

## **BAB III**

### **METODE PENELITIAN**

#### **A. Objek Penelitian**

Obyek penelitian yang digunakan oleh peneliti dalam penelitian ini adalah tingkat pendidikan, pertumbuhan ekonomi dan rasio gini di lima kabupaten/kota di Daerah Istimewa Yogyakarta periode tahun 2010-2015 sebagai variabel bebas (X). Selain itu peneliti juga menggunakan data tingkat pengangguran terbuka di Daerah Istimewa Yogyakarta periode 2010-2015 sebagai variabel terikat (Y).

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

Peneliti dalam penelitian ini menggunakan data sekunder yaitu berupa *time series* dan *cross section data* atau data panel. Data sekunder merupakan data yang disajikan dan atau dipakai oleh lembaga atau badan yang bukan pengolahnya. Sumber data sekunder dalam penelitian ini diperoleh dari Badan Pusat Statistik (BPS) Provinsi D.I. Yogyakarta.

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

Data yang digunakan pada penelitian ini dikumpulkan oleh penulis dengan menggunakan metode *library research* atau kepustakaan yaitu penelitian yang menggunakan bahan-bahan kepustakaan berupa tulisan ilmiah, artikel, jurnal, majalah, laporan-laporan penelitian ilmiah yang berhubungan dengan topic penelitian. Teknik pengumpulan data pada penelitian ini dengan melakukan pencatatan secara langsung berupa data *time series* dan *cross*

*section* dari tahun 2010 sampai dengan 2015 yang diperoleh Badan Pusat Statistik (BPS).

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

### **1. Variabel Penelitian**

Dalam sebuah penelitian menggunakan metode kuantitatif, variabel merupakan dasar dari pembahasan. Menurut Martono (2011) variabel merupakan unsur yang memiliki lebih dari satu nilai atau merupakan suatu unsur yang bervariasi.

### **2. Definisi Operasional**

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

#### **a. Variabel Dependen**

Variabel dependen adalah variabel yang dipengaruhi oleh variabel independen (Martono, 2011). Variabel dependen yang digunakan dalam penelitian ini adalah tingkat pengangguran terbuka yaitu jumlah tingkat penduduk yang sedang mencari pekerjaan, atau mereka yang mempersiapkan usaha, atau mereka yang tidak mereka yang tidak mencari pekerjaan karena tidak mungkin mendapat pekerjaan dan mereka yang sudah memiliki pekerjaan namun belum mulai bekerja dan dalam waktu yang bersamaan mereka tidak bekerja. TPT

memberikan indikasi tentang penduduk usia kerja yang termasuk kedalam kelompok pengangguran. Variabel ini menggunakan satuan persen.

b. Variabel Independen

Menurut Martono dalam Arsono (2014) Variabel independen merupakan variabel yang mempengaruhi dan menghasilkan akibat pada variabel lain. Variabel independen yang digunakan dalam penelitian ini adalah sebagai berikut :

1) Variabel Pendidikan (PND)

Variabel yang digunakan dalam penelitian ini diperoleh dari data angka melek huruf (AMH) pada penduduk yang berusia 15 tahun ke atas pada 5 kabupaten atau kota di Provinsi DIY dengan menggunakan satuan persen dalam periode 2010-2015.

2) Variabel Rasio Gini (GINI)

Variabel gini merupakan indeks ketimpangan pendapatan antar penduduk di lima kabupaten atau kota di Provinsi DIY. Suatu indeks pada umumnya tidak memiliki satuan ukuran.

3) Variabel Laju Pertumbuhan Ekonomi

Variabel dalam penelitian ini adalah perubahan berupa peningkatan ataupun penurunan aktivitas perekonomian domestik yang ditunjukkan dengan laju pertumbuhan Produk Domestik Bruto (PDRB). Variabel ini menggunakan satuan persen.

## E. Metode Analisis

Metode analisis yang digunakan di dalam penelitian ini adalah metode dengan analisis kuantitatif. Analisis kuantitatif yang digunakan menggunakan metode analisis regresi data panel. Menurut Basuki dan Yuliadi (2015) data panel adalah gabungan antara data runtut waktu (*time series*) dan data silang (*cross section*). Menurut Widarjono dalam Basuki dan Yuliadi (2015) menggunakan data panel dalam penelitian mempunyai beberapa keuntungan yaitu. Pertama, data panel yang merupakan gabungan antara *time series* dan *cross section* mempunyai *degree of freedom* yang lebih besar karena data yang tersedia lebih banyak. Kedua, menggunakan data panel dapat mengatasi masalah yang muncul karena penghilangan variabel (*omitted-variable*).

Sedangkan menurut Wibisono dalam Basuki dan Yuliadi (2015) keunggulan menggunakan data panel yaitu pertama, 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 kolinieritas (multikolinieritas) antara data semakin berkurang dan derajat kebebasan (*degree of freedom/df*) lebih tinggi sehingga hasil lebih efisien. Kelima, data panel dapat digunakan untuk mempelajari model-model perilaku yang kompleks. Dan keenam, data panel dapat meminimalkan bias yang mungkin terjadi karena agregasi individu.

Analisis regresi dalam penelitian ini diolah menggunakan program *Eviews 7.0* dengan bentuk persamaan sebagai berikut :

$$Y_i = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{4it} + \varepsilon_{it} \dots \dots \dots (3)$$

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

Dimana:

Y = variabel dependen

$\beta$  = koefisien regresi

i = *cross-section*

t = waktu / *time series*

t = Data Panel

$\varepsilon$  = *error term*

Model dalam penelitian ini penulis modifikasi disesuaikan dengan ketersediaan data di Daerah Istimewa Yogyakarta. Sehingga diperoleh persamaan yang digunakan dalam penelitian ini yaitu :

$$TPT_{it} = \beta_1 + \beta_2 PND_{it} + \beta_3 PE_{it} + \beta_4 GINI_{it} + \varepsilon_{it} \dots \dots \dots (4)$$

Dimana:

TPT : tingkat pengangguran terbuka di Provinsi D.I. Yogyakarta tahun 2010-2015

PND : angka melek huruf penduduk usia 15 tahun ke atas di Provinsi D.I. Yogyakarta tahun 2010-2015

PE : laju pertumbuhan ekonomi di Provinsi D.I. Yogyakarta tahun 2010-2015

GINI : ketimpangan distribusi pendapatan di Provinsi D.I. Yogyakarta tahun 2010-2015

Dalam metode estimasi model regresi dengan menggunakan data panel dapat dilakukan dengan tiga pendekatan (Basuki dan Yuliadi, 2015).

### 1. *Common Effect Model*

Merupakan bentuk estimasi paling sederhana karena hanya menggunakan kombinasi data *time series* dan data *cross section* tanpa memperhatikan dimensi waktu maupun individu./wilayah Sehingga mengasumsikan perilaku setiap individu sama dalam berbagai kurun waktu. Metode estimasi ini bisa menggunakan pendekatan *Ordinary Least Square* (OLS) atau teknik kuadrat terkecil dalam mengestimasi data panel.

### 2. *Fixed Effect Model*

Model ini mengasumsikan bahwa perbedaan individu dapat diakomodasi dari perbedaan intersepnya. Estimasi *Fixed Effect Model* (FEM) menggunakan teknik variabel *dummy* untuk melihat perbedaan intersep antar individu/wilayah, namun terdapat kesamaan slop antar wilayah. Teknik ini juga sering disebut sebagai *Least Square Dummy Variabel* (LSDV).

### 3. *Random Effect Model*

Model ini mengestimasi data panel dimana variabel gangguan mungkin saling berhubungan antar wilayah. Pada model ini perbedaan intersep diakomodasi oleh *error term* masing-masing wilayah. Keuntungan menggunakan model ini adalah menghilangkan heteroskedastisitas. Model ini juga biasa disebut sebagai *Error Component Model* atau teknik *Generalized Least Square (GLS)*.

Untuk menentukan model yang tepat dalam estimasi data panel perlu dilakukan pengujian terlebih dahulu (Basuki dan Yuliadi, 2015).

a. Uji Chow

Chow test yakni pengujian untuk menentukan model *Fixed Effect* atau *Random Effect* yang paling tepat dalam mengestimasi data panel.

b. Uji Hausman

Uji Hausman digunakan untuk memilih apakah model *Fixed Effect* atau *Random Effect* yang paling tepat untuk mengestimasi data panel.

c. Uji Lagrange *Multiplier* (LM)

Pengujian ini digunakan untuk mengetahui apakah model *Random Effect* atau *Common Effect (OLS)* yang paling tepat untuk mengestimasi data panel.

## F. Uji Kualitas Data

### 1. Desteksi Multikolinearitas

Multikolinearitas dapat diartikan sebagai suatu keadaan dimana satu atau lebih variabel bebas dapat dinyatakan sebagai kombinasi kolinier dari variabel yang lainnya. Uji ini bertujuan untuk mengetahui apakah dalam regresi ini ditemukan adanya korelasi antar variabel independen. Jika terjadi korelasi maka terdapat problem multikolinearitas.

Menurut Gujarati dalam Arsono (2014) penyakit multikolinieritas dalam sebuah regresi dapat dilihat dari gejala sebagai berikut :

- a. Estimasi menghasilkan nilai R kuadrat yang lebih tinggi (lebih dari 0,8) nilai F, sedangkan nilai t-statistik semua atau hampir semua variabel tidak signifikan.
- b. Melakukan regresi parsial, yaitu :
  - 1) Lakukan regresi variabel dalam level atau regresi awal sehingga didapat nilai R kuadrat
  - 2) Lakukan *auxiliary regression* pada setiap variabel bebas



- 3) Bandingkan nilai R kuadrat pada regresi awal dengan regresi parsial, jika nilai R kuadrat regresi parsial lebih tinggi maka terdapat multikolinearitas
- 4) Melakukan korelasi variabel bebas, jika nilainya lebih dari 0,8 maka terjadi multikolinearitas

## 2. Deteksi Autokorelasi

Penyebab terjadinya autokorelasi menurut Basuki dan Yuliadi (2015) yaitu adanya Kelembaman (*inertia*) yaitu adanya pola konjungtur, hal tersebut terjadi karena pada data observasi periode sebelumnya dengan periode sekarang kemungkinan besar saling ketergantungan. Kedua, yaitu adanya bias atau dengan kata lain variabel yang tidak dimasukkan. Hal itu terjadi karena variabel yang berdasar teori sangat penting perannya terhadap variabel terikat tidak dimasukkan dalam estimasi. Ketiga, adanya fenomena sarang laba-laba (*cobweb phenomenon*).

Cara mendeteksi ada atau tidaknya suatu autokorelasi dalam sebuah regresi dapat dilakukan dengan cara uji *d* Durbin Watson (*Durbin Watson d Test*). Menurut Gujarati dalam Astuti (2014) cara untuk mendeteksi adanya autokorelasi yaitu :

- a. Lakukan regresi *OLS* dan dapatkan residualnya
- b. Hitung nilai *d* (*Durbin Watson*)
- c. Dapatkan nilai  $d_L$  dan  $d_U$

d. Apabila hipotesis nol adalah tidak ada serial korelasi positif maka, jika

$d < d_L$ , tolak  $H_0$

$d < d_U$ , terima  $H_0$

$d_L = d = d_U$ , pengujian tidak meyakinkan

Apabila hipotesis nol adalah tidak serial korelasi negative, maka jika

$d > 4-d_L$ , tolak  $H_0$

$d < 4-d_U$ , terima  $H_0$

$4-d_L = d = 4-d_U$ , pengujian tidak meyakinkan

e. Apabila  $H_0$  adalah dua ujung, yaitu bahwa tidak ada serial korelasi positif manapun atau negatif maka jika

$d < d_L$ , tolak  $H_0$

$d > 4-d_L$ , tolak  $H_0$

$d_U < d < 4-d_U$ , terima  $H_0$

$d_U = d = d_L$ , pengujian tidak meyakinkan

$4-d_U = d = 4-d_L$ , pengujian tidak meyakinkan

### 3. Heteroskedastisitas

Homoskedastisitas terjadi apabila nilai probabilitas tetap sama dalam sebuah observasi  $x$ , dan varian setiap residual sama untuk setiap variabel bebas, sebaliknya apabila terjadi heteroskedastisitas maka nilai variansnya berbeda (Basuki dan Yuliadi, 2015).

Menurut Gujarati dalam Astuti (2014) untuk mendeteksi penyakit heteroskedastisitas salah satunya yaitu dilakukan dengan Uji White. Jika

nilai *chi-square* ( $X^2$ ) yang didapat lebih besar dari *chi-square* kritis maka terdapat heteroskedastisitas dalam model tersebut.

#### 4. Deteksi Normalitas

Deteksi normalitas digunakan untuk mengetahui apakah data residualnya terdistribusi secara normal atau tidak. Menurut Gujarati dalam Astuti (2014) normalitas suatu data dapat diuji dengan melihat *Normal Probability Plot (NPP)*, jika data terdistribusi normal maka NPP terlihat seperti garis lurus. Selain menggunakan NPP, kita juga dapat mendekteksi normalitas suatu data dengan membandingkan nilai *Jarque Bera (JB)* dengan  $X^2$  tabel, jika nilai J-B hitung  $> 0,05$  maka terdistribusi normal, namun jika J-B hitung  $< 0,05$  maka data terdistribusi tidak normal.

Menurut Basuki dan Yuliadi (2015) tidak semua uji asumsi klasik harus digunakan pada setiap regresi.

- a. Uji linearitas hampir tidak digunakan dalam setiap regresi karena sudah diasumsikan bahwa model bersifat linear. Walaupun harus dilakukan uji tersebut maka hanya untuk melihat sejauh mana tingkat linearitasnya.
- b. Uji normalitas bukan merupakan syarat BLUE (*Best Linier Unbias Estimator*). Beberapa pendapat menyebutkan bahwa tidak mengharuskan uji ini sebagai yang wajib diuji.
- c. Autokorelasi hanya terjadi pada data *time series*. Pengujian pada data panel akan sia-sia.

- d. Multikolinearitas perlu dilakukan pada regresi linier apabila menggunakan variabel bebas lebih dari satu. Apabila hanya terdapat satu variabel bebas maka pastilah tidak terjadi multikolinearitas.
- e. Heteroskedastisitas biasanya terjadi pada data *cross section*, dimana data panel lebih mendekati ciri-ciri data *cross section* dibandingkan *time series*.

Dari penjelasan di atas maka dapat ditarik kesimpulan bahwa pada regresi dengan menggunakan data panel tidak semua uji asumsi klasik digunakan pada metode OLS, maka dari itu peneliti hanya akan melakukan pengujian dengan uji multikolinieritas dan uji heteroskedastisitas saja.

## **G. Uji Hipotesis**

### **1. Uji Koefisien Determinasi (R-square)**

Koefisien determinasi  $R^2$  pada intinya mengukur seberapa jauh kemampuan model dalam menerangkan variasi variabel independen untuk mengukur kebaikan suatu model (*Goodness of Fit*). Nilai koefisien determinasi diantara 0 dan 1 ( $0 < R^2 < 1$ ), nilai  $R^2$  yang kecil berarti kemampuan variabel-variabel independen dalam menjelaskan variasi variabel independen sangat terbatas. Nilai yang mendekati 1 berarti variabel independen memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variasi variabel dependen (Gujarati, 2003).

Kelemahan mendasar penggunaan koefisien determinasi adalah bias terhadap jumlah variabel dependen,  $R^2$  pasti meningkat, tidak peduli apakah

variabel tersebut berpengaruh secara signifikan terhadap variabel dependen atau tidak. Oleh karena itu, banyak peneliti menganjurkan untuk menggunakan nilai *adjusted R<sup>2</sup>* pada saat mengevaluasi model regresi terbaik. Tidak seperti nilai  $R^2$ , nilai *adjusted R<sup>2</sup>* dapat naik dapat turun apabila satu variabel independen ditambahkan dalam model. Pengujian ini pada intinya adalah mengukur seberapa jauh kemampuan model dalam menerangkan variasi variabel independen.

## 2. Uji F-Statistik

Uji F-statistik ini dilakukan untuk melihat seberapa besar pengaruh variabel independen secara keseluruhan untuk bersama-sama terhadap variabel dependen. Adapun langkah-langkah yang dilakukan dalam uji ini sebagai berikut:

### a. Merumuskan Hipotesis

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ , artinya secara bersama-sama tidak ada pengaruh variabel independen terhadap variabel dependen.

$H_a: \beta_1: \beta_2: \beta_3: \beta_4 \neq 0$ , artinya secara bersama-sama ada pengaruh variabel independen terhadap variabel dependen

### b. Pengambilan Keputusan

Pengambilan keputusan dalam uji F dilakukan dengan membandingkan probabilitas pengaruh variabel independen secara simultan antara variabel dependen dengan nilai alpha yang digunakan, dalam penelitian ini penulis menggunakan alpha 0,05. Jika probabilitas variabel independen  $> 0,05$ , maka secara hipotesis  $H_0$  diterima, artinya

variabel independen secara simultan (bersama-sama) tidak berpengaruh secara nyata terhadap variabel dependen. Jika probabilitas variabel independen  $< 0,05$ , maka secara hipotesis  $H_0$  ditolak atau menerima  $H_a$ , artinya variabel independen secara simultan (bersama-sama) berpengaruh terhadap variabel dependen.

### **3. Uji t-Statistik (Uji Parsial)**

Uji t dilakukan untuk melihat signifikansi dari pengaruh variabel bebas secara individual terhadap variabel terikat dengan menganggap variabel bebas lainnya adalah konstan. Adapun langkah-langkah yang dilakukan dalam uji ini sebagai berikut:

#### **a. Merumuskan Hipotesis**

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ , artinya tidak ada pengaruh secara individu variabel independen terhadap variabel dependen.

$H_a: \beta_1 : \beta_2 : \beta_3 : \beta_4 \neq 0$ , artinya ada pengaruh secara individu variabel independen terhadap variabel dependen

### **4. Pengambilan Keputusan**

Pengambilan keputusan dalam uji T dilakukan dengan membandingkan probabilitas variabel independen terhadap variabel dependen dengan nilai alpha yang digunakan, dalam penelitian ini penulis menggunakan alpha 0,05.

Jika probabilitas variabel independen  $> 0,05$ , maka secara hipotesis  $H_0$  diterima, artinya variabel independen secara partial (sendiri) tidak berpengaruh secara nyata terhadap variabel dependen.

Jika probabilitas variabel independen  $< 0,05$ , maka secara hipotesis  $H_0$  ditolak atau menerima  $H_a$ , artinya variabel independen secara partial (sendiri) berpengaruh secara nyata terhadap variabel dependen.

Uji ini dapat dilakukan dengan membandingkan  $t$  hitung dengan  $t$  tabel. Adapun rumus untuk mendapatkan  $t$  hitung adalah sebagai berikut:

$$t \text{ hitung} = (b_i - b) / sb_i$$

Dimana:

$b_i$  = koefisien variabel independen ke- $i$

$b$  = nilai hipotesis nol

$sb_i$  = simpangan baku dari variabel independen ke- $i$

Pada tingkat signifikasnsi 5% dengan kriteria pengujian yang dilakukan sebagai berikut:

- a. Jika  $t$  hitung  $< t$  tabel maka  $H_0$  diterima dan  $H_1$  ditolak, yang artinya salah satu variabel bebas (*independent*) tidak mempengaruhi variabel terikat (*dependent*) secara signifikan.
- b. Jika  $t$  hitung  $> t$  tabel maka  $H_0$  ditolak dan  $H_1$  diterima, yang artinya salah satu variabel bebas (*independent*) mempengaruhi variabel terikat (*dependent*) secara signifikan.

