

III. METODE PENELITIAN

A. Ruang Lingkup Penelitian

Ruang lingkup penelitian ini bertujuan untuk menganalisis pengaruh pertumbuhan ekonomi, inflasi dan kualitas sumber daya manusia terhadap tingkat pengangguran di Sumatera Tahun 2009 hingga Tahun 2013.

B. Populasi dan Sampel Penelitian

Populasi adalah wilayah generalisasi yang terdiri atas, obyek/subyek mempunyai kuantitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya (Sugiono, 2009). Penelitian ini menggunakan populasi Pemerintah Provinsi-provinsi yang ada di Sumatera. Jumlah Pemerintahan Provinsi yang ada di Sumatera adalah sebanyak 10 pemerintahan. Populasi penelitian ini adalah Provinsi Nanggroe Aceh Darrusalam, Sumatera Utara, Sumatera Barat, Riau, Kepulauan Riau, Sumatera Selatan, Jambi, Bengkulu, Bangka Belitung, dan Lampung.

Sampel adalah sebagian yang diambil dari keseluruhan objek yang diteliti dan dianggap mewakili seluruh populasi (Sugiono, 2009).

Dalam penelitian ini daerah yang menjadi sampel dipilih berdasarkan *Purposive Sampling* (kriteria yang dikehendaki). Kriteria sampel dalam penelitian ini adalah sebagai berikut:

1. Provinsi di Sumatera yang masa pemerintahannya lebih dari 10 tahun.
2. Pemerintah Provinsi di Sumatera yang telah menyusun laporan keuangan, kependudukan, serta ketenagakerjaan Tahun 2009 sampai dengan Tahun 2013.
3. Pemerintah di Provinsi Sumatera yang mempunyai daftar laporan ketenagakerjaan, Laporan Hasil Pemeriksaan (LHP) atas laporan keuangan pemerintah daerah, dan kependudukan Tahun 2009 sampai dengan Tahun 2013 telah dipublikasikan melalui *website* resmi BPS.

Berdasarkan pertimbangan tersebut maka ukuran sampel pada penelitian ini yaitu sebanyak 10 Provinsi yaitu:

- | | |
|--------------------------------------|-----------------------------|
| 1. Provinsi Nanggroe Aceh Darussalam | 6. Provinsi Jambi |
| 2. Provinsi Sumatera Utara | 7. Provinsi Bengkulu |
| 3. Provinsi Sumatera Barat | 8. Provinsi Lampung |
| 4. Provinsi Riau | 9. Provinsi Bangka Belitung |
| 5. Provinsi Sumatera Selatan | 10. Provinsi Kepulauan Riau |

C. Jenis Penelitian dan Sumber Data

Penelitian ini menggunakan data sekunder bersifat runtun waktu (*time series*) dalam periode tahunan dan data antar ruang (*cross section*). Penelitian ini dilakukan dengan mengambil data sekunder dari berbagai sumber, seperti Badan Pusat Statistik (BPS), maupun sumber data sekunder lainnya. Keseluruhan data berupa data panel Tahun 2009 hingga Tahun 2013. Data panel merupakan sekelompok data individual yang diteliti selama rentang waktu tertentu sehingga memberikan informasi observasi setiap subjek dalam sampel.

D. Definisi Operasional Variabel

Untuk memberikan kejelasan mengenai penggunaan beberapa variabel dalam penelitian ini, maka perlu diberikan definisi beberapa variabel tersebut, yaitu:

1. Variabel dependen yang digunakan dalam penelitian ini adalah jumlah tingkat pengangguran terbuka, yaitu jumlah penduduk yang menganggur, yang termasuk angkatan kerja namun tidak melakukan pekerjaan atau sedang mencari kerja. Variabel jumlah tingkat pengangguran yang digunakan adalah jumlah tingkat pengangguran di Sumatera Tahun 2009 – 2013 dengan menggunakan data dari BPS.
2. Variabel independen yang digunakan dalam penelitian ini:
 - a) Pertumbuhan Ekonomi menjelaskan mengenai faktor-faktor apa saja yang menentukan kenaikan output perkapita dalam jangka panjang dan

bagaimana faktor tersebut berinteraksi satu sama lain sehingga terjadilah proses pertumbuhan ekonomi. Indikator yang digunakan untuk mengukur pertumbuhan ekonomi adalah tingkat pertumbuhan PDB. Data Pertumbuhan Ekonomi diperoleh dari BPS Provinsi-provinsi yang ada di Sumatera selama periode Tahun 2009 – 2013.

- b) Inflasi, indikator yang sering digunakan untuk mengukur tingkat inflasi adalah indeks harga konsumen (IHK). Data inflasi didapat dari BPS Provinsi-provinsi yang ada di Sumatera selama periode Tahun 2009 – 2013.
- c) Kualitas Sumber Daya Manusia (KSDM) indikator yang sering digunakan untuk mengukur Kualitas Sumber Daya Manusia adalah Indeks Pembangunan Manusia (IPM). Indeks Pembangunan Manusia (IPM) digunakan sebagai indikator pembangunan manusia. Data IPM diperoleh dari BPS Provinsi-provinsi yang ada di Sumatera selama periode Tahun 2009 – 2013.

Tabel 3.1 Nama Variabel, Simbol, Satuan Pengukuran dan Sumber Data

Nama Variabel	Simbol	Satuan Pengukuran	Sumber Data
Pengangguran	P	Persen	BPS
Pertumbuhan Ekonomi	PE	Persen	BPS
Inflasi	INF	Persen	BPS
Kualitas Sumber Daya Manusia (IPM)	KSDM	Indeks	BPS

E. Spesifikasi Model Penelitian

a. Metode Analisis Data

Metode analisis yang digunakan dalam penelitian ini adalah menggunakan analisis data panel. Ada 3 teknik pendekatan mendasar yang digunakan dalam mengestimasi model regresi dengan data panel, yaitu:

a) Model *Pooled Least Square*

Metode pendekatan ini tidak memperhatikan dimensi individu maupun waktu. Diasumsikan bahwa perilaku data antar daerah sama dalam berbagai kurun waktu (Widarjono, 2009). Model ini hanya menggabungkan kedua data tersebut tanpa melihat perbedaan antar waktu dan individu sehingga dapat dikatakan bahwa model ini sama halnya dengan metode OLS (*Ordinary Least Square*) karena menggunakan kuadrat kecil biasa. Pada beberapa penelitian data panel, model ini seringkali tidak pernah digunakan sebagai estimasi utama karena sifat dari model ini yang tidak membedakan perilaku data sehingga memungkinkan terjadinya bias, namun model ini digunakan sebagai pembanding dari kedua pemilihan model lainnya.

Analisis data menggunakan model regresi berganda yang digunakan untuk mengetahui pengaruh Pertumbuhan Ekonomi (PE), Inflasi (INF) dan Kualitas Sumber Daya Manusia (KSDM) terhadap Tingkat Pengangguran di Sumatera dengan menggunakan metode OLS (*Ordinary Least Square*).

Model umum dari analisis ini adalah:

$$P_{it} = \beta_{0i} + \beta_1 PE_{it} + \beta_2 INF_{it} + \beta_3 KSDM_{it} + \varepsilon_{it}. \quad (3.1)$$

Keterangan:

P_{it}	= Tingkat Pengangguran Provinsi, diukur dalam persen
PE_{it}	= Pertumbuhan Ekonomi Provinsi, dinyatakan dalam persen
INF_{it}	= Inflasi, dinyatakan dalam persen
$KSDM_{it}$	= Kualitas Sumber Daya Manusia, diproyeksi dengan indeks pembangunan manusia
β_0	= Intersep
β_i	= Koefisien variabel bebas
i	= Provinsi-provinsi di Sumatera
t	= Data time series
ε_{it}	= Variabel gangguan

b) Model Pendekatan Efek Tetap

Kesulitan terbesar dalam pendekatan metode kuadrat terkecil biasa adalah adanya asumsi intersep dan slope dari persamaan regresi yang dianggap konstan, baik antar daerah maupun antar waktu yang mungkin tidak beralasan. Generalisasi secara umum sering dilakukan dengan memasukkan variabel boneka (*dummy variabel*) untuk memungkinkan terjadinya perbedaan nilai parameter yang berbeda-beda baik lintas unit *cross section* maupun antar waktu. Pendekatan dengan memasukkan variabel boneka ini dikenal dengan sebutan model efek tetap (*fixed effect*) atau *Least Square Dummy Variabel* atau disebut juga *Covariance Model*.

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

$$P_{it} = \alpha_0 + \alpha_1 D_{1i} + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 D_{4i} + \alpha_5 D_{5i} + \alpha_6 D_{6i} + \alpha_7 D_{7i} + \alpha_8 D_{8i} + \alpha_9 D_{9i} + \beta_1 \ln PE_{it} + \beta_2 \ln INF_{it} + \beta_3 \ln KSDM_{it} + \varepsilon_{it} \quad (3.2)$$

Variabel di $D_{1i} = 1$ untuk objek pertama dan 0 untuk objek lainnya. Variabel $D_{2i} = 1$ untuk objek kedua dan 0 untuk objek lainnya. Variabel semu $D_{3i} = 1$ untuk objek ketiga dan 0 untuk objek lainnya dan seterusnya. Dengan menggunakan pendekatan ini, akan terjadi *degree of freedom* sebesar $NT - N - K$. Keputusan memasukkan variabel boneka ini harus didasarkan pada pertimbangan statistik. Hal tersebut disebabkan, dengan melakukan penambahan variabel boneka akan dapat mengurangi jumlah *degree of freedom* yang pada akhirnya akan mempengaruhi koefisien dari parameter yang diestimasi.

c) Model Pendekatan Efek Acak

Walaupun FEM atau LSDV mudah untuk diaplikasikan, tidak dapat dipungkiri penerapannya akan menimbulkan konsekuensi (*trade off*) yang mungkin cukup mahal. Penambahan *dummy variables* ke dalam model dapat mengurangi banyaknya derajat kebebasan (*degree of freedom*) yang pada akhirnya akan mengurangi efisiensi dari parameter yang diestimasi.

Menurut Gujarati (2010), jika *dummy variables* adalah untuk merepresentasikan ketidaktahuan tentang model yang sebenarnya, maka kita dapat menggunakan *disturbance term* untuk merepresentasikan ketidaktahuan tentang model yang

sebenarnya. Hal ini dikenal sebagai model efek acak (*random effect model* atau *REM*). Ide dasar *Random Effect Model* (REM) dalam metode ini dapat dimulai dari persamaan:

$$P_{it} = \beta_{oi} + \beta_1 P_{Eit} + \beta_2 INF_{it} + \beta_3 KSDM_{it} + u_{it}$$

Disini β_{oi} tidak diperlakukan sebagai *fixed*, kita mengasumsikan bahwa konstanta adalah variabel acak dengan nilai rata-rata β_0 . Dan nilai intersep untuk masing-masing unit *cross-section* dapat dituliskan sebagai:

$$\beta_{oi} = \beta_i + \varepsilon_i = 1, 2, \dots, N$$

dimana ε_i adalah *random error term* dengan nilai rata-rata adalah nol dan variasi adalah $\sigma^2\varepsilon$. Secara esensial, kita ingin mengatakan bahwa semua individu yang masuk ke dalam sampel diambil dari populasi yang lebih besar dan mereka memiliki nilai rata-rata yang sama untuk *intercept* (β_i) dan perbedaan individual dalam nilai *intercept* setiap individu akan direfleksikan dalam *error term* (u_i). Dengan demikian persamaan REM awal dapat dituliskan kembali menjadi:

$$P_{it} = \beta_0 + \beta_1 P_{Eit} + \beta_2 INF_{it} + \beta_3 KSDM_{it} + \varepsilon_i + u_{it}$$

$$P_{it} = \beta_0 + \beta_1 P_{Eit} + \beta_2 INF_{it} + \beta_3 KSDM_{it} + w_{it} \quad (3.3)$$

Dimana :

$$w_{it} = \varepsilon_i + u_{it}$$

Error term kini adalah w_{it} yang terdiri dari ε_i dan u_{it} . ε_i adalah *cross-section (random) error component*, sedangkan u_{it} adalah *combined error component*. Untuk alasan inilah, REM sering juga disebut *error components model (ECM)*.

Beberapa pertimbangan yang dapat dijadikan acuan untuk memilih antara *fixed effect* atau *random effect* adalah (Gujarati, 2010):

1. Bila T (banyaknya unit *time series*) besar sedangkan N (jumlah unit *cross section*) kecil, maka hasil *fixed effect* dan *random effect* tidak jauh berbeda, sehingga dapat dipilih pendekatan yang lebih mudah untuk dihitung, yaitu *fixed effect model*.
2. Bila N besar dan T kecil, maka hasil estimasi kedua pendekatan akan berbeda jauh. Apabila diyakini bahwa unit *cross section* yang dipilih dalam penelitian diambil secara acak, maka *random effect* harus digunakan. Sebaliknya apabila diyakini bahwa unit *cross section* yang dipilih dalam penelitian tidak diambil secara acak, maka harus menggunakan *fixed effect*.
3. Apabila komponen *error individual* (ε_i) berkorelasi dengan variabel bebas X, maka parameter yang diperoleh dengan *random effect* akan bias sementara parameter yang diperoleh dengan *fixed effect* tidak bias.
4. Apabila N besar dan T kecil, kemudian apabila asumsi yang mendasari *random effect* dapat terpenuhi, maka *random effect* lebih efisien dibandingkan *fixed effect*.

F. Uji Kesesuaian Model

Untuk memilih model mana yang paling tepat digunakan untuk pengolahan data panel, maka terdapat beberapa pengujian yang dapat dilakukan, antara lain:

a) Chow Test adalah pengujian untuk memilih apakah model yang digunakan *Pooled Least Square Model* atau *Fixed Effect Model*. Dalam pengujian ini dilakukan dengan hipotesis sebagai berikut:

- H0: *Pooled Least Square Model*
- H1: *Fixed Effect Model*

Dasar penolakan terhadap hipotesis nol tersebut adalah dengan menggunakan F Statistic seperti yang dirumuskan oleh Chow:

$$\text{Chow} = \frac{(\text{RRSS} - \text{URSS}) / (N - 1)}{\text{URSS} / (\text{NT} - N - K)} \sim F_{\alpha(N-1, \text{NT} - N - K)}$$

Dimana pengujian ini mengikuti distribusi F yaitu F K (N – 1, NT – N – K). Jika nilai CHOW Statistics (F Statistic) hasil pengujian lebih besar dari F Tabel, maka melakukan penolakan terhadap H0, begitu juga sebaliknya.

b) Hausman Test adalah pengujian statistik sebagai dasar pertimbangan dalam memilih apakah menggunakan *Fixed Effect Model* atau *Random Effect Model*. Pengujian ini dilakukan dengan hipotesis sebagai berikut:

- H0: *Random Effects Model*
- H1: *Fixed Effects Model*

$$H = (\beta_{REM} - \beta_{FEM})' (M_{FEM} - M_{REM})^{-1} (\beta_{REM} - \beta_{FEM}) \sim \chi^2 (k).$$

Sebagai dasar penolakan H_0 maka digunakan statistik Hausman dan membandingkannya dengan Chi square. Statistik Hausman dirumuskan dengan: Jika nilai H hasil pengujian lebih besar dari $A_2(k)$, maka cukup bukti untuk melakukan penolakan terhadap H_0 sehingga model yang digunakan adalah *Fixed Effect Model*, begitu juga sebaliknya.

G. Evaluasi Model

Untuk menghasilkan model yang efisien dan konsisten, perlu evaluasi berdasarkan kriteria ekonomi apakah hasil estimasi terhadap model regresi tidak terjadi masalah penyimpangan asumsi klasik heteroskedastisitas, multikolinearitas, dan autokorelasi.

1. Heterokedastisitas

Dalam regresi linear berganda, salah satu asumsi yang harus dipenuhi agar taksiran parameter dalam model tersebut BLUE adalah $\text{var}(u_i) = \sigma^2$ (konstan), semua *error* mempunyai variasi yang sama. Pada umumnya, heteroskedastisitas diperoleh pada data *cross section*. Jika pada model dijumpai heteroskedastisitas, maka akan membuat varians residual dari variabel tidak konstan, sehingga menyebabkan model menjadi tidak efisien meskipun tidak bias dan konsisten atau jika regresi tetap dilakukan meskipun ada masalah heteroskedastisitas, maka hasil regresi akan menjadi *misleading* (Gujarati, 2010).

Untuk menguji adanya pelanggaran asumsi heteroskedastisitas, digunakan uji *white heteroskedasticity* yang diperoleh dalam program E-views. Uji *white heteroskedasticity* dilakukan dengan membandingkan $\text{Obs} \times \text{R-Square}$ dengan χ^2 (*Chi-Square*) tabel. Jika nilai $\text{Obs} \times \text{R-Square}$ lebih kecil dari χ^2 tabel, maka tidak ada heteroskedastisitas pada model. Dalam pengolahan data panel dengan E-views 6, dapat digunakan metode *General Least Square (cross section weight)*, dan untuk mendeteksi heteroskedastisitas dilakukan dengan cara membandingkan *Sum Square Resid* pada *weighted statistics* dengan *Sum Square Resid unweighted statistics*. Jika *Sum Square Resid* pada *weighted statistics* lebih kecil dari *Sum Square Resid unweighted statistics*, maka terjadi heteroskedastisitas. Perlakuan untuk pelanggaran tersebut adalah dengan mengestimasi GLS menggunakan *White Heteroskedasticity*.

2. Multikolinieritas

Multikolinieritas adalah hubungan linear yang kuat antara variabel-variabel bebas dalam persamaan regresi berganda. Gejala multikolinieritas ini dapat dideteksi dari nilai R^2 tinggi tetapi tidak terdapat atau sedikit sekali koefisien dugaan yang berpengaruh nyata dan tanda koefisien regresi tidak sesuai dengan teori (Gujarati, 2010). Multikolinieritas dalam *pooled* data dapat di atasi dengan pemberian pembobotan (*cross section weight*) atau GLS, sehingga parameter dugaan pada taraf uji tertentu (t-statistik maupun F-hitung) menjadi signifikan.

Selain cara tadi, terdapat cara lain untuk mendeteksi gangguan multikolinieritas. Menurut Widarjono (2009), untuk mendeteksi multikolinieritas dalam sebuah model regresi berganda dapat menggunakan VIF (*Variance Inflation Factor*). Nilai VIF dapat dicari dengan menggunakan formula:

$$VIF = \frac{1}{1 - R_j^2}$$

Dimana R_j^2 diperoleh dari regresi auxiliary antara variabel bebas dengan variabel bebas sisanya (k-1). Dengan ketentuan sebagai berikut:

- $0 < VIF \leq 10$, tidak terdapat multikolinieritas
- $VIF > 10$, terdapat multikolinieritas

3. Autokorelasi

Suatu model dikatakan memiliki autokorelasi jika *error* dari periode waktu (*time series*) yang berbeda saling berkorelasi. Masalah autokorelasi ini akan menyebabkan model menjadi tidak efisien meskipun masih tidak bias dan konsisten (Agus, 2009). Autokorelasi menyebabkan estimasi standar *error* dan varian koefisien regresi yang diperoleh akan *underestimate*, sehingga R^2 akan besar serta uji-t dan uji-F menjadi tidak valid. Autokorelasi yang kuat dapat menyebabkan dua variabel yang tidak berhubungan menjadi berhubungan. Bila OLS digunakan, maka akan terlihat koefisien signifikansi dan R^2 yang besar atau juga disebut sebagai regresi lancung atau palsu.

Tabel 3.2. Uji Statistik Durbin Watson

Nilai statistik d	Hasil
$0 < d < d_L$	Menolak H_0 ; ada autokorelasi positif
$d_L \leq d \leq d_U$	Daerah keragu-raguan; tidak ada keputusan
$d_U \leq d \leq 4 - d_U$	Menerima H_0 ; tidak ada autokorelasi
$4 - d_U \leq d \leq 4 - d_L$	Daerah keragu-raguan; tidak ada keputusan
$4 - d_L \leq d \leq 4$	Menolak H_0 ; ada autokorelasi negative

Sumber: Ekonometrika Pengantar dan Aplikasinya, Agus Widarjono 2009

Untuk mendeteksi ada tidaknya autokorelasi dapat dilakukan uji Durbin

Watson (DW), yaitu dengan membandingkan nilai Durbin Watson dari model dengan DW-tabel.

Walaupun uji otokorelasi DW mudah dilakukan, namun uji ini mengandung beberapa kelemahan yaitu uji DW hanya berlaku jika variabel independen bersifat random atau stokastik. Kedua, uji DW hanya berlaku jika hubungan otokorelasi antar residual dalam order pertama atau autoregresif order pertama disingkat AR (1). Ketiga, model ini tidak dapat digunakan dalam kasus rata-rata bergerak dari residual yang lebih tinggi.

Berdasarkan kelemahan diatas, maka Breusch dan Godfrey mengembangkan uji autokorelasi yang lebih umum dan dikenal dengan uji LM atau LM-Test. Jika nilai Chi-Squares hitung lebih kecil dari nilai kritisnya, maka kita menerima hipotesis nol. Artinya model tidak mengandung unsur autokorelasi karena semua nilai ρ sama dengan nol. Dan Jika nilai probabilitas lebih besar dari nilai α yang dipilih maka kita menerima H_0 yang berarti tidak ada autokorelasi.

H. Pengujian Hipotesis

1. Uji Parsial (Uji t-statistik)

Uji ini digunakan untuk melihat signifikansi dari pengaruh variabel independen terhadap variabel dependen secara individual (Gujarati, 2010). Digunakan uji 1 arah dengan tingkat kepercayaan 95% dengan hipotesis.

- $H_0 : \beta_1 = 0$ PE tidak berpengaruh terhadap Tingkat Pengangguran di Sumatera.
 $H_a : \beta_1 < 0$ PE berpengaruh terhadap Tingkat Pengangguran di Sumatera.
- $H_0 : \beta_2 = 0$ Inflasi tidak berpengaruh terhadap Tingkat Pengangguran di Sumatera.
 $H_a : \beta_2 < 0$ Inflasi berpengaruh terhadap Tingkat Pengangguran di Sumatera.
- $H_0 : \beta_3 = 0$ Kualitas Sumber Daya Manusia tidak berpengaruh terhadap Tingkat Pengangguran di Sumatera.
 $H_a : \beta_3 < 0$ Kualitas Sumber Daya Manusia berpengaruh terhadap Tingkat Pengangguran di Sumatera.
- Jika nilai t-hitung > nilai t-tabel maka H_0 ditolak atau menerima H_a
 - Jika nilai t-hitung < nilai t-tabel maka H_0 diterima atau menolak H_a

2. Uji F-statistik

Uji statistik F pada dasarnya menunjukkan apakah semua variabel independen atau bebas yang dimasukkan dalam model mempunyai pengaruh secara bersama-sama terhadap variabel dependen/terikat (Gujarati, 2010). Berikut ini adalah langkah-langkah dalam uji-F statistik pada tingkat kepercayaan 95% dengan derajat kebebasan $df\ 1 = (k-1)$ dan $df\ 2 = (n-k)$:

$H_0 : \beta_1, \beta_2, \beta_3 = 0$ (Semua variabel *independen* tidak mampu mempengaruhi variabel *dependent* secara bersama-sama).

$H_a : \beta_1, \beta_2, \beta_3 \neq 0$ (Paling tidak salah satu koefisien regresi tidak sama dengan nol maka variabel *independen* mampu mempengaruhi variabel *dependent* secara bersama-sama).

Untuk menguji hipotesis ini digunakan F-statistik dengan kriteria pengambilan keputusan membandingkan nilai F-hitung dengan nilai F-tabel.

- Jika F-hitung $>$ F-tabel, maka H_0 ditolak
- Jika F-hitung $<$ F-tabel, maka H_0 diterima

I. Penafsiran Koefisien Determinasi

Koefisien determinasi berfungsi untuk menunjukkan seberapa baik model yang diperoleh sesuai dengan data actual (*goodness of fit*), mengukur berapa persentase variasi dalam peubah terikat mampu dijelaskan oleh informasi peubah bebas. Kisaran nilai koefisien determinasi adalah $0 \leq R^2 \leq 1$. Model dikatakan semakin baik apabila nilai R^2 mendekati 1 atau 100 %.