b. Uji Statistik F

Dalam melihat sejauh apa variabel bebas (X) mampu menjelaskan variabel terikat (Y) dalam sebuah penelitian, maka haruslah menggunakan uji f. Uji statistik f dapat disebut juga dengan uji kelayakan model regresi yang digunakan untuk menjabarkan pengaruh variabel bebas terhadap variabel terikat. Hipotesis uji f menunjukkan  $H_0$  ditolak dan  $H_1$  diterima apabila  $f_{hitung} > f_{tabel}$ , artinya terdapat pengaruh variabel bebas dan variabel terikat. Begitupun sebaliknya, apabila  $f_{hitung} < f_{tabel}$   $H_0$  diterima  $H_1$  di tolak, hal tersebut menunjukan bahwa tidak terdapat pengaruh antara variabel bebas dengan variabel terikat.

Besaran signifikansi  $\alpha = 5\%$  atau dengan derajat keyakinan 95% dengan ketentuan apabila jika nilai probability t-statistik  $< 5\%$  maka  $H_0$  ditolak dan  $H_1$  diterima, lalu apabila nilai probability t-statistik  $> 5\%$  maka  $H_0$  diterima dan  $H_1$  ditolak.

### c. Koefisien Determinasi ( $R^2$ )

Koefisien determinasi digunakan untuk melihat seberapa besar model yang digunakan dapat menjelaskan variabel terikat (Y) yang di teliti. Ketika nilai  $R^2$  mendekati angka satu maka variabel bebas (X) mampu memberikan hampir seluruh variasi variabel maka model yang digunakan sudahlah tepat. Namun apabila nilai  $R^2$  mendekati angka nol maka akses informasi dalam menjelaskan variabel terikat (Y) sangat terbatas.