

## **BAB III**

### **METODE PENELITIAN**

#### **A. Obyek Penelitian**

Dalam penelitian ini penulis melakukan penelitian yang dilakukan pada negara Indonesia. Penelitian mulai dilakukan pada bulan oktober 2018 dengan pengambilan data yang dilakukan di Provinsi Daerah Istimewa Yogyakarta. Pada penelitian ini peneliti menggunakan variabel terikat (Y) berupa data pertumbuhan ekonomi pada setiap provinsi yang ada di Indonesia pada periode tahun 2012-2016. Peneliti juga menggunakan variabel bebas (X) berupa data penanaman modal dalam negeri, jumlah tenaga kerja, angka melek huruf dan pengeluaran pemerintah pada setiap provinsi di Indonesia selama periode tahun 2012-2016.

#### **B. Jenis dan Sumber Data**

Data yang digunakan dalam penyusunan penelitian ini menggunakan data panel hasil dari analisa data sekunder yang diperoleh melalui sumber-sumber yang sistematis berupa data runtun waktu *time series* dari tahun 2012-2016. Data sekunder pada penelitian ini digunakan karena penelitian ini bersifat makro dan data sekunder lebih mudah diperoleh pada instansi-instansi terkait. Data sekunder yang digunakan dalam penelitian ini diperoleh dari Badan Pusat Statistik (BPS) Provinsi Daerah Istimewa Yogyakarta.

### **C. Teknik Pengumpulan Data**

Pengumpulan data dalam suatu penelitian untuk membantu penulis mendapatkan bahan-bahan yang akurat, realistis serta akurat. Metode yang digunakan penulis dalam mengumpulkan data ialah metode studi pustaka yang diperoleh pada instansi-instansi terkait, dan referensi berupa jurnal-jurnal ekonomi. Data yang digunakan dalam penelitian ini adalah data runtut waktu yang telah di catat dan observasi sepanjang waktu oleh instansi terkait dengan jenis data yang digunakan berupa data sekunder.

### **D. Variabel dan Definisi Operasional**

#### **1. Variabel Penelitian**

Variabel dalam penelitian ini adalah penanaman modal dalam negeri, jumlah tenaga kerja, angka melek huruf dan pengeluaran pemerintah sebagai variabel bebas (independen) serta pertumbuhan ekonomi sebagai variabel terikat (dependen).

#### **2. Definisi Operasional**

Definisi operasional menunjukkan definisi setiap variabel penelitian serta satuan alat ukur yang digunakan dalam variabel penelitian. Dalam penelitian ini definisi operasional masing-masing variabel yang digunakan adalah sebagai berikut:

##### **a) Variabel dependen (terikat)**

Variabel dependen atau terikat merupakan variabel yang dapat dipengaruhi oleh adanya variabel bebas. Dalam penelitian

ini, variabel dependen yang digunakan adalah pertumbuhan ekonomi pada setiap provinsi di Indonesia. Pengambilan data berupa Produk Domestik Regional Bruto (PDRB) pada setiap provinsi di Indonesia atas dasar harga konstan dengan memakai satuan miliar rupiah selama periode tahun 2012-2016.

**b) Variabel Independen (bebas)**

Variabel independen atau bebas merupakan variabel yang dapat mempengaruhi variabel dependen. Variabel bebas yang digunakan pada penelitian ini yaitu penanaman modal dalam negeri, jumlah tenaga kerja, dan angka melek huruf dengan definisi sebagai berikut:

**1) Variabel penanaman modal dalam negeri**

Pada variabel penanaman modal dalam negeri, data yang digunakan dalam penelitian ini diambil dari data investasi dalam negeri pada setiap provinsi di Indonesia dengan memakai satuan jutaUS\$ selama periode tahun2012-2016.

**2) Variabel jumlah tenaga kerja**

Variabel jumlah tenaga kerja yang digunakan dalam penelitian ini diambil dari data penduduk usia 15 tahun ke atas yang memasuki usia kerja pada setiap provinsi di Indonesia dengan satuan juta selama periode tahun 2012-2016.

### 3) Variabel angka melek huruf

Variabel angka melek huruf pada penelitian ini diambil dari data angka melek huruf penduduk yang berusia 15 tahun ke atas pada setiap provinsi di Indonesia dengan memakai satuan persen selama periode tahun 2012-2016.

### 4) Variabel pengeluaran pemerintah

Variabel pengeluaran pemerintah pada penelitian ini diambil dari data belanja negara atau *expenditure* yang dikeluarkan oleh pemerintah dalam negeri pada setiap provinsi di Indonesia dengan memakai satuan juta rupiah selama periode 2012-2016.

## E. Metode Analisis Data

Pada penelitian ini metode analisis yang digunakan yaitu menggunakan metode dengan analisis kuantitatif. Analisis kuantitatif yang digunakan pada penelitian ini yaitu metode analisis regresi data panel. Data panel adalah gabungan antara data runtut waktu (*time series*) dengan data silang (*cross section*).

Penggunaan data panel dalam observasi memiliki beberapa keuntungan yang diperoleh. Keuntungan penggunaan data panel menurut Widarjono (2009) yaitu, pertama data panel merupakan gabungan antara dua data yaitu data *time series* dengan *cross section* sehingga mampu

menyediakan data yang lebih banyak sehingga akan lebih menghasilkan *degree of freedom* yang lebih besar. Kedua, menggabungkan informasi dari data *time series* dan *cross section* dapat mengatasi masalah yang timbul ketika ada masalah penghilangan variabel (*omitted-variabel*).

Sedangkan menurut Wibisono (2005) regresi data panel memiliki keunggulan antara lain, pertama data panel mampu memperhitungkan heterogenitas individu secara eksplisit dengan mengizinkan variabel spesifik individu. Kedua, mampu mengontrol heterogenitas sehingga data panel dapat membangun model yang lebih kompleks. Ketiga, data panel yang memuat *cross section* yang berulang-ulang (*time series*) sehingga cocok untuk *study of dynamic adjustment*. Keempat, banyaknya jumlah observasi memiliki implikasi pada data yang lebih informatif, variatif dan kolinearitas antara data semakin berkurang dan derajat kebebasan (*degree of freedom*) lebih tinggi sehingga hasil lebih efisien. Kelima, data panel dapat digunakan untuk mempelajari model-model perilaku yang kompleks. Keunggulan yang keenam, data panel dapat meminimalkan bias yang mungkin terjadi karena agregasi individu.

Analisis model regresi dari penelitian ini sebagai berikut:

$$Y = \alpha + b_1x_{1it} + b_2x_{2it} + b_3x_{3it} + b_4x_{4it} + e$$

Dimana:

Y = variabel dependen

$\alpha$  = konstanta

$x_2$  = variabel independen 1

$x_3$  = variabel independen 2

$x_4$  = variabel independen 3

$x_5$  = variabel independen 4

$e$  = waktu term

$t$  = waktu / time series

$i$  = provinsi

Dalam penelitian ini peneliti memodifikasi model dengan menyesuaikan ketersediaan data di Daerah Istimewa Yogyakarta. Sehingga diperoleh persamaan yang digunakan dalam penelitian ini sebagai berikut:

$$PDRB = \alpha + PMDN_{it} + TK_{it} + AMH_{it} + PP_{it} + e$$

Dimana:

PDRB : Jumlah pendapatan ekonomi di setiap provinsi di Indonesia pada tahun 2012-2016.

PMDN: Penanaman modal dalam negeri di setiap provinsi di Indonesia pada tahun 2012-2016.

TK : Jumlah penduduk yang berusia 15 tahun ke atas yang bekerja di setiap provinsi di Indonesia pada tahun 2012-2016.

AMH : Jumlah penduduk angka melek huruf yang berusia 15 tahun ke atas di setiap provinsi di Indonesia pada tahun 2012-2016.

PP : Pengeluaran pemerintah di setiap provinsi di Indonesia pada tahun 2012-2016.

Dalam metode estimasi model regresi dengan menggunakan data panel dapat dilakukan dengan tiga pendekatan sebagai berikut:

### **1. *Common Effect Model***

*Common effect model* merupakan bentuk pendekatan model data panel yang paling sederhana karena hanya mengkombinasikan antara data time series dengan *cross section* tanpa memperhatikan dimensi waktu maupun individu/wilayah, sehingga dapat diasumsikan bahwa perilaku setiap individu sama dalam berbagai kurun waktu. Metode ini bisa menggunakan *Ordinary Least Square* (OLS) atau teknik kuadrat terkecil dalam mengestimasi data panel.

### **2. *Fixed Effect Model***

Pada *Fixed Effect Model* mengasmsikan bahwa perbedaan individu dapat diakomodasi dari perbedaan intersepnya. Dalam mengestimasi data panel model *Fixed Effect* menggunakan teknik variabel dummy untuk menangkap perbedaan intersep antar individu/wilayah, namun terdapat kesamaan pada slop antar wilayah. Model estimasi ini sering juga disebut sebagai *Least Squares Dummy Variable* (LSDV).

### **3. *Random Effect Model***

*Random Effect Model* akan mengestimasi data panel dimana variabel gangguan mungkin saling berhubungan antar

wilayah. Pada model ini perbedaan intersep diakomodasi oleh eror term dari masing-masing wilayah. *Model Random Effect* memiliki keuntungan yakni menghilangkan heteroskedastisitas. Model ini juga disebut sebagai *Error Component Model* (ECM) atau teknik *Generalized Least Square* (GLS).

Dalam memilih model yang tepat digunakan dalam mengelola data panel, terdapat beberapa pengujian yang dapat dilakukan sebagai berikut:

### **1. Uji Chow**

Chow test yaitu pengujian untuk memnentukan model *Fixed Effect* atau *Random Effect* yang paling tepat untuk digunakan dalam mengestimasi data panel.

### **2. Uji Hausman**

Hausman test yaitu pengujian statistik untuk memilih model yang paling tepat digunakan antara model *Fixed effect* atau *Random Effect*.

### **3. Uji Lagrange Multiplier**

Pengujian ini digunakan untuk mengetahui apakah model *Random Effect* lebih baik daripada



metode *Common Effect* (OLS) yang paling tepat untuk mengestimasi data panel.

## **F. Uji Kualitas Data**

### **1. Uji Multikolinearitas**

Menurut Basuki (2017) bahwa salah satu asumsi regresi linier klasik adalah tidak adanya multikolinearitas sempurna (*no perfect multicollinearity*) yaitu tidak adanya hubungan linier antara variabel bebas atau variabel penjelas dalam suatu model regresi. Menurut Frisch dalam Basuki dan Yuliadi (2015) suatu model regresi dikatakan terkena multikolinearitas apabila terjadi hubungan linier antara variabel bebas dengan variabel terikat.

Multikolinearitas memiliki konsekuensi yaitu invalidnya signifikansi variabel maupun besaran koefisien variabel dan konstanta. Terjadinya multikolinearitas diduga apa bila menghasilkan nilai R kuadrat yang tinggi atau lebih dari 0,8, nilai F tinggi, sedangkan nilai t-statistik semua atau hampir semua variabel penjelas tidak signifikan.

### **2. Uji Heteroskedastisitas**

Menurut Basuki dan Yuliadi (2015) Homoskedastisitas akan terjadi apabila distribusi probabilitas tetap sama dalam sebuah observasi  $x$ , dan varians setiap residual adalah sama untuk semua nilai sama untuk setiap variabel penjelas atau bebas. Sebaliknya, apabila terjadi heterokedastisitas maka nilai variansnya akan berbeda-beda.

Kesimpulan dari heteroskedastisitas ialah situasi dimana varians bernilai tidak konstan. Heteroskedastisitas memiliki konsekuensi yaitu biasanya varians konstan sehingga uji signifikansi menjadi invalid. Heteroskedastisitas dapat dideteksi dengan salah satu cara yaitu dengan melakukan uji Glesjer (Basuki dan Yuliadi, 2015).

### 3. Uji Autokorelasi

Autokorelasi dalam Basuki dan Yuliadi (2015) disebabkan karena adanya kelembaban (*inertia*) yaitu diman data runtun waktu ekonomi menunjukkan adanya pola konjungtur, hal tersebut biasanya terjadi pada data observasi pada periode sebelumnya dengan periode sekarang sehingga kemungkinan besar akan menyebabkan saling ketergantungan (*interdependence*). Selanjutnya, yaitu adanya *bias specification* atau disebut dengan kasus variabel yang tidak dimasukkan. Penyebab hal ini adalah tidak masuknya variabel yang menurut teori ekonomi, karena dalam memasukkan estimasi diperlukan variabel yang berdasar pada teori. Terakhir yaitu adanya fenomena sarang laba-laba (*cobweb phenomenon*). Fenomena ini rata-rata terjadi pada penawaran komoditi sektor pertanian, karena reaksi penawaran terhadap perubahan harga terjadi setelah melalui suatu tenggang waktu.

Penggunaan ekonometrika mungkin akan merubah bentuk fungsi persamaan regresinya, misalnya dalam bentuk log atau *first*

*difference* untuk mengetahui ada-tidaknya autokorelasi (Basuku dan Yuliadi 2015). Pendeteksian autokorelasi dilakukan dengan membandingkan antara nilai statistik Durbin Watson hitung dengan Durbin Watson tabel, mekanisme dari uji Durbin Watson adalah sebagai berikut:

- a. Mendapatkan nilai residual dengan melakukan regresi OLS.
- b. Menghitung nilai  $d$  (yang di maksud yaitu Durbin Watson).
- c. Dapatkan nilai kritis  $d_L$  dan  $d_u$ .
- d. Apabila diasumsikan yaitu dengan hipotesis sebesar 0 (nol) adalah bahwa tidak ada serial korelasi positif, maka jika
  - $d < d_L$ , maka  $H_0$  ditolak atau tidak diterima
  - $d < d_u$ , maka  $H_0$  diterima
  - $d = d_L = d_u$ , maka pengujian ini meragukan atau tidak meyakinkan.
- e. Apabila hipotesis nol adalah tidak ada serial korelasi negatif, maka jika
  - $d > 4-d_L$ , maka  $H_0$  di tolak
  - $d < 4-d_u$ , maka  $H_0$  diterima
  - $d = 4-d_u = 4-d_L$ , maka pengujian tidak meyakinkan
- f. Apabila  $H_0$  merupakan dua ujung, berarti diketahui bahwa tidak ada serial korelasi baik positif maupun negatif, maka jika
  - $d < d_L$ , maka  $H_0$  ditolak
  - $d > 4-d_L$ , maka  $H_0$  ditolak

$d_u < d < 4-d_u$ , maka  $H_0$  diterima

$d_L = d = d_u$ , pengujian tidak meyakinkan

$4-d_u = d = 4-d_L$ , pengujian juga tidak meyakinkan

#### 4. Uji Normalitas

Fungsi uji normalitas menurut Basuki dan Yuliadi (2015) adalah untuk mengetahui apakah residualnya berdistribusi normal atau tidak. Hal tersebut dilakukan dengan membandingkan antara *Jarque Bera* (JB) dengan  $X^2$  tabel, maka jika

- a. Nilai probabilitas dari *Jarque Bera* (JB)  $> 0,05$ , maka residualnya berdistribusi secara normal.
- b. Nilai probabilitas dari *Jarque Bera* (JB)  $< 0,05$ , maka residualnya berdistribusi secara tidak normal.

Uji asumsi klasik yang digunakan dalam regresi linier dengan pendekatan *Ordinary Least Square* (OLS) meliputi uji multikolinearitas, heteroskedastisitas, autokorelasi dan normalitas. Walaupun demikian, tidak semua uji asumsi klasik harus dilakukan pada setiap model regresi linier dengan pendekatan OLS (Basuki dan Yuliadi, 2015).

- a. Uji multikolinearitas perlu dilakukan pada regresi linier yang memiliki lebih dari satu variabel bebas. Jika hanya memiliki satu variabel bebas maka dapat disebut tidak mungkin terjadi multikolinearitas.

- b. Uji heteroskedastisitas biasanya terjadi pada data *cross section*, yaitu dimana data panel lebih dekat ke ciri data *cross section* dibandingkan dengan data *time series*.
- c. Uji autokorelasi tidak berlaku dan akan sia-sia pada data yang bersifat selain *time series* (*cross section* atau data panel). Pengujian ini hanya berlaku pada data *time series* saja.
- d. Uji normalitas pada dasarnya bukan merupakan syarat *Best Linier Unbias Estimator* (BLUE) dan banyak pendapat yang tidak mengharuskan bahwa uji normalitas sebagai sesuatu yang wajib dipenuhi.

Dari penjelasan di atas maka dapat ditarik kesimpulan bahwa tidak semua uji asumsi klasik yang digunakan pada metode *Ordinary Least Square* (OLS) dapat dipakai pada data panel, maka dari itu dalam penelitian ini peneliti hanya akan melakukan pengujian dengan uji multikolinearitas dan uji heteroskedastisitas.

## **G. Uji Hipotesis**

### **1. Uji Kecocokan Model (*Goodnes of Fit Test*)**

Nilai koefisien determinasi ( $R^2$ ) dalam uji kecocokan model digunakan untuk menjelaskan seberapa besar variasi perubahan variabel endogen (Basuki dan Yuliadi, 2015). Saat nilai  $R^2$  sedang tinggi maka model yang dirumuskan dapat menerangkan seberapa

besar variasi perubahan dari variabel endogen. Sebaliknya jika nilai  $R^2$  sedang rendah berarti model yang dirumuskan lemah dalam menjelaskan seberapa besar variasi perubahan variabel endogen.  $R^2$  memiliki batasan yaitu antara nol dengan satu ( $0 \leq R^2 \leq 1$ ) sehingga nilai  $R^2$  tidak pernah negatif. Perumusan nilai  $R^2$  adalah sebagai berikut:

$$R^2 = SSR/TSS = 1 - \frac{\sum e_i^2}{\sum y_i^2} \dots\dots\dots$$

Dimana:

$R^2$  = Koefisien determinasi

RSS = *Residual sum of squares*

TSS = *Total sum of squares*

N = Jumlah observasi

K = Jumlah parameter (termasuk intersept)

## 2. Uji Signifikan Parameter (*Partial test* atau Uji t)

Uji t digunakan untuk mengetahui seberapa pengaruh secara parsial dari masing-masing variabel independen terhadap variabel endogen (Basuki dan Yuliadi, 2015). Sebagaimana prosedur uji t dapat dilihat sebagai berikut:

$$H_0 : \beta_1 = 0$$

$$H_a : \beta_1 \text{ tidak sama dengan } 0 \text{ (} i = 1, 2, 3, \dots, n \text{)}$$

$$\text{Dimana } t \text{ hitung} = \beta_1 / \text{Se}(\beta_1) \text{ dan } \text{Se}(\beta_1) = \sqrt{\text{var}(\beta_1)}$$

Jika nilai  $t$  hitung  $>$  nilai  $t$  kritis, maka  $H_0$  ditolak dan berarti variabel independen berpengaruh secara signifikan terhadap variabel endogen. Sebaliknya jika nilai  $t$  hitung  $<$  nilai  $t$  kritis, maka  $H_0$  diterima dan berarti variabel independen tidak berpengaruh secara signifikan terhadap variabel endogen. Dalam satu penelitian terdapat toleransi kesalahan ( $\alpha$ ) maksimum yaitu sebesar 10 persen atau disebut dengan taraf kepercayaan minimum 90 persen dan jika nilainya dibawah ini berarti tidak signifikan.

### 3. Uji Signifikasi Garis Regresi Secara Keseluruhan (*Overall Test*)

Uji F dilakukan untuk mengetahui apakah variabel independen secara bersama-sama berpengaruh signifikan terhadap variabel dependen (Basuki dan Yuliadi, 2015). Pengujian F dapat dilakukan dengan kaidah sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_n$$

Ha : Paling sedikit terdapat satu  $\beta_1$  yang tidak sama dengan nol

$$F = \frac{R^2/(K - 1)}{(1 - R^2)/(n - k)} \dots \dots \dots$$

Dimana:

$k$  = Jumlah variabel dalam model regresi

$n$  = Jumlah observasi

Jika nilai  $F$  hitung  $> F$  tabel pada derajat kesalahan ( $\alpha$ ) dengan derajat kebebasan  $k-1$  dan  $N-1$ , maka  $H_0$  ditolak dan berarti variabel independen secara bersama-sama berpengaruh secara signifikan terhadap variabel endogen. Sebaliknya jika  $F$  hitung  $< F$  tabel, maka  $H_0$  diterima dan berarti variabel independen secara simultan tidak berpengaruh secara signifikan terhadap variabel endogen.