

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Jenis dan cara pengumpulan data**

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu data yang diperoleh melalui data yang telah diteliti dan dikumpulkan oleh pihak lain yang berkaitan dengan permasalahan penelitian data sekunder dapat diperoleh dari berbagai sumber seperti Badan pusat statistik Daerah Istimewa Yogyakarta, Dinas Transmigrasi dan Tenaga kerja Daerah Istimewa Yogyakarta, artikel, buku dan website yang mendukung penelitian ini. Data sekunder dalam penelitian ini adalah variabel dependen jumlah pengangguran terdidik di Daerah Istimewa Yogyakarta dan Variabel Independen adalah Investasi asing, investasi dalam negeri, presentase lulusan SMA/SMK dan UMK.

#### **3.2 Variabel dan Definisi Variabel**

##### **3.2.1 Variabel Penelitian**

Penelitian ini bertujuan untuk mengetahui factor yang mempengaruhi pengangguran terdidik yang ada di Provinsi Yogyakarta tahun 2011-2016 mengambil data dari BPS DIY.

##### **3.2.2 Variabel Dependen**

Variabel dependen dalam penelitian ini adalah

## 1. Pengangguran Terdidik (Y)

Menurut BPS (2008) tingkat pengangguran terdidik adalah rasio jumlah pencari kerja yang berpendidikan SLTA ke atas sebagai kelompok terhadap besarnya angkatan kerja pada kelompok tersebut.

### 3.2.3 Variabel Independen

#### 1. Investasi Asing

Investasi Penanaman Modal Asing adalah realisasi kumulatif investasi penanaman modal asing (PMA) Kabupaten/Kota di Provinsi Yogyakarta selama periode 2011-2016. Kabupaten Kulon Progo, Kabupaten Sleman, Kabupaten Gunung Kidul, Kabupaten Bantul dan Kota Yogyakarta yang bersumber dari Badan Pusat Statistik Provinsi Yogyakarta dengan satuan ribu rupiah.

#### 2. Investasi Dalam Negeri

Investasi Penanaman Modal Dalam Negeri adalah realisasi kumulatif investasi penanaman Modal Dalam Negeri Kabupaten/Kota di Provinsi Yogyakarta selama periode 2011-2016. Kabupaten Kulon Progo, Kabupaten Sleman, Kabupaten Gunung Kidul, Kabupaten Bantul dan Kota Yogyakarta yang bersumber dari Badan Pusat Statistik Provinsi Yogyakarta dengan satuan ribu rupiah.

### 3. Presentase kelulusan Sekolah SMA/SMK

Presentase kelulusan Sekolah SMA/SMK Kabupten/Kota di Provinsi Yogyakarta adalah menghitung jumlah lulusan SMA/SMK dibagi jumlah pengangguran terdidik dikali 100 persen selama periode 2011-2016. Diantaranya Kabupaten Bantul, Kabupaten Sleman, Kabupaten Kulon Progo, Kabupaten Gunung Kidul dan Kota Yogyakarta yang bersumber dari Badan Pusat Statistik Provinsi Yogyakarta dengan satuan persen.

### 4. UMK

Upah Minimum adalah upah bulanan terendah yang terdiri atas upah pokok termasuk tunjangan tetap yang ditetapkan oleh gubernur sebagai jaring pengaman. UMK di Kabupten/Kota di Provinsi Yogyakarta selama periode 2011-2016. Diantaranya Kabupaten Bantul, Kabupaten Sleman, Kabupaten Kulon Progo, Kabupaten Gunung Kidul dan Kota Yogyakarta yang bersumber dari Badan Pusat Statistik Provinsi Yogyakarta dengan satuan rupiah..

## 3.3 Metode analisis data

### 3.3.1 Data Panel

Untuk metode analisis data, menggunakan data panel dan alat pengolah data yang digunakan adalah *evIEWS 9*. Data panel adalah gabungan antara data runtut waktu (*time series*) dan data silang (*cross section*). Menurut Agus Widarjono (2013) adalah dalam penggunaan data panel memiliki observasi yang memiliki beberapa keuntungan yang di dapatkan antara lain adalah

- a. Data panel adalah merupakan sebuah gabungan dari data *time series* dan *cross section* yang akan menyediakan sebuah data yang lebih banyak akibatnya akan lebih mendapatkan hasil *degree of freedom* dengan jumlah yang lebih besar.
- b. Dengan menggabungkan berbagai informasi dari sebuah dan *time series* dan *cross section* dapat mengatasi berbagai permasalahan yang ditimbulkan oleh ada masalah-masalah penghilangan variable.

Menurut Suliyanto (2011) kelebihan data panel dibanding data *time series* dan *cross section* yaitu.

- a. Panel data memiliki tingkat heterogenitas yang lebih tinggi oleh karena itu data tersebut melihat beberapa individu dalam waktu beberapa waktu. Dengan panel kita dapat mengestimasi karakteristik untuk setiap individu berdasarkan heterogenitasnya.
- b. Panel data dapat memberikan data yang lebih informatif, lebih bervariasi, serta memiliki tingkat kolinieritas yang rendah. Hal ini disebabkan oleh gabungan dari data *time series* dan data *cross section*.
- c. Data panel paling cocok untuk perubahan study dinamis karena panel data pada dasarnya adalah data *cross section* yang diulang-ulang (*series*).
- d. Panel data dapat mendeteksi dan mengukur pengaruh yang tidak dapat diobservasi dengan data *time series* murni atau data *cross section* murni.

- e. Panel data mampu mempelajari model perilaku yang lebih kompleks. antara lain fenomena perubahan teknologi lebih baik observasinya dengan menggunakan data panel dari pada dengan *cross section* atau sekedar *time series* saja.

Menurut sulisyanto berdasarkan keseimbangan data, data panel dapat dikelompokkan menjadi dua adalah

- a. Panel data seimbang

Panel data seimbang jika setiap *cross section* memiliki jumlah observasi time misalnya ada lima unit *cross section*, masing-masing individu memiliki time series sebanyak sepuluh tahun.

- b. Panel data tidak seimbang.

Panel data tidak seimbang jika setiap unitnya tidak memiliki jumlah observasi *time series* yang sama. Misalnya ada lima unit *cross section*, masing-masing individu observasinya *time series*. Ada yang memiliki observasi delapan tahun, sembilan tahun dan sepuluh tahun.

Berikut adalah susunan persamaan untuk estimasi dengan menggunakan model regresi data panel.

Model persamaan data panel yang merupakan gabungan dari data *cross section* dan data *time series* adalah sebagai berikut:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{3i} + \beta_n X_{nit} + e_{it}$$

dimana:

$Y_{it}$  = Pengangguran Terdidik

$X_{1it}$  = Investasi Asing

$X_{2it}$  = Investasi Dalam Negeri

$X_{3it}$  = Presentase lulusan SMA/SMK

$X_{4it}$  = UMK

$\beta_0$  :Konstanta

$i$  = Kabupaten/kota

$t$  = Tahun

$e$  =Residual

### 3.3.2 Metode *Common Effects Model/Pooled Least Square (PLS)*

Menurut Sriyana (2014) Asumsi pertama yang dikenalkan dalam regresi data panel dengan metode *common effects/ PLS* adalah asumsi yang menganggap bahwa intersep dan slope selalu tetap baik antar waktu maupun antar individu. Setiap individu ( $n$ ) yang diregresi untuk mengetahui hubungan antara variabel dependen dengan variabel independennya akan memberikan nilai intersep maupun *slope* yang sama besarnya. Begitupula dengan waktu ( $t$ ), nilai *intersep* dan *slope* dalam persamaan regresi yang menggambarkan hubungan antara variabel dependen dan variabel-variabel independennya adalah sama untuk setiap waktu. Hal ini dikarenakan dasar yang digunakan dalam regresi data panel ini yang mengabaikan pengaruh individu dan waktu pada model yang dibentuknya. Sistematis model *common effects* adalah menggabungkan antara data *time series* dan data *cross-section* ke dalam data panel (*pooled data*). Dari data tersebut kemudian di regresi dengan metode OLS, dengan melakukan regresi semacam ini maka hasilnya tidak dapat mengetahui perbedaan baik antar individu maupun antar waktu disebabkan oleh pendekatan yang digunakan mengabaikan dimensi individu maupun waktu yang mungkin saja memiliki pengaruh.

Regresi model *commom effects* ini berasumsi bahwa *intersep* dan *slope* adalah tetap sepanjang waktu dan individu, adanya perbedaan *intersep* dan *slope* diasumsikan akan dijelaskan oleh variabel gangguan (*error* atau *residual*). Dalam persamaan matematis asumsi tersebut dapat dituliskan  $\beta_0$  (*slope*) dan  $\beta_k$  (*intersep*) akan sama (konstan) untuk setiap data *time series* dan *cross section*. Persamaan matematis untuk model *common effects* akan mengestimasi  $\beta_1$  dan  $\beta_k$  dengan model berikut:

$$Y_{it} = \beta_0 + \sum_{k=1}^n \beta_k X_{kit} + \varepsilon_{it}$$

Keterangan:

$i$  : adalah banyaknya observasi

(1,2,...,n)  $t$  : adalah banyaknya waktu

(1,2,...,t)  $n \times t$  : adalah banyaknya data panel

$\varepsilon$  : adalah residual

### 3.3.3 Metode *Fixed Effect Model* (FEM)

Model (*fixed effects*) adalah suatu regresi yang menunjukkan perbedaan konstanta antar obyek, meskipun dengan koefisien regresi yang sama (Sriyana, 2014). Ada 2 asumsi yang ada dalam model regresi (*fixed effects*) yaitu:

1. Asumsi *slope* konstan tetapi *intersep* bervariasi antar unit.

*Intersep* pada suatu hasil regresi sangat mungkin berubah untuk setiap individu dan waktu, pada pendekatan (*fixed effects*) metode dapat dilakukan dengan variabel semu (*dummy*) untuk menjelaskan adanya perbedaan antar intersep. Model ini dapat di regresi dengan teknik *Least Squares Dummy Variables* (LSDV).

2. Asumsi slope konstan tetapi intersep bervariasi antar individu/unit dan antar periode waktu, Pendekatan dari metode estimasi regresi data panel ini adalah asumsi tentang intersep yang berubah baik antar individu obyek analisis maupun antar waktu, namun slope masih diasumsikan konstan/sama.

Persamaan model ini adalah sebagai berikut:

$$Y_{it} = \alpha_i + \beta_j X_{it}^j + \sum_{i=2}^n \alpha_i D_i + U_{it}$$

$i = 1, \dots, N$  dan  $t = 1, \dots, K$

Keterangan :

$Y_{it}$  : adalah variabel terikat di waktu  $t$  untuk unit cross section  $I$

$\alpha_i$  : adalah intersep yang berubah – ubah antar unit cross section

$X_{it}^j$  : adalah variabel bebas ke- $j$  di waktu  $t$  untuk unit cross section  $I$

$\beta_j$

$U_{it}$  : adalah parameter untuk variabel bebas ke- $j$

: adalah komponen error di waktu  $t$  untuk unit cross section  $I$

Keputusan memasukan variabel  $D_i$ ) pada pendekatan *fixed effects* tidak dapat dipungkiri akan mengurangi jumlah *degree of freedom* yang pada akhirnya akan mempengaruhi efisiensi dari parameter yang diestimasi.

### 3.3.4 Metode *Random Effects* Model (REM)

Model ini lebih dikenal sebagai model *Generalized Least Squares* (GLS). Model ini di asumsikan bahwa perbedaan intersep dan konstanta disebabkan residual atau *error*, sebagai akibat perbedaan antar unit dan antar periode waktu yang terjadi secara random. Karena hal inilah, model efek acak (*random effects*) sering juga disebut model komponen *error* (*error component model*) (Sriyana, 2014).

Ada 2 asumsi yang digunakan dalam *model random effects* (REM) adalah

1. *Intersep* dan *slope* berbeda antar individu, pada asumsi ini intersep dan slope yang di analisis hanya dilihat dari perbedaan antar obyek antar individu saja, adanya perbedaan intersep dan koefisien regresi berdasarkan perubahan waktu masih dikesampingkan.
2. *Intersep* dan *slope* berbeda antar individu/unit dan periode waktu, asumsi ini menjelaskan adanya perbedaan hasil estimasi intersep dan slope yang di analisis terjadi karena perbedaan antar obyek individu analisis sekaligus karena adanya perubahan antar periode waktu.

## 3.4 Pemilihan Uji data

### 3.3.4 Pemilihan model data panel

### a. Chow Test

*Chow test* (Uji Chow) yakni pengujian untuk menentukan model *Fixed Effects* atau *Random Effects* yang paling tepat digunakan dalam mengestimasi data panel.

#### Hipotesis

H0 : *Common Effects Model* atau pooled OLS

H1 : *Fixed Effects Model*

Dasar penolakan terhadap hipotesis diatas adalah dengan membandingkan perhitungan F-statistik dengan F-tabel. Perbandingan dipakai apabila hasil F hitung lebih besar (>) dari F tabel maka H0 ditolak yang berarti model yang paling tepat digunakan adalah *Fixed Effect Model*. Begitupun sebaliknya, jika F hitung lebih kecil (<) dari F table maka H0 diterima dan model yang digunakan adalah *Common Effect Model* (Widarjono, 2009). Perhitungan F statistik didapat dari Uji Chow dengan rumus (Baltagi, 2005):

positif dan signifikan dugaan hipotesis Positif sementara terbukti benar adanya.

$$F = \frac{((SEE_1 - SEE_2) / (n - 1))}{((SEE_2) / (n - t - k))}$$

Dimana:

SSE1 : Sum Square Error dari model *Common Effects*

SSE2 : Sum Square Error dari model *Fixed Effects*

n : Jumlah Kabupaten/kota (*cross section*)

nt : Jumlah *cross section* x jumlah *time series*

k : Jumlah variabel independent

### **b. Uji Hausman**

Pengujian ini membandingkan model *fixed effects* dengan *random effects* dalam menentukan model yang terbaik untuk digunakan sebagai model regresi data panel (Gujarati, 2012). Hausman *test* menggunakan program yang serupa dengan Chow *test* yaitu program *Eviews9*. Hipotesis yang dibentuk dalam Hausman *test* adalah sebagai berikut :

$H_0$  : Model *Random Effects*

$H_1$  : Model *Fixed Effects*

$H_0$  ditolak jika *P-value* lebih kecil dari nilai  $\alpha$ . Sebaliknya,  $H_0$  diterima jika *P-value* lebih besar dari nilai  $\alpha$ . Nilai  $\alpha$  yang digunakan sebesar 5%..

### **3.3.5 Uji statistik**

#### **a) Uji determinan $R^2$**

Dalam sebuah fungsi  $R^2$  adalah fungsi dengan yang tidak pernah menurunkan dari penjelasan yang ada jumlah variabelnya, kecuali variabel yang di tambah dengan regresor lainnya secara sempurna. Koefisien determinan ( $R^2$ ) untuk menjelaskan seberapa besar proporsi variabel dependent di jelaskan oleh variabel independen. Dalam hal ini kita dapat mengukur seberapa besar proporsi variabel dependen di jelaskan oleh semua variabel independen. Nilai  $R^2$  berkisaran antara  $0 < R^2 < 1$  semakin besar  $R^2$  semakin bisa menjelaskan hubungan variabel dependen dan variabel independen.

## **b) Uji t**

Uji t digunakan untuk menguji pengaruh variabel independen terhadap variabel dependen secara individu, dapat dibuat hipotesis sebagai berikut :

### **a. Variabel Investasi Asing**

$H_0 : \beta_1 \geq 0$ , artinya tidak ada pengaruh variabel jumlah Investasi Asing terhadap variabel Pengangguran Terdidik.

$H_1 : \beta_1 < 0$ , artinya terdapat pengaruh variabel jumlah Investasi Asing terhadap variabel Pengangguran Terdidik

### **b. Variabel Investasi Dalam Negeri**

$H_0 : \beta_2 \geq 0$ , artinya tidak ada pengaruh variabel Investasi Dalam Negeri terhadap variabel Pengangguran Terdidik

$H_1 : \beta_2 < 0$ , artinya terdapat pengaruh variabel Investasi Dalam Negeri terhadap variabel Pengangguran Terdidik

### **c. Variabel Presentase Kelulusan SMA/SMK**

$H_0 : \beta_3 \geq 0$ , artinya tidak ada pengaruh variable Presentase Kelulusan SMA/SMK terhadap variable

Pengangguran Terdidik

$H_1 : \beta_3 < 0$ , artinya terdapat pengaruh variable Presentase Kelulusan SMA/SMK terhadap variable

Pengangguran Terdidik

### **d. Variabel UMK**

$H_0 : \beta_4 \geq 0$ , artinya tidak ada pengaruh variable UMK variable Pengangguran Terdidik

$H_1 : \beta_4 < 0$ , artinya terdapat pengaruh variable UMK terhadap variable Pengangguran Terdidik.

Dalam Uji t ini dilakukan dengan membandingkan t hitung dengan t tabel. Apabila t hitung  $>$  t kritis, maka  $H_0$  ditolak maka variabel independen secara individual mempengaruhi variabel dependen. Sebaliknya apabila t hitung  $<$  t kritis maka variabel independen secara individual tidak mempengaruhi variabel dependen (Widarjono, 2009)

c) Uji F

Uji F dilakukan untuk mengetahui apakah variabel-variabel independen secara keseluruhan signifikan secara statistik dalam mempengaruhi variabel dependen. Apabila nilai F hitung lebih besar dari nilai F kritis maka variabel-variabel independen secara keseluruhan berpengaruh terhadap variabel dependen (Widarjono, 2009).

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

$H_1$ : minimal ada satu koefisien regresi tidak sama dengan nol

Dengan membandingkan nilai prob f-stat dengan  $\alpha$  ( $0,05=5\%$ ), jika prob f-statistik  $<$   $\alpha$  maka menolak  $H_0$  maka variabel independen secara serentak mempengaruhi variabel dependen. Sebaliknya apabila prob f-statistik  $>$   $\alpha$  maka variabel independen secara serentak tidak mempengaruhi variabel dependen.