

BAB III

METODE PENELITIAN

A. Objek penelitian

Penelitian ini mencakup dua puluh lima kabupaten dan kota yang ada di Sumatera Utara. Data diambil dari Badan Pusat Statistik Sumatera Utara yang diketahui bahwa terdapat tiga puluh dua kabupaten dan kota. Tetapi, dalam penelitian ini peneliti hanya mengambil dua puluh lima kabupaten dan kota, dikarenakan ketersediaan data yang tidak lengkap.

B. Jenis Data

Pada penelitian ini, penulis menggunakan data sekunder. Data sekunder adalah data yang dikumpulkan dari berbagai sumber yang telah ada, sehingga penulis hanya menggunakan data tersebut. Data sekunder diperoleh dari beberapa sumber, seperti Badan Pusat Statistik (BPS) dan naskah publikasi BPS Sumatera Utara. Data yang didapatkan berupa Nilai Output, Jumlah unit usaha, PDRB dan penyerapan tenaga kerja sektor industri manufaktur kategori industri besar dan menengah.

Penelitian ini mengambil tahun 2011 sampai dengan 2015, dan sebagai periode akhir dalam penelitian ini tahun 2015 dikarenakan data terkini yang tersedia dari beberapa sumber hanya sampai tahun tersebut. Dengan demikian,

penelitian ini menggunakan data panel dengan basis data tahunan dari kabupaten dan kota di provinsi Sumatera Utara pada tahun 2011-2015.

C. Teknik Pengambilan Data

Data dari penelitian ini merupakan data sekunder didapatkan dari berbagai basis data dan laporan statistik yang akurat, yakni melalui situs resmi Badan Pusat Statistik (BPS).

D. Teknik Pengumpulan Data

Dalam penelitian ini penulis menggunakan teknik dokumentasi, yaitu mengambil data dan informasi terkait dengan melihat kembali laporan-laporan tertulis baik berupa angka maupun keterangan. Teknik dokumentasi dalam penelitian ini digunakan untuk mencari data nilai output, jumlah unit usaha, PDRB, dan penyerapan tenaga kerja sektor industri manufaktur kategori industri besar dan menengah

E. Definisi Variabel Operasional

Penelitian ini terdiri dari variabel dependen dan variabel independen, yang mana dalam penelitian ini memiliki satu variabel dependen dan empat variabel independen. Variabel Independen adalah variabel bebas atau bisa disebut variabel yang mempengaruhi, sedangkan variabel dependen atau variabel terikat adalah variabel yang dipengaruhi oleh variabel independen. Adapun variabel dependen dan independen dalam penelitian ini meliputi:

1. Penyerapan Tenaga Kerja

Jumlah tenaga kerja yang dimaksud adalah tenaga kerja yang bekerja pada industri besar dan menengah di provinsi Sumatera Utara dan dinyatakan dalam satuan orang data di peroleh dari naskah publikasi yang berjudul Sumatera Utara dalam angka dari tahun 2013-2017 yang diterbitkan oleh BPS Sumatera Utara.

2. Jumlah unit usaha

Jumlah unit usaha adalah total keseluruhan perusahaan Industri besar dan menengah di provinsi Sumatera Utara berdasarkan kab/kota, Jumlah unit usaha di hitung dalam (satuan). data di peroleh dari naskah publikasi yang berjudul Sumatera Utara dalam angka dari tahun 2013-2017 yang diterbitkan oleh BPS Sumatera Utara.

3. Nilai Output

Nilai Output merupakan kemampuan tenaga kerja dalam menghasilkan barang dan jasa. Nilai output di hitung dalam (000 Rp). Data di peroleh dari publikasi Badan Pusat Statistik Sumatera Utara yang berjudul karakteristik industri Besar dan sedang dari tahun 2011-2015.

4. Produk Domestik Regional Bruto (PDRB)

Produk Domestik Regional Bruto (PDRB) merupakan suatu indikator ekonomi makro yang dapat memberikan gambaran tentang keadaan perekonomian di suatu daerah atau regional. PDRB yang digunakan adalah

PDRB atas dasar harga konstan dari tahun 2011-2015 di Sumut. Data di dapat melalui naskah publikasi di BPS Sumut. Adapun satuannya (juta rupiah)

F. Alat Analisis

Alat analisis yang digunakan untuk menjawab permasalahan atau hipotesis dalam penelitian ini adalah regresi Data Panel dengan cara menguji secara statistik terhadap variabel-variabel yang telah dikumpulkan dengan menggunakan program StataSE13. Hasil analisis diharapkan dapat digunakan untuk mengetahui besarnya pengaruh beberapa variabel bebas terhadap variabel terikat.

G. Metode Analisis Data

Data panel merupakan gabungan antara data runtut waktu (*time series*) dan data silang (*cross section*). Menurut Widarjono dalam Basuki&Yuliadi (2015), penggunaan data panel dalam sebuah observasi mempunyai beberapa keuntungan yang diperoleh. Pertama, data panel yang merupakan gabungan data *time series* dan *cross selection* mampu menyediakan data yang lebih banyak sehingga akan lebih menghasilkan (*degree of freedom*) yang lebih besar. Kedua, data panel dapat mengatasi masalah yang timbul akibat masalah penghilangan variabel (*omitted variabel*).

Menurut Wibisino (2015), data panel memiliki beberapa keuntungan sebagai berikut:

1. Data panel dapat memperhitingkan heterogenitas individu secara eksplisit dengan mengizinkan variabel spesifik individu.

2. Kemampuan dalam mengontrol heterogenitas ini selanjutnya menjadikan data panel dapat digunakan untuk menguji dan membangun model perilaku lebih kompleks.
3. Cocok digunakan sebagai studi penyesuaian dinamis karena didasari oleh observasi *cross section* yang berulang-ulang.
4. Tinggi rendahnya jumlah observasi menyajikan data yang lebih informatif, variatif dan kolnieritas data semakin berkurang denga *degree of freedom* lebih tinggi sehingga hasil estimasi akan lebih baik.
5. Dapat digunakan untuk mempelajari model prilaku yang kompleks.
6. Dapat digunakan untuk meminimalisir bias yang mungkin ditimbulkan oleh agregasi data individu.

Adapun model regresi panel dalam penelitian ini sebagai berikut:

$$Y = a + b_1X_1it + b_2X_2it + b_3X_3it + e$$

Keterangan:

- Y : Penyerapan Tenaga Kerja (ton)
- α : Konstanta
- X_1 : Jumlah Unit Usaha(satuan):
- X_2 : Nilai Output (000 Rupiah)
- X_3 : PDRB (Juta Rupiah)
- e : *Error term*

t : Waktu

i : Kabupaten dan Kota

H. Model Estimasi

Metode estimasi dengan data panel dapat dilakukan dengan tiga tahap, antara lain:

1. Uji Asumsi Klasik

a. Multikolinearitas

Uji multikolinearitas bertujuan untuk menguji apakah model regresi di temukan adanya hubungan antar variabel bebas atau independen. Model regresi yang baik seharusnya tidak terjadi korelasi diantara variabel independen. Jika terjadi multikolinearitas dalam model, estimator masih bersifat *Best Linear Unbiased Estimator* (BLUE) namun estimator mempunyai varian dan ovarian yang besar sehingga sulit didapatkan estimasi yang tepat (Widarjono, 2013).

Uji multikolinearitas adalah sebuah situasi yang menunjukkan adanya korelasi atau hubungan antara variabel penjelas dan saling bergerak satu sama lain. Di dalam praktiknya, sulit untuk melakukan perkiraan yang dapat diandalkan dari masing-masing koefisien individu dan melihat besarnya kesalahan pada perkiraan. Dengan kata lain, hal tersebut akan berakibat pada salahnya kesimpulan hubungan antar variabel. Multikolinearitas meningkatkan varian parameter perkiraan sehingga dapat

menyebabkan kurangnya signifikansi variabel penjelas walaupun model yang digunakan benar. Aturan dalam uji multikolinearitas adalah jika VIF melebihi 5 atau 10 maka hasil regresi mengandung multikolinieritas (Monthgomery, 2001).

b. Heteroskedisitas

Uji Heteroskedastisitas bertujuan untuk menguji apakah model regresi mempunyai variabel pengganggu yang tidak konstan atau heteroskedastisitas. Model regresi yang baik mengandung homokedastisitas atau tidak terjadi heteroskedastistas. Gejala ini lebih sering terjadi pada data *cross section* (Widarjono, 2013). Varian variabel pengganggu yang tidak konstan atau heterokedastisitas disebabkan oleh residual pada variabel independen di dalam model. Adapun bentuk fungsi variabel gangguan adalah sebagai berikut:

$$\sigma_i^2 = \sigma_i^2 x_i^2 e^{ui}$$

Dimana $e = 2,718$

Uji heteroskedastisitas dapat dilakukan melalui uji White dengan meregresi terhadap residual kuadrat yang prosedurnya sebagai berikut:

H_0 : Tidak ada heteroskedastisitas

H_1 : Ada heteroskedastitas

Jika nilai signifikansi lebih besar dari derajat kepercayaan 0,05 maka dapat dikatakan tidak terjadi heteroskedastisitas. Sedangkan uji autokorelasi dan normalitas sebaiknya tidak dilakukan karena hasilnya tidak akan memberikan makna sama sekali. Hal ini disebabkan karena pada dasarnya uji normalitas hanya digunakan pada data primer dan uji autokorelasi untuk data *time series* dengan periode waktu 20 sampai 30 tahun lebih (Baltagi, 2008). Dalam penelitian ini penulis menggunakan data sekunder berbasis data panel dengan kurun waktu hanya 5 tahun.

2. Pemilihan Model

a. Common Effect

Model *common effect* merupakan pendekatan data panel yang paling sederhana. Model ini tidak memperhatikan dimensi individu maupun waktu sehingga diasumsikan bahwa perilaku antar individu sama dalam berbagai kurun waktu. Model ini hanya mengkombinasikan data *cross section* serta *time series*. Metode ini dapat menggunakan pendekatan *Ordinary Least Square* (OLS) atau teknik kuadrat terkecil untuk mengestimasi model pada data panel. Berikut persamaan regresi model common effect:

$$Y_{it} = \alpha + X_{it}\beta + \varepsilon_{it}$$

Dimana:

$i = 25$ kab/kota di provinsi Sumatera Utara

$t = 2011$ hingga 2015

proses estimasi dilakukan secara terpisah setiap *cross unit section* yang dapat dilakukan dengan asumsi komponen error pada kuadrat terkecil.

b. Fixed Effect

Model ini mengasumsikan bahwa terdapat efek yang berbeda antarindividu. Perbedaan itu dapat diakomodasi melalui perbedaan pada intersepnya. Dalam model *fixed effect* setiap parameter merupakan parameter yang tidak diketahui dan akan di estimasi dengan menggunakan teknik variabel *dummy* yang dinamakan *Least Square Dummy Variable (LSDV)*, yang mampu mengakomodasikan efek waktu yang sistematis. Hal ini dilakukan melalui penambahan variabel *dummy* di dalam model.

c. Random Effect Model

Model ini menjelaskan efek spesifik dari setiap individu sebagai bagian dari komponen error yang bersifat acak dan tidak berkorelasi dengan variabel penjelas yang teramati. Model ini disebut dengan *Error Component Model* (ECM). Persamaan dalam model ECM ini dapat dituliskan sebagai berikut:

$$Y_{it} = \alpha + X_{1it}\beta + w_{it}$$

$i = 25$ kab/kota di provinsi Sumatera Utara

$t = 2011$ hingga 2015

Dimana :

$$W_{it} = \varepsilon_{it} + u_i; E(W_{it}) = 0; E(W_{it}^2) = \alpha^2 + a_u^2;$$

$$E(W_{it}, W_{it-1}) = 0; i \neq j; E(u_i, \varepsilon_{it}) = 0;$$

$$E(\varepsilon_i, \varepsilon_{is}) = E(\varepsilon_i, \varepsilon_{jt}) = E(\varepsilon_i, \varepsilon_{js}) = 0$$

Meskipun komponen error bersifat homoskedastisitas, nyatanya terdapat

korelasi antara W_{it} dan $W_{i,t-1}$ yakni:

$$cross(W_{it}, W_{i,(t-1)}) = \frac{a_u^2}{a^2 + a_u^2}$$

Oleh karena itu, metode OLS tidak bisa digunakan untuk mendapatkan estimator yang efisien bagi model *random effects*. Metode yang tepat untuk mengestimasi model *random effects* adalah *Generalized Least Square* (GLS) dengan asumsi homoskedastik dan tidak ada korelasi *cross sectional*.

d. Uji Chow

Uji chow digunakan untuk menentukan apakah model *Common Effect* atau *Fixed Effect* yang paling tepat digunakan dalam estimasi data panel. Hipotesis dalam uji Chow adalah sebagai berikut:

H_0 : *Common Effect Model* atau *Pooled OLS*

H_1 : *Fixed Effect Model*

Dasar penolakan terhadap hipotesis di atas ialah dengan membandingkan perhitungan F-statistik dengan F-tabel. Perbandingan digunakan apabila hasil F-statistik lebih kecil dibanding F-tabel, sehingga H_0 tidak ditolak yang berarti model yang digunakan ialah *Common Effect Model* Menurut Baltagi (2008), perhitungan F-statistik didapatkan melalui Uji Chow dengan rumus sebagai berikut:

$$F = \frac{\frac{(SE_1 - SSE_2)}{(n - 1)}}{\frac{SSE_2}{(nt - n - k)}}$$

Dimana :

SSE_1 : *Sum Square Error* dari model *Common Effect*

SSE_2 : *Sum Square Error* dari model *Fixed Effect*

n : Jumlah n (*cross section*)

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

k : Jumlah variabel independen

Sedangkan F-tabel didapat dari:

$$F_{tabel} = \{a: df(n - 1), nt - n - k\}$$

Dimana:

A : Tingkat signifikansi yang dipakai

n : Jumlah unit *cross section*

nt : Jumlah *cross section* dikali *time series*

k : Jumlah variabel independen

3. Uji Signifikansi

Uji signifikansi dilakukan guna melihat apakah hipotesis akan ditolak atau tidak. Terdapat tiga cara dalam uji signifikansi, yakni

a. Uji t

Uji t atau biasa dikenal dengan uji parsial digunakan untuk menguji bagaimana pengaruh masing-masing variabel independen

dengan variabel dependen. Uji ini dapat dilakukan dengan membandingkan t-hitung dengan t-tabel atau dengan melihat signifikansi masing-masing t-hitung.

b. Uji F

Uji F dilakukan guna melihat pengaruh variabel independen secara bersama-sama mempengaruhi variabel dependen dengan membandingkan F-hitung dengan F-tabel. Apabila F-hitung lebih dari F-tabel, maka H_0 ditolak.

c. Koefisien Determinasi

Koefisien determinasi dilakukan untuk mengukur seberapa jauh kemampuan sebuah model menjelaskan variasi dalam variabel dependen. Nilai koefisien determinasi berkisar antara nol dan satu. Apabila nilai R^2 nya mendekati nol berarti variasi variabel dependennya sangat terbatas. Apabila nilainya mendekati satu berarti variabel independennya dapat menjelaskan segala informasi dari variabel dependen.