

BAB III

METODOLOGI PENELITIAN

A. Tujuan Penelitian

Berdasarkan masalah-masalah yang telah peneliti rumuskan, maka tujuan penelitian ini adalah untuk:

1. Mengetahui besarnya pengaruh jumlah penduduk bekerja terhadap pendapatan asli daerah.
2. Mengetahui besarnya pengaruh PDRB terhadap pendapatan asli daerah.
3. Mengetahui besarnya pengaruh jumlah penduduk bekerja dan PDRB terhadap pendapatan asli daerah.

B. Objek Penelitian

Penelitian ini dilaksanakan dengan mengambil data Jumlah Penduduk Bekerja Banten, PDRB Banten dan Pendapatan Asli Daerah Banten di BPS Pusat (Badan Pusat Statistik) yang berada di DKI Jakarta.

Objek penelitian dalam penelitian ini adalah 6 Kabupaten/Kota yang ada di Provinsi Banten. Lokasi ini dipilih karena jumlah pendapatan asli daerah di Banten merupakan terkecil kedua di Pulau Jawa setelah Provinsi D.I Yogyakarta, setelah itu Provinsi Banten merupakan provinsi penyanggah Ibukota DKI Jakarta serta merupakan daerah industri yang maju.

C. Metode Penelitian

Metode yang digunakan dalam penelitian adalah metode ekspos fakto dengan

pendekatan korelasional. Metode ini dipilih karena merupakan metode yang sistematis dan empirik. Metode ekspos fakto adalah suatu penelitian yang dilakukan untuk meneliti peristiwa yang telah terjadi dan kemudian meruntut ke belakang untuk mengetahui faktor-faktor yang dapat menimbulkan kejadian tersebut.³⁸ Data yang digunakan adalah data panel dengan menggabungkan data *time series* (rentang waktu) dan *cross section* (data silang) dari tahun 2006 hingga 2012. Waktu ini dipilih karena merupakan interval waktu yang paling baik yakni setelah Provinsi Banten terlepas oleh Provinsi Jawa Barat .

D. Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang bersifat kuantitatif yaitu data yang telah tersedia dalam bentuk angka. Sedangkan data yang digunakan dalam penelitian ini adalah data panel yang merupakan gabungan antara data *time series* (runtun waktu) dan data *cross section* (silang).³⁹ Data *time series* sebanyak tujuh tahun dari tahun 2006-2012 dan data *cross section* sebanyak 6 kabupaten/kota di Banten. Data tersebut diperoleh dari studi kepustakaan yang bersumber dari publikasi yang diterbitkan oleh Badan Pusat Statistik (BPS) untuk memperoleh data jumlah penduduk, produk domestik regional bruto dan pendapatan asli daerah.

E. Operasionalisasi Variabel Penelitian

Data pada penelitian ini diperoleh dengan cara mengumpulkan data sekunder yang didapat dari catatan jumlah penduduk Banten, PDRB Banten dan

³⁸ Sugiyono. Metode Penelitian Bisnis. Jakarta: Alfabeta. 2004. p. 7

³⁹Nachrowi, *Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan* (Jakarta: LPFE UI, 2006), p. 309

pendapatan asli daerah Banten di BPS Pusat (Badan Pusat Statistik) yang berada di DKI Jakarta.

a. Pendapatan Asli Daerah

1. Definisi Konseptual

Pendapatan asli daerah adalah semua pendapatan yang di usahakan dan dikelola oleh daerah serta bersumber dari daerah itu sendiri yang merupakan total dari hasil-hasil pajak daerah, dinas-dinas daerah, dan pendapatan lainnya diluar pendapatan non-asli.

2. Definisi Operasional

Data Pendapatan asli daerah dalam penelitian ini adalah semua penerimaan yang diperoleh dari sektor pajak daerah, retribusi daerah, hasil perusahaan milik daerah, hasil pengelolaan kekayaan daerah yang dipisahkan, dan lain-lain yang sah dari sumber-sumber penerimaan daerah yang berdasarkan UU dan selama kurun waktu dari tahun 2006 sampai 2012.

b. Jumlah Penduduk Bekerja

1. Definisi Konseptual

Tenaga kerja adalah penduduk dalam usia kerja (berusia 15-64 tahun) atau jumlah seluruh penduduk dalam suatu negara yang dapat memproduksi barang/jasa jika ada permintaan terhadap tenaga mereka, dan jika mereka mau berpartisipasi dalam aktifitas tersebut.

2. Definisi Operasional

Data Jumlah penduduk bekerja dalam penelitian ini adalah semua orang yang berusia 15-64 tahun dan sedang bekerja di provinsi Banten selama kurun waktu dari tahun 2004 sampai 2012.

c. **Produk Domestik Regional Bruto**

1. Definisi Konseptual

Produk domestik regional bruto adalah nilai brang dan jasa-jasa yang di produksi di dalam negara tersebut dalam satu tahun tertentu, barang-barang dan jasa-jasa ini di produksi bukan saja oleh perusahaan milik penduduk negara tersebut tetapi oleh penduduk negara lain yang bertempat tinggal di negara tersebut.

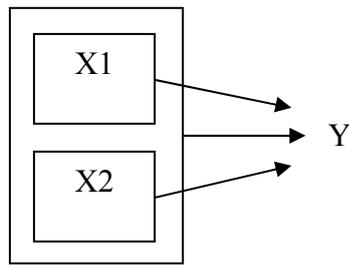
2. Definisi Operasional

Data PDRB dalam penelitian ini adalah data PDRB menurut Atas Dasar Harga berlaku atau secara nominal per kabupaten/kota di Banten dari tahun 2006-2012 yang diterbitkan oleh Badan Pusat Statistik Pusat.

F. **Konstelasi Pengaruh Antar Variabel**

Variabel penelitian ini terdiri dari tiga variabel yaitu variabel bebas (jumlah penduduk bekerja digambarkan dengan simbol X1 dan PDRB digambarkan dengan simbol X2) dan variabel terikat pendapatan asli daerah yang digambarkan dengan simbol Y.

Sesuai dengan hipotesis yang diajukan bahwa terdapat pengaruh variabel X1 dan X2 terhadap variabel Y, maka konstelasi pengaruh variabel X1 dan X2 terhadap variabel Y adalah:



Keterangan:

Variabel Bebas (X2) : PDRB

(X1) : Jumlah Penduduk Bekerja

Variabel Terikat (Y) : Pendapatan Asli Daerah

—————> : Menunjukkan Arah Pengaruh

G. Teknik Analisis Data

1. Metode Analisis

a. Analisis Data Panel

Analisis dengan menggunakan panel data adalah kombinasi antar deret waktu (*time series*) dan kerat lintang (*cross section*). Gujarati menyatakan bahwa untuk menggambarkan data panel secara singkat, misalnya pada data *cross section*, nilai dari satu variabel atau lebih dikumpulkan untuk beberapa unit sampel pada suatu waktu. Dalam data panel, unit *cross section* yang sama disurvei dalam beberapa waktu. Dalam model panel data, persamaan model dengan menggunakan data *cross section* dapat ditulis sebagai berikut :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} \quad ; \quad i = 1, 2, \dots, N \quad \dots \dots \dots (3.1)$$

dimana N adalah banyaknya data *cross section*

Sedangkan persamaan model dengan *time series* adalah

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_t \quad ; t = 1, 2, \dots, T \quad \dots\dots\dots (3.2)$$

dimana T adalah banyaknya data time-series

Mengingat data panel merupakan gabungan dari time series dan cross section, maka model dapat ditulis dengan :

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 X_{it} + \mu_{it} \quad \dots\dots\dots (3.3)$$

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

Keterangan :

- | | |
|--------------------|--|
| Y | = variabel pendapatan asli daerah |
| X1 | = jumlah penduduk bekerja |
| X2 | = PDRB |
| i | = cross section |
| t | = time series |
| β_0 | = konstanta |
| β_1, β_2 | = koefisien yang dicari untuk mengukur pengaruh variabel X1 dan X2 |
| μ | = kesalahan pengganggu |

Model tersebut dapat ditransformasikan kedalam persamaan logaritma :

$$\text{Log}Y = \beta_0 + \beta_1 \text{Log}X_1 + \beta_2 \text{Log}X_2 + \mu$$

Keterangan:

- | | |
|---|--------------------------|
| Y | = pendapatan asli daerah |
|---|--------------------------|

β_0 = Konstanta

X_1 = jumlah penduduk bekerja

X_2 = PDRB

β_1, β_2 = Koefisien yang dicari untuk mengukur pengaruh variabel X_1 dan X_2

μ = Kesalahan pengganggu

Log = logaritma

Pemilihan model ini didasarkan pada penggunaan model logaritma natural (Ln). Damodar Gujarati menyebutkan bahwa salah satu keuntungan dari penggunaan logaritma natural adalah memperkecil bagi variabel-variabel yang diukur karena penggunaan logaritma dapat memperkecil salah satu penyimpangan dalam asumsi OLS (Ordinary Least Square) yaitu heterokedastisitas.⁴⁰

Penggunaan data panel pada dasarnya merupakan solusi akan ketidaktersediaan data time series yang cukup panjang untuk kepentingan analisis ekonometrika. Menurut Hsiao dalam Greene keunggulan penggunaan data panel dibandingkan deret waktu dan kerta lintang adalah:

1) Dapat memberikan peneliti jumlah pengamatan yang besar, meningkatkan degrees of freedom (derajat kebebasan), data memiliki variabilitas yang besar dan mengurangi kolinearitas antara variabel penjelas, dimana dapat menghasilkan ekonometri yang efisien.

⁴⁰Damodar Gujarati, *Ekonometrika Dasar* (Jakarta: Erlangga, 1997)

- 2) Data panel data, data lebih informatif, lebih bervariasi, yang tidak dapat diberikan hanya oleh data cross section dan time series saja.
- 3) Panel data dapat memberikan penyelesaian yang lebih baik dalam inferensi perubahan dinamis dibandingkan data cross section.

b. Estimasi Model

1) *Model Common Effect*

Model *common effect* atau *pooled regression* 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. Sesuatu yang secara realistis tentunya kurang dapat diterima. Karena itu, model ini sangat jarang digunakan dalam analisis data panel.

Berdasarkan asumsi struktur matriks varians-covarians residual, maka pada model *common effect*, terdapat 4 metode estimasi yang dapat digunakan, yaitu:

- a) *Ordinary Least Square (OLS)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat homoskedatik dan tidak ada *cross sectional correlation*.
- b) *General Least Square (GLS)/ Weight Least Square (WLS): Cross Sectional Weight*, jika struktur matriks varians-kovarians residual diasumsikan bersifat heteroskedastik dan tidak ada *cross sectional correlation*,

c) *Feasible Generalized Least Square (FGLS)/ Seemingly Uncorrelated Regression (SUR)* atau *Maximum Likelihood Estimator (MLE)*, jika struktur matriks varians-kovarians residual diasumsikan bersifat heterokedastik dan ada *cross sectional correlation*,

d) *Feasible Generalized Least Square (FGLS)* dengan proses *autoregressive (AR)* pada error term-nya, jika struktur matriks varians-kovarians residulnya diasumsikan bersifat heteroskedastik dan ada korelasi antar waktu pada residualnya.

2) *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 efek spesifik individu α_i dan diasumsikan berkorelasi dengan variabel penjelas yang teramati X_{it} .

Berdasarkan asumsi struktur matriks varians-kovarians residual, maka pada model *fixed effect*, terdapat 3 metode estimasi yang dapat digunakan, yaitu :

1) *Ordinary Least Square (OLS/LSDV)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat homoskedatik dan tidak ada *cross sectional correlation*.

2) *Weighted Least Square (WLS)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan tidak ada *cross sectional correlation*.

3) *Seemingly Uncorrelated Regression (SUR)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan ada *cross sectional correlation*

3) **Model Random Effect**

Pendekatan ini mengasumsikan *unobservable individual effect* (u_{it}) tidak berkorelasi dengan *regressor* (X) atau dengan kata lain u_{it} diasumsikan bersifat random. Sebelum model diestimasi dengan model yang tepat, terlebih dahulu dilakukan uji spesifikasi apakah *fixed effect* atau *random effect* atau keduanya memberikan hasil yang sama.

2. **Uji Metode Estimasi data panel**

Sebelum menentukan metode estimasi data panel yang akan digunakan dalam penelitian ini, maka harus dilakukan beberapa pengujian. Untuk menentukan apakah model panel data dapat diregresi dengan metode *common effect*, metode *Fixed Effect* (FE) atau metode *Random Effect* (RE), maka dilakukan uji-uji sebagai berikut:

1) **Uji Chow**

Uji Chow dapat digunakan untuk memilih teknik dengan metode pendekatan *Pooled Least Square* (PLS) atau metode *Fixed Effect* (FE). Prosedur Uji Chow adalah sebagai berikut:

a. Buat hipotesis dari Uji Chow

- model *common effect*
- model *Fixed Effect*

b. Menentukan kriteria uji

Apabila nilai F statistik $>$ F tabel, maka hipotesis ditolak yang artinya kita harus memilih teknik FE.

Apabila nilai F statistik $<$ F tabel, maka hipotesis diterima yang artinya kita harus memilih teknik PLS.

2) Uji Hausman

Uji Hausman digunakan untuk memilih antara metode pendekatan *Fixed Effect* (FE) atau *Random Effect* (RE). Prosedur Uji Hausman adalah sebagai berikut:

- a. Buat hipotesis dari Uji Hausman: $=random\ effect$ dan $=fixed\ effect$.
- b. Menentukan kriteria uji: apabila *Chi-square* statistik $>$ *Chi-square* tabel dan *p-value* signifikan, maka hipotesis ditolak, sehingga metode FE lebih tepat untuk digunakan. Dan apabila *Chi-square* statistik $<$ *Chi-square* tabel dan *p-value* signifikan, maka hipotesis diterima, sehingga metode RE lebih tepat untuk digunakan.

3. Pengujian Asumsi Klasik

Menurut Greene “uji asumsi klasik dilakukan karena dalam model regresi perlu memperhatikan adanya penyimpangan-penyimpangan atas asumsi klasik, karena pada hakekatnya jika asumsi klasik tidak dipenuhi maka variabel-variabel yang menjelaskan akan menjadi tidak efisien.”⁴¹ Konsekuensi yang muncul ketika membangun model regresi dengan data panel adalah bertambahnya komponen residual, karena adanya dimensi cross section dan time series pada data. Kondisi ini menyebabkan matriks varian kovarian residual menjadi sedikit lebih kompleks

⁴¹ William H. Greene, *Econometric Analysis* (New York : New York University, 2002), p. 307

bila dibandingkan dengan model regresi klasik yang hanya menggunakan data *cross section* atau data *time series*.

a. Uji Normalitas

Uji normalitas bertujuan untuk menguji apakah dalam model regresi variabel bebas dan variabel terikat mempunyai distribusi normal atau tidak. Menurut Imam Ghozali, Jika data tidak berdistribusi normal maka uji statistik menjadi tidak valid dan statistik parametrik tidak dapat digunakan.⁴²

Ada beberapa metode untuk mengetahui normal atau tidak gangguan (μ) antara lain J-B test dan metode grafik. Penelitian ini akan menggunakan metode J-B test yang dilakukan dengan menghitung skewness dan kurtosis, apabila J-B hitung < nilai X^2 (chi-square) tabel, maka nilai residual berdistribusi normal. Model untuk mengetahui uji normalitas adalah :

$$JB = n \left[\frac{\mu_3^2}{6\mu_2^3} + \frac{(\mu_4 - 3)}{24} \right]$$

Keterangan :

n = jumlah sampel

2 = varians

3 = slewness

4 = kurtosis

Jarque-Bera test mempunyai distribusi chi square dengan derajat bebas dua. Jika hasil Jarque-Berra test lebih besar dari nilai chi-square pada $\alpha=5$ persen,

⁴² Imam Ghozali, *Ekonometrika Teori, Konsep dan Aplikasi dengan SPSS 17* (Semarang: Universitas Diponegoro, 2007), hal. 110

maka H_0 ditolak yang berarti tidak berdistribusi normal. Jika hasil Jarque-Beta test lebih kecil dari nilai chi square pada $\alpha=5$ persen, maka H_0 diterima yang berarti error term berdistribusi normal.

b. Uji Heterokedastisitas

Uji heteroskedastisitas berarti uji terhadap ada tidaknya heteroskedastisitas dapat dilakukan dengan menggunakan uji white. Uji white menggunakan residual kuadrat sebagai variabel dependen, dan variabel independennya terdiri atas variabel independen yang sudah ada, ditambah dengan kuadrat variabel independen, ditambah lagi dengan perkalian dua variabel independen.

Untuk melihat apakah data mengandung heteroskedastisitas dapat dilihat pada nilai probabilitas dari Obs*R-squared. Jika nilai probabilitas dari Obs*R-squared lebih besar dari $\alpha=5\%$ (0,05), maka data terbebas dari heteroskedastisitas yang artinya data bersifat homokedastisitas. Jika probabilitas dari Obs*R-squared lebih kecil dari $\alpha=5\%$ (0,05), maka data mengandung heteroskedastisitas yang artinya data tidak bersifat homokedastisitas, sehingga perlu diperbaiki.

Dalam pengambilan keputusannya adalah jika titik-titik dalam *scatterplot* membentuk suatu pola yang jelas dan teratur, maka terdapat heteroskedastisitas pada model penelitian. Namun jika titik-titik tersebar secara acak (*random*), tidak berpola, serta data menyebar di atas dan di bawah angka 0 pada sumbu Y, maka tidak terdapat heteroskedastisitas pada model penelitian.

c. Uji Multikolinearitas

Uji ini hanya digunakan untuk regresi berganda, dimana tujuannya adalah untuk melihat apakah pada model regresi ditemukan adanya korelasi antara variabel-variabel bebas. Multikolinearitas ada pada setiap persamaan regresi, disini yang akan diuji bukanlah ada atau tidaknya multikolinearitas, tetapi menentukan seberapa banyak atau parah multikolinearitas itu ada. Salah satu cara menghitung multikolinearitas adalah dengan *Variance Inflation Factor* (VIF). Menghitung *Variance Inflation Factor* untuk koefisien bisa dengan menggunakan rumus :⁴³

$$VIF = \frac{1}{(1-Rt^2)}$$

Keterangan :

R^2 = koefisien determinasi pada *auxiliary regression*

Menganalisis derajat multikolinearitas dengan cara mengevaluasi nilai VIF. Apabila nilai VIF > 10 dan tolerance < 0,1 maka terjadi multikolinearitas. Sebaliknya, jika VIF < 10 dan tolerance > 0,1 maka tidak terjadi multikolinearitas.⁴⁴

4. Pengujian Kriteria Statistik

a. Pengujian Signifikansi Simultan (Uji-F)

⁴³ Sarwoko, *Dasar-dasar Ekometrika* (Yogyakarta: ANDI, 2005), p.120

⁴⁴ Duwi Priyatno, *Buku Saku SPSS Analisis Statistik Data* (Jakarta: MediaKom, 2011), p. 288

Uji F atau uji koefisien regresi secara serentak, yaitu untuk mengetahui pengaruh variabel independen secara serentak terhadap variabel dependen, apakah pengaruhnya signifikan atau tidak.⁴⁵

Hipotesis penelitiannya:

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

Artinya variabel X1 dan X2 secara serentak tidak berpengaruh terhadap Y.

$$H_0 : \beta_1 \neq \beta_2 \neq 0$$

Artinya variabel X1 dan X2 secara serentak berpengaruh terhadap Y.

Kriteria pengambilan keputusannya, yaitu:

- a. $F_{hitung} \leq F_{tabel}$, maka H_0 diterima
- b. $F_{hitung} > F_{tabel}$, maka H_0 ditolak

Nilai F – hitung dapat diperoleh dengan rumus:

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

Keterangan:

R^2 = koefisien determinasi (residual)

K = Jumlah variabel independen ditambah intercept dari suatu model persamaan

N = jumlah sampel

b. Uji t (Partial Test)

Uji t untuk mengetahui pengaruh variabel independen secara parsial terhadap variabel dependen, apakah pengaruhnya signifikan atau tidak.⁴⁶ Selain itu, uji

⁴⁵ Duwi Priyanto, SPSS Analisa Korelasi, Regresi dan Multivariate (Yogyakarta: Gava Media, 2009), hal. 48

statistik t pada dasarnya menunjukkan seberapa jauh pengaruh satu variabel independen secara individual dalam menerangkan variasi variabel dependen. Dengan uji statistik t maka dapat diketahui apakah pengaruh masing-masing variabel independen terhadap variabel dependen sesuai hipotesis atau tidak.

1) Hipotesis pengujian :

$$H_0 : \beta_1 \leq 0$$

$$H_i : \beta_1 > 0$$

Kriteria pengujian:

- a) Jika $t_{hitung} > t_{tabel}$, H_0 ditolak, maka salah satu variabel independen mempengaruhi variabel dependen secara signifikan
- b) Jika $t_{hitung} \leq t_{tabel}$, H_0 diterima, maka salah satu variabel independen tidak mempengaruhi variabel dependen secara signifikan.

5. Uji Koefisien Determinasi (R^2)

Nilai koefisien determinasi untuk mengetahui besarnya presentasi variabel terikat yang disebabkan oleh variabel bebas. Dengan kata lain, koefisien determinasi menunjukkan ragam naik turunnya Y yang diterangkan oleh pengaruh linier X. Dalam hal ini ragam naik turunnya Y seluruhnya disebabkan oleh X. Perhitungan koefisien determinasi dapat dihitung dengan rumus:⁴⁷

⁴⁶ Duwi Priyanto, *op.cit* , p.50

⁴⁷ Nachrowi Djalal Nachrowi, *Penggunaan Teknik Ekonometrika* (Jakarta: Raja Grafindo persada, 2008), p.22

$$R^2 = \frac{EES}{TSS}$$

Keterangan :

EES (*Explained of Sum Squared*) : jumlah kuadrat yang dijelaskan

TSS (*Total Sum of Squares*) : total jumlah kuadrat

Dimana nilai R^2 terletak diantara 0 sampai dengan 1, nilai $0 \leq R^2 \leq 1$. Jika $R^2 = 0$, berarti variabel bebas tidak bisa menjelaskan variabel perubahan variabel terikat, maka model dapat dikatakan buruk. Jika $R^2 = 1$, berarti variabel bebas mampu menjelaskan variabel perubahan variabel terikat dengan sempurna. Kondisi seperti dua hal tersebut hampir sulit diperoleh. Kecocokan model dapat dikatakan lebih baik kalau R^2 semakin dekat dengan 1.