

## **BAB III**

### **METODOLOGI PENELITIAN**

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

Berdasarkan masalah-masalah yang telah peneliti rumuskan sebelumnya, maka tujuan penelitian ini adalah untuk mendapatkan pengetahuan yang tepat (shahih, benar, valid) dan dapat dipercaya (dapat diandalkan atau reliabel) untuk mengetahui sejauh mana pengaruh jumlah penduduk, tingkat pendidikan, dan upah minimum terhadap pengangguran di Provinsi Banten.

#### **B. Waktu dan Tempat Penelitian**

Waktu penelitian berlangsung pada tahun 2010-2014 berdasarkan data yang diperoleh di lembaga statistik. Alasan dilakukan penelitian pada waktu tersebut, karena keterbatasan data diikuti dengan meningkatnya jumlah pengangguran dengan tingkat pendidikan masyarakat yang semakin meningkat.

Penelitian akan dilakukan di Provinsi Banten. Provinsi Banten dipilih karena provinsi Banten memiliki jumlah penduduk yang meningkat pesat setiap tahunnya, tingkat pendidikan di Provinsi Banten juga mengalami kenaikan, namun pengangguran yang ada di Provinsi Banten masih tinggi dibandingkan dengan provinsi-provinsi lain di Indonesia. Selain itu juga karena angkatan kerja yang meningkat yang tidak didukung dengan kesempatan kerja yang tersedia, hal ini perlu mendapat perhatian serius dari pemerintah. Sehingga dirasakan sesuai dengan judul penelitian.

### **C. Metode Penelitian**

Metode yang digunakan dalam penelitian ini adalah metode *expos facto*, yang merupakan suatu penelitian yang dilakukan untuk meneliti peristiwa yang terjadi dan kemudian menurut ke belakang untuk mengetahui faktor yang dapat menimbulkan kejadian tersebut. Metode ini (*expos facto*) digunakan untuk memperoleh data sekunder dan dengan mengambil data sekunder yang diperoleh dari badan atau lembaga pengolah data.

Sedangkan pendekatan yang digunakan adalah pendekatan regresi berganda karena banyaknya faktor yang mempengaruhi variabel tak bebas, yaitu pengangguran sebagai variabel terikat, tingkat upah sebagai variabel bebas pertama, pertumbuhan penduduk sebagai variabel bebas kedua, dan tingkat pendidikan sebagai variabel bebas ketiga. Metode ini dipilih karena masalah yang ditemukan oleh peneliti merupakan masalah yang dapat diperoleh melalui data badan pusat statistik melalui data pada tahun 2010-2014 di Provinsi Banten. Untuk mengetahui seberapa besar pengaruh antara jumlah penduduk, tingkat pendidikan dan upah minimum terhadap tingkat pengangguran terbuka.

### **D. Jenis dan Sumber Data**

Jenis data yang digunakan adalah data sekunder yakni data lima tahun pada 2010-2014 dari delapan kabupaten/kota di Provinsi Banten. Data yang dikumpulkan mencakup data-data di Provinsi Banten dari tahun ke tahun. Dengan kata lain data yang digunakan adalah data panel, yang menggabungkan data *cross section* dan data *time series*. Penggunaan data tahunan di Provinsi Banten dipilih

untuk melihat jumlah pengangguran, jumlah penduduk, tingkat pendidikan dan upah minimum. Data sekunder tersebut diperoleh dari Badan Pusat Statistik (BPS) dan Permendagri.

## **E. Operasionalisasi Variabel Penelitian**

### **1. Pengangguran (Variabel Y)**

#### a) Definisi Konseptual

Pengangguran adalah seseorang yang sudah digolongkan dalam angkatan kerja, yang secara aktif sedang mencari pekerjaan pada suatu tingkat upah tertentu, tetapi tidak dapat memperolehnya.

#### b) Definisi Operasional

Pengangguran merupakan seseorang dalam usia kerja yang tidak memiliki pekerjaan dan sedang mencari pekerjaan di Provinsi Banten tahun 2010-2014 yang diukur dalam satuan jiwa. Data diambil dari BPS.

### **2. Jumlah Penduduk (Variabel X1)**

#### a) Definisi Konseptual

Penduduk adalah orang, baik sebagai perseorangan (individu) maupun sebagai kelompok yang bertempat tinggal dan menetap disuatu wilayah.

#### b) Definisi Operasional

Penduduk adalah orang yang bertempat tinggal menetap di suatu wilayah yang diukur dengan menggunakan data jumlah penduduk di Provinsi Banten tahun 2010 – 2014 yang diukur dalam satuan jiwa dan data diambil dari BPS.

### **3. Tingkat Pendidikan (Variabel X2)**

#### a) Definisi Konseptual

Tingkat pendidikan SMA adalah tingkat pendidikan sekolah lanjutan tingkat atas yang dicapai penduduk suatu negara.

#### b) Definisi Operasional

Tingkat pendidikan adalah jenjang pendidikan tertinggi yang ditamatkan oleh penduduk, data tingkat pendidikan yang diambil merupakan lulusan Sekolah Menengah Atas (SMA) Kabupaten/Kota di Provinsi Banten tahun 2010-2014 yang datanya diambil dari BPS (Permendagri).

### **4. Upah Minimum (Variabel X3)**

#### a) Definisi Konseptual

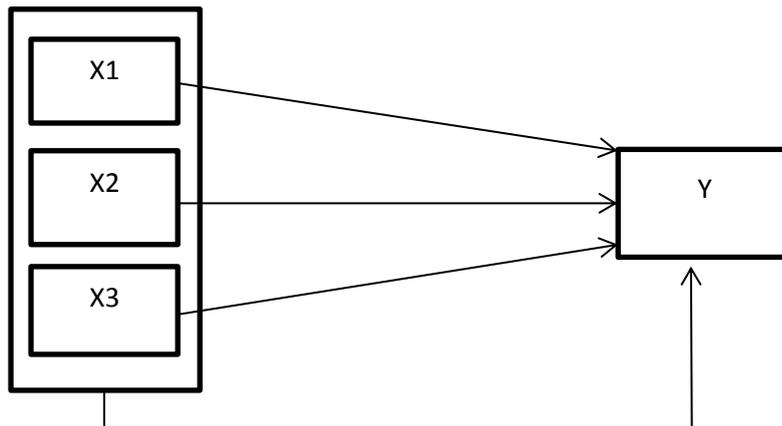
Upah minimum merupakan imbalan minimum yang diterima seseorang atas jasa kerja yang diberikannya bagi pihak lain, diberikan seluruhnya dalam bentuk uang atau sebagian dalam bentuk uang dan sebagian dalam bentuk natura.

#### b) Definisi Operasional

Upah minimum kabupaten/kota adalah upah minimum yang berlaku di daerah kabupaten/kota, yang diterima oleh pekerja per bulan yang diukur dalam satuan rupiah dengan data upah minimum yang berlaku di Provinsi Banten tahun 2010-2014 dan datanya diambil dari BPS.

## F. Konstelasi Pengaruh antar Variabel

Konstelasi pengaruh antar variabel dalam penelitian ini bertujuan untuk memberikan arah atau gambaran dari penelitian ini, yang dapat digambarkan sebagai berikut:



**Gambar III. 1**  
**Konstelasi Pengaruh antar Variabel**

Keterangan:

- X1 : Jumlah penduduk
- X2 : tingkat pendidikan
- X3 : upah minimum
- Y : pengangguran
- : arah hubungan

## G. Teknik Analisis Data

### 1. Analisis Data Panel

Data panel (*panel pooled data*) merupakan gabungan data *time series* dan data *cross section*. Dengan kata lain, data panel merupakan unit-unit individu yang sama yang diamati dalam kurun waktu tertentu. Jika kita memiliki T periode waktu ( $t = 1, 2, \dots, T$ ) dan N jumlah individu ( $i = 1, 2, \dots, N$ ), maka dengan data panel kita akan memiliki total unit observasi sebanyak NT. Penggunaan data panel

pada dasarnya merupakan solusi akan ketidaksediaan data time series yang cukup panjang untuk kepentingan analisis ekonometrika.

Terdapat tiga metode pada estimasi model menggunakan data panel, yaitu model *common effect*, *fixed effect*, dan *random effect*.

a. Model *common effect*

Model *common effect* merupakan model regresi data panel yang paling sederhana. Model ini pada dasarnya mengabaikan struktur panel dari data, sehingga diasumsikan bahwa perilaku antar individu sama dalam berbagai kurun waktu atau dengan kata lain pengaruh spesifik dari masing-masing individu diabaikan atau dianggap tidak ada. Dengan demikian, akan dihasilkan sebuah persamaan regresi yang sama untuk setiap unit cross section. Persamaan regresi untuk model *common effect* dapat dituliskan sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it} \quad i = 1, 2, \dots, N \quad t = 1, 2, \dots, T$$

Keterangan:

- Y : variabel dependen
- $\alpha$  : koefisien regresi
- X : variabel independen
- $\beta$  : estimasi parameter (koefisien)
- $\varepsilon$  : *error term*
- N : jumlah (individu)
- T : jumlah periode waktu

b. Model *Fixed effect*

Jika model *common effect* cenderung mengabaikan struktur panel dari data dan pengaruh spesifik masing-masing individu, maka model *fixed effect* adalah sebaliknya. Pada model ini, terdapat model spesifik individu  $\alpha_i$  dan diasumsikan berkorelasi dengan variabel penjelas yang teramati  $X_{it}$ .

Secara umum, pendekatan *fixed effect* dapat dituliskan dalam persamaan sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + W_{ti} + \delta Z_{it} + \varepsilon_{it}$$

Keterangan:

$Y_{it}$	: variabel terikat untuk individu ke-i dan waktu ke-t
$X_{it}$	: variabel bebas untuk individu ke-i dan waktu ke-t
$W_{it}$ dan $Z_{it}$	: variabel dummy yang didefinisikan sebagai berikut
$W_{it}$	: 1; untuk individu i, $i = 1, 2, \dots, N = 0$ ; lainnya
$Z_{it}$	: 1; untuk individu t; $t = 1, 2, \dots, N = 0$ ; lainnya
$\varepsilon_{it}$	: <i>error term</i> untuk individu ke-i dan waktu ke-t

c. Model *Random effect*

Pada model *random effect*, efek spesifik dari masing-masing individu  $\alpha_i$  diperlakukan sebagai bagian dari komponen error yang bersifat acak dan berkorelasi dengan variabel penjelas yang teramati  $X_{it}$ . Dengan demikian, persamaan model *random effect* dapat dituliskan sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}; \varepsilon_{it} = u_i + v_i + w_{it}$$

Keterangan:

$u_i$	: komponen <i>error cross section</i>
$v_i$	: komponen <i>error time series</i>
$w_{it}$	: komponen <i>error gabungan</i>

Asumsi dasar model ini adalah perbedaan nilai intersep antar unit cross section dimasukkan ke dalam error. Karena hal ini, model *random effect* sering disebut dengan *Error Component Model (ECM)*. Model ini diestimasi dengan metode *Generalized Least Square (GLS)*. Intersep model ini bervariasi terhadap individu dan waktu, namun slopenya konstan terhadap individu dan waktu. Penggunaan pendekatan *random effect* tidak mengurangi derajat kebebasan

sebagaimana terjadi pada model *fixed effect* yang akan berakibat pada parameter hasil estimasi akan menjadi lebih efisien.

#### d. Uji Kesesuaian Model

Untuk mengetahui kesesuaian atau kebaikan model dari ketiga metode pada teknik estimasi model dengan data panel digunakan *Chow test*, *LM test*, dan *Hausman test* yang digunakan untuk menguji kesesuaian model antara model yang diperoleh dari *common effect* dengan model yang diperoleh dari metode *fixed effect*. *LM test* digunakan untuk menguji kesesuaian model yang diperoleh dari *common effect* dengan model yang terbaik yang diperoleh dari metode *random effect*. Sedangkan *Hausman test* digunakan terhadap model yang terbaik yang diperoleh dari metode *fixed effect* dengan model yang diperoleh dari *random effect*.<sup>1</sup>

## 2. Uji Persyaratan Analisis

Uji persyaratan analisis yang digunakan adalah uji normalitas. Uji normalitas bertujuan untuk mengetahui apakah variabel independen dan variabel dependen dalam model regresi mempunyai distribusi normal atau tidak. Pengujian normalitas dalam penelitian ini dilakukan dengan *uji jarque-bera*. Uji *jarque-bera* memiliki nilai probability, jika hasil uji *jarque-bera* lebih besar dari probability pada signifikan 0,05 maka data berdistribusi normal. Model regresi yang baik hendaknya berdistribusi normal.

---

<sup>1</sup> Winarno, *Analisis Ekonometrika dan Statistika dengan Eviews* (Yogyakarta: UPP STIM YKPM, 2007), hal.21.

### 3. Deteksi Asumsi Klasik

#### a. Deteksi Multikolinearitas

Multikolinearitas digunakan untuk mengetahui ada atau tidaknya penyimpangan asumsi klasik multikolinearitas, yaitu adanya hubungan linear antar variabel independen dalam model regresi. Prasyarat yang harus terpenuhi dalam model regresi adalah tidak adanya multikolinearitas. Ada beberapa metode pengujian yang dapat digunakan, diantaranya; (1) dengan melihat *inflation factor* (*VIF*) pada model regresi, (2) dengan membandingkan nilai koefisien determinasi individual ( $r^2$ ) dengan nilai determinasi serentak ( $R^2$ ) dan dengan melihat nilai eigenvalue dan condition index. Pada uji ini menggunakan VIF pada model regresi. Santoso mengungkapkan bahwa “pada umumnya jika VIF lebih besar dari 5, maka variabel tersebut mempunyai persoalan multikolinearitas”.<sup>2</sup> Ghozali mengungkapkan bahwa “Dalam model regresi yang baik seharusnya tidak terjadi korelasi diantara variabel bebas”.<sup>3</sup>

#### b. Deteksi Heteroskedastisitas

Uji heteroskedastisitas dilakukan untuk menguji apakah dalam sebuah model regresi menjadi ketidaksamaan varians residual dari satu pengamatan ke pengamatan yang lain tetap, maka disebut heteroskedastisitas. Ghozali mengungkapkan bahwa “Jika titik-titik menyebar di atas dan di bawah angka 0 pada sumbu Y tanpa membentuk pola tertentu, maka tidak terjadi heteroskedastisitas”.<sup>4</sup>

---

<sup>2</sup> Duwi Priyatno, *Mandiri Belajar SPSS*, (Yogyakarta: Mediakom, 2008), hal. 39.

<sup>3</sup> Imam Ghozali, *Aplikasi Analisis Multivariate dengan Program IBM SPSS 19*, (Semarang: Badan Penerbit Universitas Diponegoro, 2011), hal.105-106.

<sup>4</sup> *Ibid.*, hal. 139.

c. Deteksi Autokorelasi

Autokorelasi merupakan gangguan pada fungsi regresi yang berupa korelasi diantara faktor gangguan. Kemungkinan adanya autokorelasi antara nilai-nilai variabel  $\varepsilon$  yang berurutan. Dalam penelitian ini, uji untuk mengetahui ada atau tidaknya autokorelasi dengan menggunakan uji *Breusch-Godfrey Serial Correlation LM Test*.

#### 4. Regresi Linier Berganda

Penelitian ini menggunakan teknik analisis regresi linear berganda tiga prediktor. Persamaan regresi yang digunakan adalah sebagai berikut:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \varepsilon$$

Untuk mencari koefisien regresi  $a$ ,  $b_1$ ,  $b_2$  dan  $b_3$  digunakan persamaan simultan sebagai berikut:

$$(i) \quad a = \hat{Y} - \beta_1X_1 - \beta_2X_2 - \beta_3X_3$$

$$(ii) \quad b_1 = \frac{n\sum x_1 - \Sigma(x_1) \Sigma(y)}{n(\Sigma x_1^2) - (\Sigma x_1)^2}$$

$$(iii) \quad b_2 = \frac{n\sum x_2 - \Sigma(x_2) \Sigma(y)}{n(\Sigma x_2^2) - (\Sigma x_2)^2}$$

$$(iv) \quad b_3 = \frac{n\sum x_3 - \Sigma(x_3) \Sigma(y)}{n(\Sigma x_3^2) - (\Sigma x_3)^2}$$

Keterangan:

$Y$  : pengangguran (variabel terikat)

$a$  : koefisien titik potong intersep

$b_1$  : koefisien regresi tingkat upah

$b_2$  : koefisien regresi jumlah penduduk

$b_3$  : koefisien regresi tingkat pendidikan

$X_1$  : jumlah penduduk

$X_2$  : tingkat pendidikan

$X_3$  : upah minimum

## 5. Uji Hipotesis

Hipotesis merupakan jawaban sementara mengenai rumusan masalah penelitian yang belum diketahui kebenarannya. Hipotesis digunakan dengan kalimat pernyataan bukan pertanyaan. Dalam hipotesis terdapat hipotesis nihil atau nol hipotesis ( $H_0$ ) yang menyatakan tidak adanya hubungan antar variabel dan hipotesis alternatif atau hipotesis kerja ( $H_1$ ) yang menyatakan ada hubungan antar variabel. Setelah ada hipotesis, langkah selanjutnya adalah menguji hipotesis. Hipotesis adalah pengujian yang bertujuan untuk mengetahui apakah kesimpulan pada sampel data berlaku untuk populasi.<sup>5</sup>

### a. Uji Signifikansi Parsial (Uji t)

Uji t digunakan untuk mengetahui apakah variabel bebas secara parsial berpengaruh signifikan terhadap variabel terikat. Dengan demikian, bagi setiap nilai koefisien regresi dapat dihitung nilai t-nya. Prayitno mengungkapkan bahwa “Untuk mencari  $t_{hitung}$  dapat dicari dengan menggunakan rumus”<sup>6</sup>:

$$t_{hitung} = \frac{b_i}{s_{b_i}}$$

Keterangan:

$b_i$  = Koefisien regresi variabel i

$s_{b_i}$  = Standar error variabel i

Hasil yang diperoleh kemudian dibandingkan dengan  $t_{tabel}$  dengan taraf signifikansi ( $\alpha$ ) adalah 0,05 dan derajat kebebasan ( $n-K$ ).

Hipotesis pengujian :

$$H_0 : \beta_1 = 0$$

---

<sup>5</sup> Duwi Priyanto, *Paham Analisis Statistik Data dengan SPSS* (Yogyakarta: Mediakom, 2010), hal. 9

<sup>6</sup> *Ibid*, hal. 68

$$H_i : \beta_1 \neq 0$$

Kriteria pengujian :

- $H_0$  diterima jika  $t_{hitung} \leq t_{tabel}$ , maka koefisien regresi dikatakan tidak signifikan. Artinya salah satu variabel independen tidak mempunyai pengaruh yang signifikan terhadap variabel dependen.
- $H_0$  diterima jika  $t_{hitung} > t_{tabel}$ , maka koefisien regresi dikatakan signifikan, artinya salah satu variabel independen mempunyai pengaruh yang signifikan terhadap variabel dependen.

b. Uji Signifikansi Simultan (Uji F)

Uji F digunakan untuk menguji apakah variabel independen secara bersama-sama berpengaruh secara signifikan terhadap variabel dependen. Metode yang digunakan dalam uji ini adalah dengan cara membandingkan antara  $F_{hitung}$  dengan  $F_{tabel}$  atau  $F_{\alpha;n+k-1;nT-n-k}$  pada tingkat kesalahan 5% dengan hipotesis:

$$\bullet H_o : \beta_1 + \beta_2 + \beta_3 = 0$$

$$\bullet H_i : \beta_1 + \beta_2 + \beta_3 \neq 0$$

Hipotesis nol ditolak jika  $F_{hitung} > F_{tabel}$ , maka seluruh variabel independen berpengaruh signifikan terhadap variabel dependen secara simultan dan sebaliknya. Prayitno mengungkapkan bahwa “Untuk menguji kedua hipotesis ini digunakan nilai statistik F yang dihitung dengan rumus sebagai berikut”<sup>7</sup>:

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

---

<sup>7</sup> Priyanto, *op.cit*, hal. 67

Keterangan:

$R^2$  : koefisien determinasi

$k$  : jumlah variabel bebas

$n$  : jumlah data

c. Koefisien Determinasi ( $R^2$ )

Koefisien Determinasi ( $R^2$ ) digunakan untuk mengukur seberapa besar kemampuan model dalam menerangkan variasi variabel dependen. Atau dengan kata lain koefisien determinasi mengukur seberapa baik model yang dibuat mendekati fenomena variabel dependen yang sebenarnya. Koefisien determinasi dirumuskan sebagai berikut:

$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

Keterangan:

ESS : jumlah kuadrat yang dijelaskan

RSS : jumlah kuadrat residual

TSS : jumlah kuadrat total

Penggunaan  $R^2$  sebagai ukuran kelayakan suatu model adalah bahwa  $R^2$  tidak pernah menurun dengan penambahan regresor, sebaliknya justru meningkat. Untuk mengatasi pernyataan ini, suatu instrumen mengukur nilai kebaikan model telah dikembangkan. Ukuran tersebut merupakan modifikasi dari  $R^2$  ini memberikan penalty bagi penambahan variabel penjelas yang tidak menurunkan residual secara signifikan. Ukuran tersebut adalah adjusted  $R^2$  yang dapat dihitung dengan rumus:

$$R^2 = 1 - (1 - R^2) \frac{nT - 1}{nT - n - k}$$

Keterangan:

$R^2$  : koefisien determinasi

$n$  : jumlah observasi

$T$  : jumlah waktu

$k$  : banyaknya variabel tanpa intersep