

BAB III

METODE PENELITIAN

A. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain studi panel data yang menggabungkan analisis *cross-sectional* dan *time series*. Pendekatan ini dipilih karena mampu menangkap variasi antar provinsi sekaligus perkembangan variabel penelitian dari waktu ke waktu. Data yang digunakan mencakup 34 provinsi di Indonesia selama periode 2019-2023. Penelitian ini mengamplifikasi dua teknik analisis utama, yaitu *Ordinary Least Squares* (OLS) untuk mengestimasi hubungan dasar antara variabel dependen dan independen, serta *Two-Stage Least Squares* (2SLS) untuk mengatasi potensi masalah endogenitas. Pemilihan metode 2SLS didasarkan pada kemungkinan adanya hubungan timbal balik antara tingkat kriminalitas dan jumlah wisatawan.

B. Jenis dan Sumber Data

Penelitian ini menggunakan data sekunder yang diperoleh dari berbagai sumber resmi dan terpercaya. Data pariwisata, termasuk jumlah wisatawan mancanegara serta kapasitas fasilitas akomodasi (jumlah kamar dan tempat tidur pada hotel berbintang dan non-bintang), diperoleh dari publikasi Statistik Pariwisata yang diterbitkan oleh Badan Pusat Statistik (BPS). Data kriminalitas, yang diukur berdasarkan jumlah total kejahatan per 100.000 penduduk, diambil dari laporan resmi yang dipublikasikan oleh Kepolisian Negara Republik Indonesia (POLRI).

Sementara itu, variabel ekonomi seperti Produk Domestik Bruto (PDB) per kapita provinsi dan nilai tukar rupiah terhadap dolar Amerika Serikat diperoleh dari publikasi BPS dan Bank Indonesia. Adapun data PDB per kapita dunia sebagai indikator ekonomi global diambil dari basis data World Bank. Penelitian ini juga menggunakan variabel dummy untuk mengontrol efek karakteristik wilayah tertentu, seperti letak geografis (provinsi berada di Jawa-Bali) dan tingkat urbanisasi (rasio penduduk perkotaan lebih dari 50%).

Penelitian memiliki keterbatasan data, khususnya dalam hal pengukuran jumlah wisatawan. Variabel dependen yang digunakan adalah jumlah wisatawan mancanegara yang menginap di hotel berbintang dan non-bintang, yang hanya merupakan proksi dari jumlah total wisatawan yang masuk ke provinsi. Ini menimbulkan dua keterbatasan utama:

- a. Tidak mencakup wisatawan non-hotel, seperti mereka yang menginap di rumah keluarga, *homestay*, penginapan informal, atau bahkan tidak menginap
- b. Data kejahatan terbatas pada kejahatan yang dilaporkan, sehingga tidak mencerminkan angka kejahatan laten (*unreported crime*).

C. Populasi dan Sampel

1. Populasi

Populasi adalah keseluruhan objek atau subjek penelitian yang memiliki karakteristik tertentu dan menjadi fokus dalam penelitian. Populasi mencakup semua individu, kejadian, atau objek yang relevan

dengan tujuan penelitian.⁴³ Populasi dalam penelitian ini mencakup seluruh provinsi yang ada di Indonesia selama periode penelitian.

2. Sampel

Sampel adalah bagian dari populasi yang dipilih untuk mewakili keseluruhan populasi dalam penelitian. Sampel digunakan ketika populasi terlalu besar untuk diteliti seluruhnya.⁴⁴ Sampel yang digunakan berjumlah 34 provinsi dengan data lengkap selama enam tahun pengamatan (2019-2023). Pemilihan periode ini didasarkan pada fenomena pandemik yang melanda dunia dan ketersediaan data yang konsisten dari berbagai sumber resmi. Kriteria inklusi sampel meliputi ketersediaan data kriminalitas dan pariwisata yang tercatat secara resmi di setiap provinsi. Penggunaan data panel memungkinkan penelitian ini untuk mengontrol heterogenitas antar provinsi yang tidak terobservasi sekaligus menganalisis dinamika perubahan dari tahun ke tahun.

D. Variabel Penelitian

Penelitian ini menggunakan tiga kelompok variabel utama, yaitu variabel dependen, variabel independen, dan variabel kontrol.

1. Variabel Dependen

Tingkat kejahatan di tingkat provinsi, yang diukur berdasarkan total kejadian kejahatan per 100.000 penduduk. Data kejahatan ini merupakan

43 Primadi Candra Susanto et al., "Konsep Penelitian Kuantitatif: Populasi, Sampel, Dan Analisis Data (Sebuah Tinjauan Pustaka)," *Jurnal Ilmu Multidisplin* 3, no. 1 (2024): 1–12, <https://doi.org/10.38035/jim.v3i1.504>.

44 Nidia Suriani, Risnita, and M. Syahrani Jailani, "Konsep Populasi Dan Sampling Serta Pemilihan Partisipan Ditinjau Dari Penelitian Ilmiah Pendidikan," *Jurnal IHSAN : Jurnal Pendidikan Islam* 1, no. 2 (2023): 24–36, <https://doi.org/10.61104/ihsan.v1i2.55>.

agregasi dari kejahatan kekerasan dan kejahatan properti yang dilaporkan kepada pihak kepolisian. Penggunaan rasio per 100.000 penduduk dimkasudkan untuk memastikan komparabilitas antar provinsi.

2. Variabel Independen

Pariwisata yang diproksikan melalui jumlah tamu mancanegara yang menginap di hotel berbintang maupun non-bintang. Indikator ini mencerminkan dinamika permintaan sektor pariwisata di setiap provinsi. Selain itu, dari sisi penawaran, penelitian juga mengukur jumlah kamar dan tempat tidur yang tersedia di hotel berbintang dan non-bintang sebagai proksi kapasitas fasilitas akomodasi.

3. Variabel Kontrol

Untuk memastikan bahwa hubungan antara variabel independen dan dependen tidak terdistorsi oleh pengaruh luar, penelitian ini menyertakan sejumlah variabel kontrol. Variabel kontrol yang digunakan dikelompokkan ke dalam tiga kategori utama:

- a. Karakteristik wilayah mencakup letak geografis provinsi (dengan dummy Jawa-Bali), serta tingkat urbanisasi (dummy kota dengan rasio penduduk perkotaan $> 50\%$). Tujuannya adalah mengontrol efek heterogenitas spasial antar provinsi.
- b. Fasilitas Pariwisata yang terdiri atas jumlah dan kapasitas fasilitas akomodasi seperti hotel berbintang dan non-bintang, yang diringkas menggunakan *Principal Component Analysis* (PCA) untuk

menghasilkan indikator komposit yang merepresentasikan kualitas dan kapasitas fasilitas wisata.

- c. Karakteristik Ekonomi: meliputi indikator ekonomi makro seperti Produk Domestik Bruto (PDB) per kapita provinsi, nilai tukar riil rupiah terhadap dolar Amerika Serikat (sebagai proksi harga domestik), dan PDB per kapita dunia (sebagai proksi pendapatan global), yang dapat memengaruhi baik aktivitas wisata maupun tingkat kriminalitas.

4. Variabel Instrumen

Untuk mengatasi potensi endogenitas, khususnya kemungkinan hubungan dua arah antara kriminalitas dan pariwisata, penelitian ini menggunakan PDB per kapita provinsi, sebagai variabel instrumen. Variabel ini dipilih karena memenuhi syarat sebagai instrumen yang valid yakni berkorelasi dengan tingkat kriminalitas, namun tidak memengaruhi jumlah wisatawan secara langsung. Secara teoritis, PDB per kapita mencerminkan kondisi ekonomi lokal yang lebih baik, yang dapat menurunkan insentif untuk melakukan tindak kriminal, namun tidak memengaruhi preferensi wisatawan dalam memilih destinasi, sehingga memenuhi asumsi eksogenitas yang dibutuhkan dalam analisis instrumental.

Dengan pengelompokan variabel-variabel tersebut, penelitian ini diharapkan dapat mengidentifikasi dan menganalisis hubungan antara kriminalitas dan pariwisata secara komprehensif dan terkontrol, sehingga hasil yang diperoleh lebih valid dan dapat diinterpretasikan secara akurat.

E. Teknis Analisis Data

Proses pengolahan data dan analisis regresi panel dalam penelitian ini menggunakan Stata 17, perangkat lunak statistik yang banyak dipakai dalam analisis data panel. *Software* ini memungkinkan estimasi model regresi panel secara akurat dan efisien serta menghasilkan *output* yang jelas dan mudah diinterpretasikan, sehingga mendukung validitas dan reliabilitas hasil penelitian.

1. Uji Asumsi Klasik pada model Regresi Linear Biasa (OLS)

Sebelum melakukan analisis regresi metode *Ordinary Least Squares* (OLS), penelitian ini akan melakukan uji asumsi klasik untuk memastikan validitas dan reliabilitas hasil estimasi. Uji asumsi klasik yang dilakukan meliputi:

a. Uji Normalitas

Uji ini dilakukan untuk mengetahui apakah residual mengikuti distribusi normal. Dalam penelitian ini, uji normalitas dilakukan menggunakan perintah *sktest* pada Stata. Namun, karena model yang digunakan berbasis data panel, uji normalitas hanya bersifat opsional dan tidak menjadi syarat konsistensi estimasi parameter ketika jumlah observasi besar (*asymptotic normality* berlaku).

b. Uji Multikolinearitas

Uji ini dilakukan untuk mengetahui apakah terdapat korelasi yang tinggi antar variabel independen dalam model. Uji multikolinearitas

dapat dilakukan dengan melihat nilai *Variance Inflation Factor* (VIF).

Jika nilai $VIF < 10$, maka tidak terdapat multikolinearitas yang serius.

c. Uji Heteroskedastisitas

Untuk menguji apakah terdapat ketidaksamaan varians error antar observasi, digunakan uji *Breusch Pagan (estat hettest)* atau alternatif White test. Jika ditemukan indikasi heteroskedastisitas, maka model akan diestimasi ulang menggunakan *robust standard errors*

d. Uji Autokorelasi

Karena penelitian ini menggunakan data panel, maka uji autokorelasi dilakukan menggunakan uji *Wooldridge* untuk panel data dengan perintah *xtserial*. Uji ini mendeteksi adanya korelasi serial dalam residual di tiap unit observasi (provinsi) sepanjang waktu. Jika autokorelasi ditemukan, pendekatan *standard error* yang sesuai (seperti *cluster robust standard error*) akan diterapkan untuk menjaga validitas inferensi.

Seluruh pengujian ini bertujuan untuk memastikan bahwa model OLS dan 2SLS yang digunakan tidak mengalami bias estimasi atau masalah inferensi akibat pelanggaran asumsi klasik, khususnya dalam konteks panel data.

2. Analisis Regresi OLS

Model dasar yang digunakan adalah regresi OLS untuk mengestimasi hubungan antara tingkat kriminalitas (variabel independen) dengan jumlah wisatawan (variabel dependen). Model ini dirumuskan sebagai berikut:

$$\ln Tourist_{it} = \beta_0 + \beta_1 \ln Crime_{it} + \beta_2 Pcatot_{it} + \beta_3 \ln NilaiTr_{it} + \beta_4 \ln GDPw_t + \beta_5 DumUrban_{it} + \beta_6 DumJawaBali_i + \eta_i + \varepsilon_{it}$$

- $\ln Tourist_{it}$ = Jumlah kunjungan wisatawan mancanegara di provinsi ke- i pada tahun ke- t
- β_0 = Konstanta/intersep, yaitu nilai rata-rata jumlah wisatawan ketika semua variabel independen bernilai nol.
- $\ln Crime_{it}$ = Tingkat kriminalitas di provinsi ke- i pada tahun ke- t
- $Pcatot_{it}$ = Indikator fasilitas pariwisata di provinsi ke- i pada tahun ke- t
- $\ln NilaiTr_{it}$ = Nilai tukar Rupiah terhadap Dollar Amerika pada tahun ke- t
- $\ln GDPw$ = PDB dunia Perkapita pada tahun ke- t
- $DumUrban_{it}$ = Tingkat Urbanisasi dari suatu provinsi ($>50\%=1$ dan $<50\%=0$)
- $DumJawaBali_{it}$ = Wilayah yang berada di pulau Jawa dan Bali (Jawa dan Bali=1 dan luar Jawa Bali=0)
- $\beta_1, \beta_2, \beta_3$ = Koefisien regresi yang menunjukkan seberapa besar pengaruh masing-masing variabel dependen

- μ_i = Efek tetap (*fixed effect*) untuk provinsi ke- i yang menangkap karakteristik khusus yang tetap/tidak berubah dari masing-masing provinsi/daerah.
- ε_{it} = Error term, yaitu komponen acak yang menangkap faktor-faktor lain yang memengaruhi jumlah wisatawan

Mengingat struktur data yang digunakan dalam penelitian ini adalah data panel (34 provinsi selama periode 2019-2023), maka diperlukan pendekatan yang mampu mengontrol heterogenitas yang tidak teramati di masing-masing provinsi. Oleh karena itu, penelitian ini menggunakan model regresi panel dengan pendekatan *Fixed Effect* (FE). Pendekatan ini memungkinkan estimasi yang lebih akurat karena menghilangkan bias akibat karakteristik tetap provinsi yang tidak diobservasi, seperti faktor sosial-budaya, tingkat keamanan umum, atau daya tarik historis suatu daerah. Asumsi dasar dari pendekatan ini adalah bahwa karakteristik provinsi dapat berkorelasi dengan variabel independen, sehingga metode *Random Effect* (RE) dianggap kurang tepat. Selain itu, pendekatan FE juga akan diterapkan dalam estimasi model instrumental (2SLS) untuk menjaga konsistensi dan validitas hasil estimasi pada data panel.

3. Analisis model *Two-Stages Least Squares* (2SLS)

Penggunaan metode dalam penelitian ini didasarkan pada pertimbangan teoretis dan metodologis. Secara konseptual, hubungan antara kriminalitas dan pariwisata bersifat timbal balik. Dari sisi

permintaan (*demand side*), meningkatnya kriminalitas menurunkan minat wisatawan untuk berkunjung karena faktor keamanan merupakan determinan penting dalam keputusan perjalanan. Sebaliknya, dari sisi penawaran (*supply side*), peningkatan jumlah wisatawan justru dapat menciptakan lebih banyak peluang kejahatan, karena wisatawan sering dipersepsikan sebagai target layak bagi pelaku kriminal. Interaksi dua arah ini menimbulkan potensi masalah endogenitas dalam model, sehingga estimasi menggunakan metode *Ordinary Least Squares* (OLS) berisiko menghasilkan koefisien yang bias dan inkonsisten. Oleh karena itu, penelitian ini menggunakan pendekatan 2SLS dengan Produk Domestik Regional Bruto (PDRB) per kapita sebagai variabel instrumen. PDRB per kapita dipilih karena secara teoritis berkorelasi dengan tingkat kriminalitas melalui dimensi kesejahteraan ekonomi, tetapi tidak memengaruhi secara langsung keputusan wisatawan mancanegara dalam memilih destinasi. Dengan demikian, 2SLS mampu mengatasi *simultaneity bias* dan memberikan estimasi hubungan kausal yang lebih sahih antara kriminalitas dan kunjungan wisatawan mancanegara.

a. Tahap 1

$$\ln Crime_{it} = \pi_0 + \pi_1 \ln PDRBcapita_{it} + \pi_2 Pcatot_{it} + \pi_3 \ln NilaiTr_t + \pi_4 \ln GDPw_t + \pi_5 DumUrban_{it} + \pi_6 DumJawaBali_i + \eta_i + v_{it}$$

- $\ln Crime_{it}$ = variabel dependen tahap pertama, yaitu logaritma natural tingkat kriminalitas pada provinsi i tahun t .

- $\pi_0 \rightarrow$ konstanta regresi.
- $\ln PDB_{it}$ = Variabel instrument, logaritma natural PDRB per kapita provinsi ke- i pada tahun ke- t
- $Pcatot_{it}$ = Indikator fasilitas pariwisata di provinsi ke- i pada tahun ke- t
- $\ln Nilait_r$ = Nilai tukar Rupiah terhadap Dollar Amerika pada tahun ke- t
- $\ln GDPw_t$ = logaritma natural PDB dunia pada tahun ke- t
- $DumUrban_{it}$ = Tingkat Urbanisasi dari suatu provinsi ($>50\%=1$ dan $<50\%=0$)
- $DumJawaBali_{it}$ = Wilayah yang berada di pulau Jawa dan Bali (Jawa dan Bali=1 dan luar Jawa Bali=0)
- η_i = Efek tetap (*fixed effect*) untuk provinsi ke- i yang menangkap karakteristik khusus yang tetap/tidak berubah dari masing-masing provinsi/daerah.
- v_{it} = Error term, yaitu komponen acak yang menangkap faktor-faktor lain yang memengaruhi jumlah wisatawan

b. Tahap 2

$$\ln Tourist_{it} = \beta_0 + \beta_1 \widehat{\ln Crime_{it}} + \beta_2 Pcatot_{it} + \beta_3 \ln NilaiTr_t + \beta_4 \ln GDPw_t + \beta_5 DumUrban_{it} + \beta_6 DumJawaBali_i + \eta_i + \varepsilon_{it}$$

- $\ln Tourist_{it}$ = variabel dependen tahap kedua, yaitu logaritma natural jumlah wisatawan mancanegara di provinsi i tahun t .

- β_0 = Konstanta/intersep, yaitu nilai rata-rata jumlah wisatawan ketika semua variabel independen bernilai nol.
- $\widehat{\ln Crime_{it}}$ = nilai prediksi kriminalitas dari tahap pertama, digunakan sebagai pengganti kriminalitas aktual untuk mengatasi masalah endogenitas.
- $Pcatot_{it}$ = Indikator fasilitas pariwisata di provinsi ke- i pada tahun ke- t
- $\ln Nilaitr_{it}$ = Nilai tukar Rupiah terhadap Dollar Amerika pada tahun ke- t
- $\ln GDP_w$ = PDB dunia Perkapita pada tahun ke- t
- $DumUrban_{it}$ = Tingkat Urbanisasi dari suatu provinsi ($>50\%=1$ dan $<50\%=0$)
- $DumJawaBali_{it}$ = Wilayah yang berada di pulau Jawa dan Bali (Jawa dan Bali=1 dan luar Jawa Bali=0)
- $\beta_1, \beta_2, \beta_3$ = Koefisien regresi yang menunjukkan seberapa besar pengaruh masing-masing variabel dependen
- μ_i = Efek tetap (*fixed effect*) untuk provinsi ke- i yang menangkap karakteristik khusus yang tetap/tidak berubah dari masing-masing provinsi/daerah.
- ε_{it} = Error term, yaitu komponen acak yang menangkap faktor-faktor lain yang memengaruhi jumlah wisatawan