



UNIVERSITAS INDONESIA

**ANALISIS REGRESI SPASIAL (SAR/SEM) PADA SEKTOR
PEKERJAAN KESEHATAN DAN SOSIAL DI INGGRIS**

LAPORAN

AMMAR HANAFI (2206051582)
NORMAN MOWLANA AZIZ (2206025470)
KIRONO DWI SAPUTRO (2106656365)
DEVANA SOLEA (2306262402)

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI SARJANA STATISTIKA
DEPOK
FEBRUARI 2026

Daftar Isi

BAB I PENDAHULUAN	4
1.1 Latar Belakang	4
1.2 Rumusan Masalah	5
1.3 Tujuan	5
BAB II LANDASAN TEORI	6
2.1 Statistika Deskriptif	6
2.1.1 Ukuran Pemusatan	6
2.1.2 Ukuran Penyebaran	6
2.2 Regresi Linear Berganda (Ordinary Least Squares)	7
2.3 Matriks Pembobot Spasial (Spatial Weights Matrix)	7
2.3.1 Rook Contiguity	7
2.3.2 Queen Contiguity	8
2.3.3 Bishop Contiguity	8
2.4 Indeks Moran (Moran's I)	8
2.5 Uji Lagrange Multiplier (LM)	9
2.6 Spatial Lag Model (SAR)	9
2.7 Spatial Error Model (SEM)	10
2.8 Pendugaan Parameter (Maximum Likelihood Estimation)	11
BAB III METODE PENELITIAN	12
3.1 Data dan Variabel	12
3.2 Tahapan Analisis	12
BAB IV ANALISIS DAN PEMBAHASAN	13
4.1 Analisis Deskriptif dan Eksplorasi Data (EDA)	13
4.1.1 Statistik Deskriptif	13
4.1.2 Visualisasi Data	13
4.1.3 Peta Sebaran	15
4.2 Hasil Regresi OLS (Global)	16
4.3 Uji Diagnostik Regresi Global	17
4.4 Diagnostik Dependensi Spasial	17
4.5 Hasil Regresi Spasial (SEM)	18
4.6 Perbandingan Model	20
4.7 Interpretasi	20
BAB V KESIMPULAN	21
DAFTAR PUSTAKA	24
LAMPIRAN	25

Daftar Tabel

1	Sampel Data (5 Baris Pertama)	12
2	Statistik Deskriptif Variabel Penelitian	13
3	Ringkasan Hasil Regresi OLS	16
4	Hasil Uji Diagnostik OLS	17
5	Variance Inflation Factor (VIF)	17
6	Hasil Uji Diagnostik Spasial	17
7	Ringkasan Hasil Spatial Error Model (SEM)	18
8	Perbandingan Kebaikan Model	20

Daftar Gambar

1	Histogram Distribusi Variabel	14
2	Boxplot Variabel (Deteksi Outlier)	14
3	Matriks Korelasi Antar Variabel	15
4	Peta Sebaran Tenaga Kerja Sektor Kesehatan (Quantile)	16
5	Peta Residual: OLS vs SAR vs SEM	18
6	Peta Residual Model SEM	19
7	Peta Klaster LISA (Local Indicators of Spatial Association)	21

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sektor kesehatan dan pekerjaan sosial merupakan salah satu sektor strategis yang memiliki peran fundamental dalam mendukung kualitas hidup masyarakat serta keberlanjutan pembangunan ekonomi suatu wilayah. Keberadaan tenaga kerja pada sektor ini tidak hanya berkontribusi terhadap peningkatan derajat kesehatan masyarakat, tetapi juga berperan dalam memperkuat sistem perlindungan sosial dan stabilitas sosial ekonomi secara keseluruhan. Dalam konteks pembangunan wilayah, distribusi tenaga kerja pada sektor *Human Health and Social Work* sering kali menunjukkan variasi yang cukup besar antar daerah. Variasi tersebut dipengaruhi oleh berbagai faktor struktural, seperti tingkat pendidikan, perkembangan sektor keuangan dan asuransi, sektor informasi dan komunikasi, serta aktivitas real estat.

Secara teoritis, wilayah dengan tingkat pendidikan yang lebih tinggi cenderung memiliki kapasitas sumber daya manusia yang lebih baik untuk mendukung sektor kesehatan dan layanan sosial. Pendidikan berperan dalam meningkatkan kompetensi tenaga kerja, memperluas akses terhadap informasi, serta mendorong inovasi dalam penyediaan layanan publik. Selain itu, sektor keuangan dan asuransi berpotensi mempengaruhi pembiayaan layanan kesehatan, sementara sektor real estat mencerminkan dinamika pembangunan wilayah yang dapat berdampak pada kebutuhan dan penyeediaan fasilitas kesehatan. Dengan demikian, hubungan antar sektor ekonomi tersebut bersifat kompleks dan saling terkait.

Namun demikian, dalam analisis data kewilayahan, terdapat karakteristik khusus yang membedakannya dari data non-spasial, yaitu adanya kemungkinan ketergantungan spasial (*spatial dependence*). Ketergantungan spasial mengacu pada kondisi di mana nilai suatu variabel pada suatu wilayah tidak berdiri sendiri, melainkan dipengaruhi oleh nilai variabel pada wilayah sekitarnya. Fenomena ini sejalan dengan Hukum Geografi Pertama Tobler yang menyatakan bahwa “segala sesuatu saling berhubungan, tetapi hal-hal yang berdekatan memiliki hubungan yang lebih kuat dibandingkan yang berjauhan”. Dalam konteks tenaga kerja sektor kesehatan, wilayah yang berdekatan secara geografis sering kali memiliki karakteristik ekonomi dan sosial yang serupa, sehingga memungkinkan terbentuknya pola pengelompokan (*clustering*).

Apabila ketergantungan spasial diabaikan, penggunaan model regresi linear klasik atau *Ordinary Least Squares* (OLS) berpotensi menghasilkan estimator yang tidak efisien dan bias dalam inferensi statistik. Salah satu asumsi penting dalam OLS adalah independensi error, yaitu tidak adanya korelasi antar residual observasi. Jika terdapat autokorelasi spasial pada residual, maka varians estimator menjadi tidak akurat sehingga uji signifikansi parameter dapat menyesatkan. Oleh karena itu, diperlukan pendekatan pemodelan yang secara eksplisit mengakomodasi struktur spasial dalam data.

Regresi spasial merupakan pengembangan dari model regresi klasik yang dirancang untuk menangani adanya interaksi atau dependensi antar wilayah. Dua model yang umum digunakan adalah *Spatial Autoregressive Model* (SAR), yang memasukkan pengaruh variabel dependen dari wilayah tetangga ke dalam model, serta *Spatial Error*

Model (SEM), yang mengakomodasi ketergantungan spasial pada komponen error. Pemilihan model yang tepat sangat penting untuk memastikan bahwa hubungan antar variabel ekonomi yang dianalisis dapat diinterpretasikan secara valid dan konsisten.

Dengan mempertimbangkan pentingnya sektor kesehatan dan pekerjaan sosial dalam pembangunan wilayah serta potensi adanya pola spasial dalam distribusinya, penelitian ini menggunakan pendekatan regresi spasial untuk mengidentifikasi faktor-faktor yang mempengaruhi variasi tenaga kerja pada sektor tersebut. Pendekatan ini diharapkan mampu memberikan gambaran yang lebih komprehensif mengenai dinamika ekonomi kewilayahan dan menjadi dasar dalam perumusan kebijakan pembangunan yang lebih tepat sasaran.

1.2 Rumusan Masalah

1. Apakah terdapat ketergantungan spasial (spatial dependence) dalam model proporsi pekerja sektor kesehatan dan sosial di Inggris?
2. Bagaimana pengaruh variabel sektor pekerjaan (Education, Finance, Info-Comm, Real Estate) terhadap tingkat pekerja di sektor Kesehatan dan Sosial (Human Health and Social Work)?
3. Model spasial manakah (SAR atau SEM) yang terbaik untuk menjelaskan data tersebut?

1.3 Tujuan

Tujuan dari analisis ini adalah untuk mengidentifikasi dan memodelkan hubungan antara variabel-variabel prediktor dengan variabel respon dengan memperhitungkan efek spasial, serta menentukan model terbaik antara OLS, SAR, dan SEM.

BAB II

LANDASAN TEORI

2.1 Statistika Deskriptif

Statistika deskriptif merupakan cabang ilmu statistika yang digunakan untuk menggambarkan karakteristik data tanpa melakukan generalisasi terhadap populasi yang lebih luas. Dalam penelitian spasial, statistika deskriptif tidak hanya menjelaskan distribusi numerik tetapi juga pola geografis antar wilayah.

2.1.1 Ukuran Pemusatan

Mean Mean atau rata-rata dihitung dengan rumus:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (1)$$

di mana x_i adalah nilai pengamatan ke- i dan n adalah jumlah observasi.

Median Median adalah nilai tengah setelah data diurutkan.

Kuartil Kuartil membagi data menjadi empat bagian:

- Q1 (25%)
- Q2 (Median)
- Q3 (75%)

2.1.2 Ukuran Penyebaran

Varians

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad (2)$$

Standar Deviasi

$$s = \sqrt{s^2} \quad (3)$$

Koefisien Variasi

$$CV = \frac{s}{\bar{x}} \times 100\% \quad (4)$$

Koefisien variasi digunakan untuk membandingkan tingkat variasi antar variabel dengan skala yang berbeda.

2.2 Regresi Linear Berganda (Ordinary Least Squares)

Regresi linear berganda digunakan untuk menganalisis pengaruh beberapa variabel independen terhadap satu variabel dependen.

Model umum regresi linear berganda adalah:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i \quad (5)$$

dengan:

- Y_i = variabel dependen
- X_{ki} = variabel independen ke- k
- β_k = parameter regresi
- ε_i = error

2.3 Matriks Pembobot Spasial (Spatial Weights Matrix)

Matriks pembobot spasial W adalah komponen kunci dalam analisis spasial yang merepresentasikan struktur hubungan antar lokasi. Dalam analisis ini, digunakan pembobot **Queen Contiguity**, di mana lokasi i dan j dianggap bertetangga jika mereka berbagi sisi atau sudut perbatasan yang sama.

Matriks pembobot spasial (W) digunakan untuk merepresentasikan hubungan kedekatan antar wilayah. Elemen matriks w_{ij} menunjukkan tingkat keterkaitan antara wilayah i dan wilayah j .

Secara umum, matriks pembobot spasial didefinisikan sebagai:

$$W = [w_{ij}] \quad (6)$$

dengan:

$$w_{ij} = \begin{cases} 1, & \text{jika wilayah } i \text{ bertetangga dengan wilayah } j \\ 0, & \text{lainnya} \end{cases} \quad (7)$$

Untuk menjaga konsistensi skala, dilakukan normalisasi baris:

$$w_{ij}^* = \frac{w_{ij}}{\sum_j w_{ij}} \quad (8)$$

Sehingga setiap baris memiliki jumlah bobot sebesar 1.

2.3.1 Rook Contiguity

Rook contiguity mendefinisikan dua wilayah sebagai bertetangga apabila keduanya memiliki batas sisi (edge) yang sama. Dengan kata lain, hubungan spasial hanya terjadi apabila terdapat perpotongan garis batas wilayah.

Secara matematis:

$$w_{ij}^{rook} = \begin{cases} 1, & \text{jika } i \text{ dan } j \text{ berbagi sisi} \\ 0, & \text{lainnya} \end{cases} \quad (9)$$

Metode ini cenderung menghasilkan jumlah tetangga yang lebih sedikit dibandingkan Queen contiguity.

2.3.2 Queen Contiguity

Queen contiguity mendefinisikan dua wilayah sebagai bertetangga apabila keduanya memiliki batas sisi atau titik sudut (vertex) yang sama. Konsep ini menyerupai pergerakan bidak ratu dalam permainan catur.

$$w_{ij}^{queen} = \begin{cases} 1, & \text{jika } i \text{ dan } j \text{ berbagi sisi atau titik sudut} \\ 0, & \text{lainnya} \end{cases} \quad (10)$$

Queen contiguity biasanya menghasilkan matriks dengan koneksiitas lebih tinggi dibandingkan rook.

2.3.3 Bishop Contiguity

Bishop contiguity mendefinisikan dua wilayah sebagai bertetangga hanya apabila keduanya berbagi titik sudut (vertex) tetapi tidak berbagi sisi. Konsep ini menyerupai pergerakan bidak gajah dalam permainan catur.

$$w_{ij}^{bishop} = \begin{cases} 1, & \text{jika } i \text{ dan } j \text{ berbagi titik sudut saja} \\ 0, & \text{lainnya} \end{cases} \quad (11)$$

Model bishop menghasilkan struktur ketetanggaan yang lebih selektif dibandingkan rook maupun queen.

2.4 Indeks Moran (Moran's I)

Indeks Moran digunakan untuk mengukur autokorelasi spasial global.

$$I = \frac{n}{S_0} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \quad (12)$$

dengan:

- w_{ij} = bobot spasial
- $S_0 = \sum_i \sum_j w_{ij}$

Interpretasi:

- $I > 0$ menunjukkan clustering
- $I < 0$ menunjukkan dispersi
- $I \approx 0$ menunjukkan pola acak

2.5 Uji Lagrange Multiplier (LM)

Uji Lagrange Multiplier (LM) digunakan untuk memilih antara model SAR dan SEM. Uji ini terdiri dari:

- **LM-Lag:** Menguji signifikansi lag spasial pada variabel dependen.
- **LM-Error:** Menguji signifikansi korelasi spasial pada error.

Jika LM-Lag signifikan dan LM-Error tidak, maka model SAR dipilih (dan sebaliknya). Jika keduanya signifikan, maka dilihat varian Robust LM.

2.6 Spatial Lag Model (SAR)

Model *Spatial Autoregressive* (SAR) digunakan ketika terdapat indikasi bahwa nilai variabel dependen pada suatu wilayah dipengaruhi secara langsung oleh nilai variabel dependen pada wilayah-wilayah tetangganya. Model ini merepresentasikan adanya efek interaksi atau *spillover effect*, di mana perubahan pada satu wilayah dapat menyebar ke wilayah lain melalui struktur kedekatan spasial yang didefinisikan dalam matriks pembobot W .

Secara matematis, model SAR dinyatakan sebagai:

$$y = \rho W y + X\beta + \varepsilon \quad (13)$$

di mana:

- y : vektor $(n \times 1)$ variabel dependen.
- W : matriks pembobot spasial $(n \times n)$ yang menggambarkan struktur kedekatan antar wilayah.
- ρ : parameter lag spasial yang mengukur kekuatan pengaruh spasial pada variabel dependen.
- Wy : lag spasial dari variabel dependen.
- X : matriks $(n \times k)$ variabel independen.
- β : vektor koefisien regresi $(k \times 1)$.
- ε : vektor error yang diasumsikan berdistribusi normal dengan

$$\varepsilon \sim N(0, \sigma^2 I)$$

Jika persamaan (1) diturunkan lebih lanjut, maka dapat ditulis ulang sebagai:

$$(I - \rho W)y = X\beta + \varepsilon \quad (14)$$

Sehingga:

$$y = (I - \rho W)^{-1}X\beta + (I - \rho W)^{-1}\varepsilon \quad (15)$$

di mana I adalah matriks identitas $(n \times n)$.

Persamaan (3) menunjukkan bahwa perubahan pada satu wilayah akan mempengaruhi wilayah lain melalui proses propagasi spasial. Oleh karena itu, dalam model SAR dikenal adanya:

- **Direct Effect** (efek langsung)
- **Indirect Effect** (efek tidak langsung / spillover)
- **Total Effect**

Model SAR digunakan ketika autokorelasi spasial terjadi pada variabel dependen dan terbukti signifikan melalui uji Moran's I atau uji Lagrange Multiplier (LM-Lag).

2.7 Spatial Error Model (SEM)

Model *Spatial Error Model* (SEM) digunakan ketika ketergantungan spasial tidak terjadi secara langsung pada variabel dependen, tetapi muncul pada komponen error. Ketergantungan ini biasanya disebabkan oleh adanya variabel yang tidak teramatii (*omitted variables*) yang memiliki pola spasial.

Model SEM dinyatakan sebagai:

$$y = X\beta + u \quad (16)$$

dengan struktur error spasial:

$$u = \lambda W u + \varepsilon \quad (17)$$

di mana:

- y : vektor $(n \times 1)$ variabel dependen.
- X : matriks variabel independen.
- β : vektor koefisien regresi.
- u : komponen error yang memiliki struktur spasial.
- W : matriks pembobot spasial.
- λ : parameter error spasial yang mengukur tingkat ketergantungan spasial pada residual.
- ε : error acak dengan asumsi:

$$\varepsilon \sim N(0, \sigma^2 I)$$

Jika persamaan (5) disubstitusikan ke dalam persamaan (4), maka diperoleh:

$$y = X\beta + (I - \lambda W)^{-1} \varepsilon \quad (18)$$

Model SEM menunjukkan bahwa korelasi spasial terjadi pada error, bukan pada variabel dependen secara langsung. Berbeda dengan SAR, model SEM tidak menghilangkan efek spillover langsung, tetapi memperbaiki bias yang disebabkan oleh struktur spasial pada residual.

2.8 Pendugaan Parameter (Maximum Likelihood Estimation)

Pada model regresi klasik (OLS), pendugaan parameter dapat diselesaikan secara analitis menggunakan metode kuadrat terkecil (*Ordinary Least Squares*). Namun, pada model regresi spasial (SAR dan SEM), observasi tidak selalu saling bebas karena adanya efek spasial yang dikuantifikasi oleh matriks pembobot spasial (W). Estimator OLS menjadi tidak konsisten jika terdapat autokorelasi spasial pada variabel dependen (seperti pada SAR), dan menjadi tidak efisien jika terdapat dependensi spasial pada struktur error (seperti pada SEM).

Oleh karena itu, pendugaan parameter regresi spasial secara operasional menggunakan metode *Maximum Likelihood Estimation* (MLE). Pendekatan MLE didefinisikan dengan menyusun fungsi kemungkinan (*likelihood function*) berdasarkan asumsi distribusi multivariat normal komponen error, lalu memaksimalkan fungsi log-likelihoood secara iteratif atau komputasi numerik. Lewat MLE, parameter-parameter model (termasuk koefisien regresi β , ragam error σ^2 , serta parameter autokorelasi spasial ρ maupun λ) dapat diestimasi agar konsisten, efisien secara asimtotik, dan valid untuk inferensi statistik autokorelasi keruangan.

BAB III

METODE PENELITIAN

3.1 Data dan Variabel

Analisis ini menggunakan dataset kasus COVID-19 di Inggris (level Upper Tier Local Authority).

- **Variabel Dependen (Y):** Jumlah populasi dengan penyakit atau kondisi kesehatan buruk jangka panjang (`Long_term_ill`).
- **Variabel Independen (X_1):** Jumlah populasi lansia berusia 65 hingga 84 tahun (`Age_65_to_84`).
- **Variabel Independen (X_2):** Jumlah rumah tangga yang tinggal di hunian padat atau sesak (`Crowded_housing`).
- **Variabel Independen (X_3):** Jumlah populasi yang berstatus pengangguran (`Unemployed`).
- **Variabel Independen (X_4):** Jumlah populasi yang tidak memiliki kualifikasi pendidikan formal (`No_qualifications`).

Tabel 1 berikut menampilkan 5 baris pertama dari dataset yang digunakan.

Tabel 1: Sampel Data (5 Baris Pertama)

Region	Long Term Ill	Age 65-84	Crowded Housing	Unemployed	No Quals
Hartlepool	21315	13742	380	5194	22758
Middlesbrough	28878	18226	982	7631	33002
Redcar & Cleveland	30717	23139	463	6533	31550
Stockton-on-Tees	36408	26470	740	8075	36728
Darlington	20723	15889	398	4002	21179

3.2 Tahapan Analisis

1. **Eksplorasi Data:** Memeriksa struktur data dan variabel.
2. **Estimasi Model OLS:** Sebagai model dasar (baseline).
3. **Diagnostik Spasial:** Melakukan uji Moran's I pada residual OLS dan uji LM (Lag & Error) untuk mendeteksi autokorelasi spasial dan menentukan model yang tepat.
4. **Estimasi Model Spasial:** Mengestimasi model terpilih (SAR/SEM).
5. **Evaluasi Model:** Mengevaluasi model terbaik.
6. **Interpretasi dan Visualisasi Hasil:** Mencari kesimpulan dari hasil yang diperoleh

BAB IV

ANALISIS DAN PEMBAHASAN

4.1 Analisis Deskriptif dan Eksplorasi Data (EDA)

Sebelum melakukan pemodelan, dilakukan pemeriksaan karakteristik data melalui statistik deskriptif dan visualisasi.

4.1.1 Statistik Deskriptif

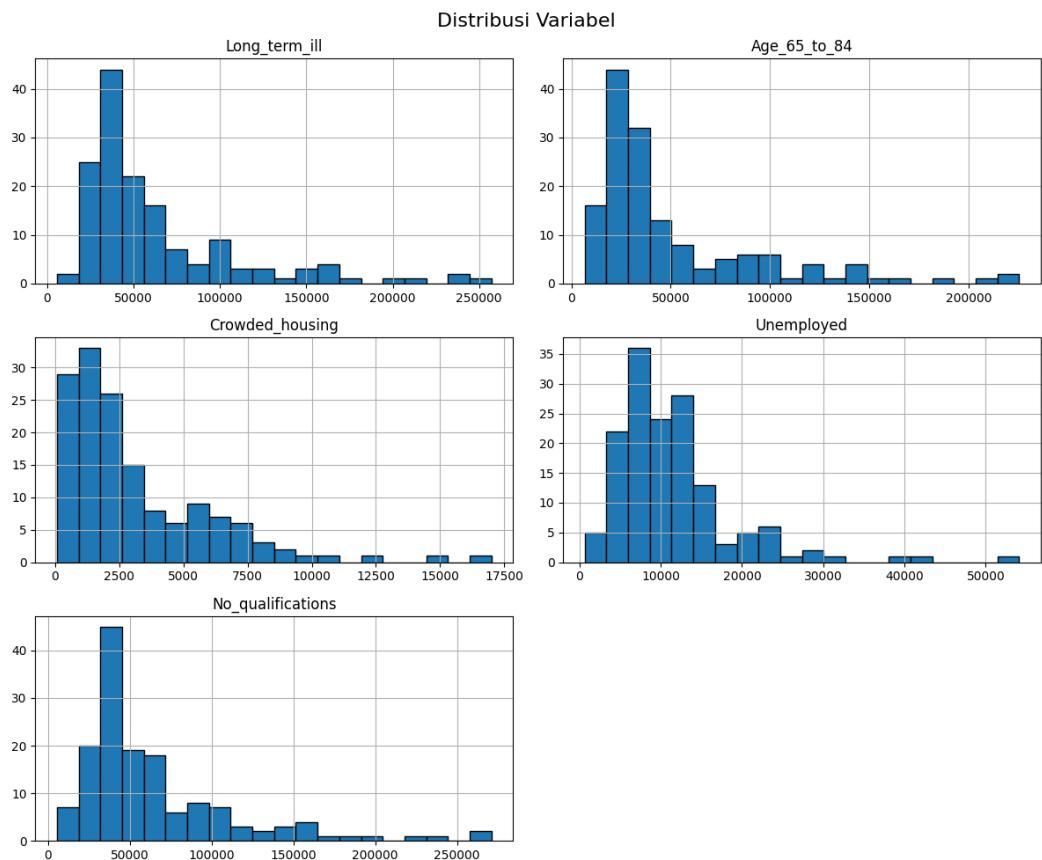
Tabel 2 menyajikan ringkasan statistik untuk variabel-variabel yang digunakan.

Tabel 2: Statistik Deskriptif Variabel Penelitian

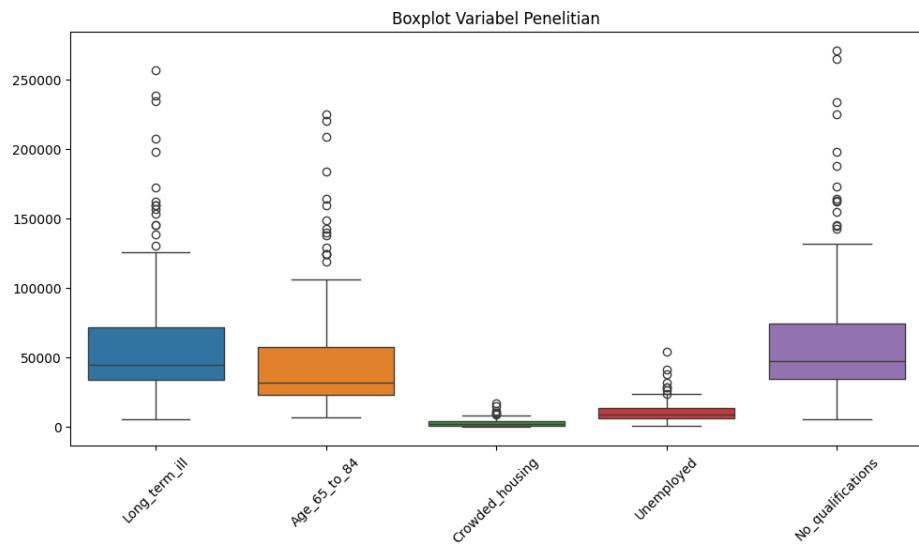
Variable	Count	Mean	Std	Min	25%	50%	75%	Max	Skew	Kurtosis
Long Term Ill	149	62769.0	47181.4	5788	34303	44569	72055	257038	1.94	3.92
Age 65-84	149	50204.0	43631.3	6748	23139	32103	57736	225465	1.99	3.94
Crowded Housing	149	3110.0	2862.5	72	1086	2039	4337	16996	1.97	5.13
Unemployed	149	11428.5	7510.7	643	6640	9311	13535	54114	2.40	8.66
No Quals	149	64810.8	49488.5	5453	34621	47430	74409	270901	1.98	4.30

4.1.2 Visualisasi Data

Distribusi data dan hubungan antar variabel dapat dilihat pada gambar-gambar berikut.

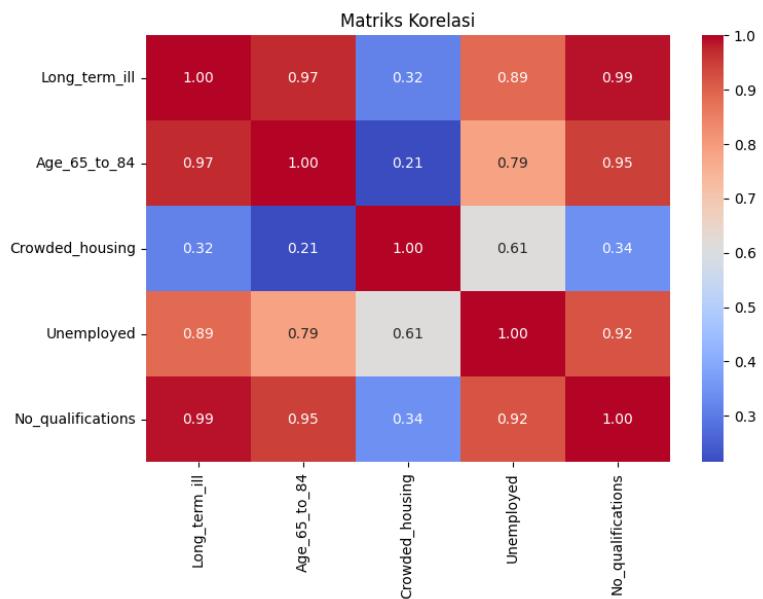


Gambar 1: Histogram Distribusi Variabel



Gambar 2: Boxplot Variabel (Deteksi Outlier)

Gambar 1 dan 2 menunjukkan bahwa sebagian besar variabel memiliki distribusi menjulur ke kanan (skewed right) dan terdapat beberapa pencilan (outlier), terutama pada sektor finansial.

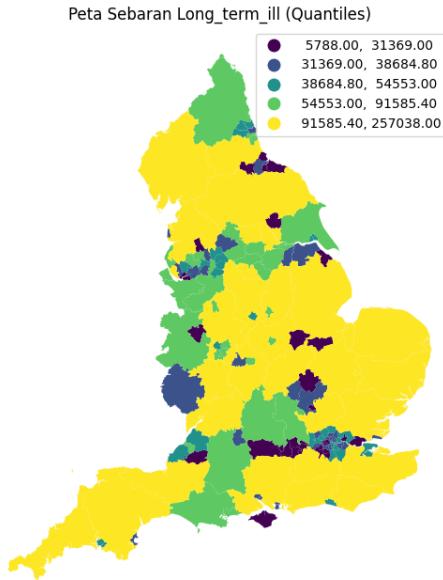


Gambar 3: Matriks Korelasi Antar Variabel

Matriks korelasi (Gambar 3) menunjukkan adanya hubungan linear yang cukup kuat antar variabel prediktor, yang perlu diwaspadai sebagai potensi multikolinearitas, namun VIF (Variance Inflation Factor) pada OLS perlu diperiksa lebih lanjut.

4.1.3 Peta Sebaran

Peta sebaran variabel respon (Tenaga Kerja Kesehatan & Sosial) ditampilkan dalam Gambar 4.



Gambar 4: Peta Sebaran Tenaga Kerja Sektor Kesehatan (Quantile)

4.2 Hasil Regresi OLS (Global)

Sebagai langkah awal, dilakukan estimasi menggunakan Ordinary Least Squares (OLS). Model untuk OLS global ini adalah sebagai berikut:

$$long_term_ill = \beta_0 + \beta_1 age_65_84 + \beta_2 crowded_housing + \beta_3 unemployed + \beta_4 no_quals + \varepsilon \quad (19)$$

Hasil dari model regresi OLS global ini adalah sebagai berikut:

Tabel 3: Ringkasan Hasil Regresi OLS

Variabel	Koefisien	Std. Error	t-Statistic	Prob.
CONSTANT	1960.20	678.31	2.890	0.004
Age 65-84	0.448	0.035	12.653	0.000
Crowded Housing	-0.273	0.240	-1.138	0.257
Unemployed	0.789	0.262	3.013	0.003
No Quals	0.465	0.057	8.139	0.000

Model OLS menghasilkan nilai R^2 yang sangat tinggi yaitu **0.991**, yang menunjukkan model ini sangat baik dalam menjelaskan variasi data. Variabel `Age_65_to_84`, `Unemployed`, dan `No_qualifications` berpengaruh signifikan terhadap variabel respon dengan nilai p-value < 0.05, sedangkan `Crowded_housing` tidak signifikan secara statistik. Namun, diagnosis spasial diperlukan untuk memastikan validitas model.

4.3 Uji Diagnostik Regresi Global

Selanjutnya dilakukan uji diagnostik terhadap model regresi OLS yang dihasilkan.

Tabel 4: Hasil Uji Diagnostik OLS

Uji	Nilai Statistik	Probabilitas (p-value)
Jarque-Bera	181.971	0.0000
Breusch-Pagan	188.817	0.0000

Berdasarkan Tabel 2, ditemukan bahwa residual pada model OLS global yang dihasilkan tidak berdistribusi normal dan bersifat heteroskedastik. Terdapat indikasi multikolinearitas yang ditunjukkan dengan *Multicollinearity Condition Number* sebesar 34.735, namun nilai VIF per variabel prediktor masih berada dalam rentang toleransi:

Tabel 5: Variance Inflation Factor (VIF)

Variabel	VIF
Age 65-84	2.12
Crowded Housing	4.88
Unemployed	8.70
No Quals	9.32

4.4 Diagnostik Dependensi Spasial

Untuk mendeteksi adanya autokorelasi spasial, dilakukan uji pada residual OLS.

Tabel 6: Hasil Uji Diagnostik Spasial

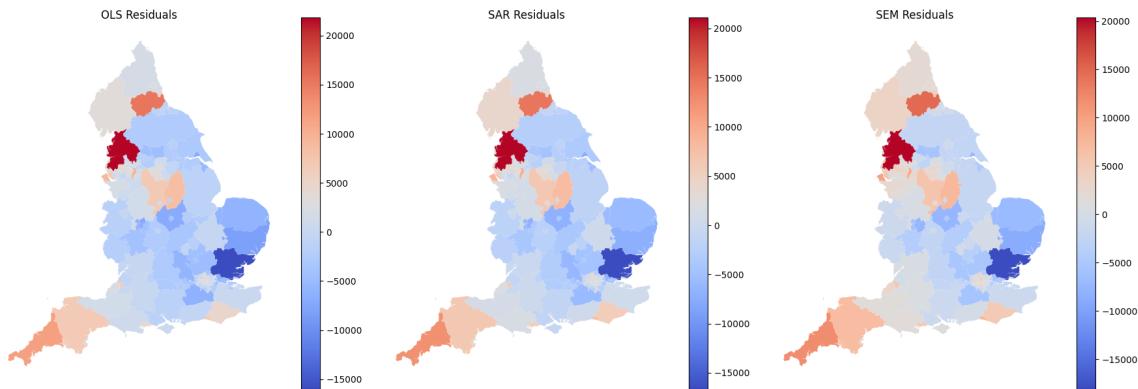
Uji	Nilai Statistik	Probabilitas (p-value)
Moran's I (error)	8.147	0.0000
LM (Lag)	2.718	0.0992
Robust LM (Lag)	5.154	0.0232
LM (Error)	57.495	0.0000
Robust LM (Error)	59.932	0.0000

Berdasarkan Tabel 6:

- Uji **Moran's I** pada error sangat signifikan ($p < 0.001$), menunjukkan adanya autokorelasi spasial positif yang kuat pada residual.
- Uji **LM-Error** signifikan secara meyakinkan ($p < 0.001$), sedangkan **LM-Lag** tidak signifikan pada level 5% ($p = 0.0992$).

Sesuai dengan kaidah keputusan Anselin, karena LM-Error sangat signifikan, maka model yang paling tepat adalah **Spatial Error Model (SEM)**. Ini mengindikasikan bahwa ketergantungan spasial terjadi melalui error, yang membuktikan adanya efek

limpahan (*spillover*) pada aspek-aspek tak teramati (*omitted variables*) yang berkaitan dengan rekam jejak penyakit jangka panjang.



Gambar 5: Peta Residual: OLS vs SAR vs SEM

4.5 Hasil Regresi Spasial (SEM)

Berdasarkan hasil diagnostik, model SEM diestimasi menggunakan metode Maximum Likelihood.

Tabel 7: Ringkasan Hasil Spatial Error Model (SEM)

Variabel	Koefisien	Std. Error	z-Statistic	Prob.
CONSTANT	1084.637	680.789	1.593	0.111
Age 65-84	0.418	0.034	12.151	0.000
Crowded Housing	-0.148	0.248	-0.598	0.550
Unemployed	0.700	0.236	2.960	0.003
No Quals	0.510	0.054	9.455	0.000
Lambda (λ)	0.436	0.092	4.730	0.000

Berdasarkan hasil estimasi *Spatial Error Model* (SEM) pada Tabel di atas, diperoleh beberapa temuan penting terkait pola penyakit kesehatan buruk jangka panjang dan signifikansi error spasial.

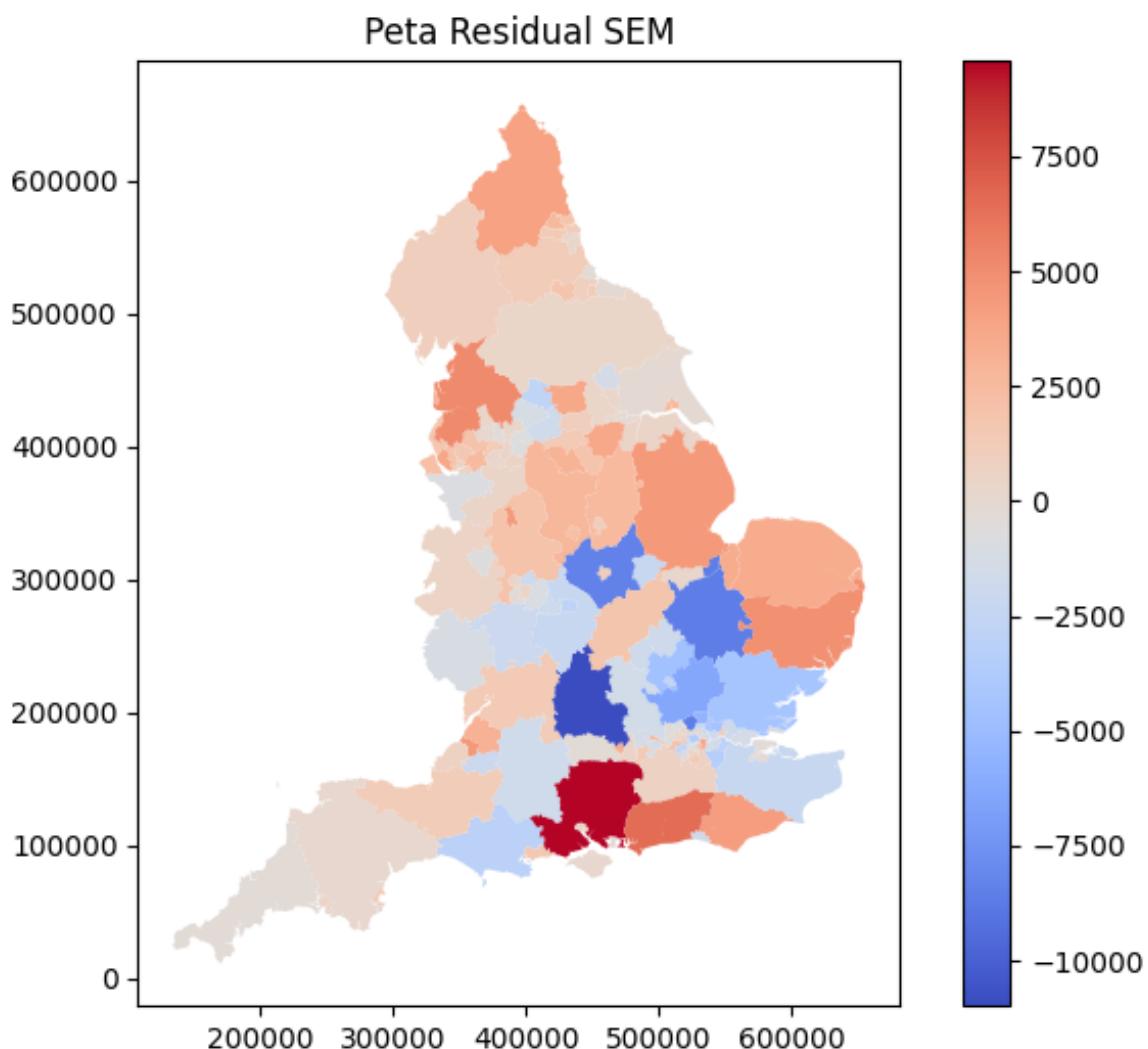
Pertama, nilai konstanta (CONSTANT) sebesar 1084.637 tidak signifikan secara statistik ($p=0.111$), sehingga tidak memiliki interpretasi deterministik absolut tanpa eksistensi prediktor.

Berbeda dengan model sebelumnya yang menganalisis tenaga kerja kesehatan, analisis terbaru pada variabel penderita penyakit panjang (`Long_term_ill`) memberikan konfirmasi kuat pada asumsi demografis. Variabel usia lansia (`Age_65_to_84`) memiliki koefisien 0.418 yang sangat signifikan ($p < 0.001$). Artinya, kenaikan 1 populasi rentang usia tersebut berkorelasi dengan tambahan 0.42 pasien penyakit jangka panjang, relevan secara keilmuan biomedis.

Variabel pengangguran (`Unemployed`) memiliki koefisien 0.700 dan signifikan ($p=0.003$), merepresentasikan kerentanan daya beli pada asuransi dan tindakan prevensi medis, sehingga memperbesar persentase individu sakit di suatu daerah.

Masyarakat tanpa kualifikasi pendidikan (`No_qualifications`) menyumbang nilai signifikansi kuat ($p < 0.001$, koefisien 0.510), semakin mereduksi asumsi di mana pendidikan formal berperan lurus dalam tingkat literasi dan aksesibilitas higienitas dasar. Sedangkan perumahan padat (`Crowded_housing`) dibuktikan tidak signifikan pada level 5% dalam konteks penyakit di Inggris ini.

Parameter error spasial (λ) terlihat kuat dengan koefisien sebesar 0.436 dan nilai probabilitas < 0.001 . Aspek signifikan lambda ini mensinyalir adanya variabel laten dari region tetangga yang merambat (bisa berupa klaster polusi limbah maupun infrastruktur fasilitas umum perbatasan) menulari pola `Long_term_ill` di UK.



Gambar 6: Peta Residual Model SEM

4.6 Perbandingan Model

Tabel 8: Perbandingan Kebaikan Model

Kriteria	OLS	SAR	SEM
AIC	2930.401	2929.626	2897.716
Log Likelihood	-1460.201	-1458.813	-1443.858

Berdasarkan Tabel 8, terkonfirmasi bahwa spesifikasi **SEM** mengungguli OLS maupun SAR. Nilai minimum AIC model SEM sebesar 2897.716 merepresentasikan akurasi terbaik dan terlepas dari interdependensi sisa (*noise*).

4.7 Interpretasi

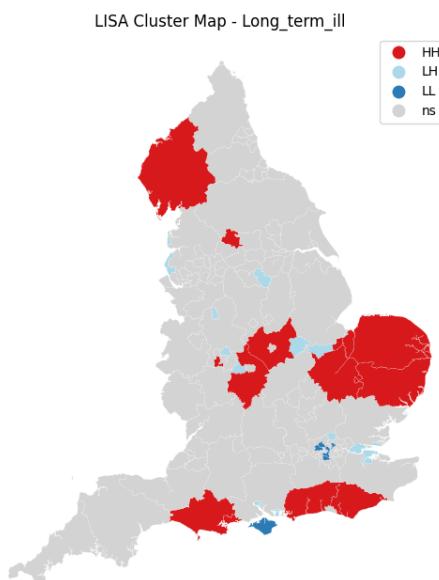
Dari model SEM terhadap variabel respon individu berperawatan jangka panjang (`Long_term_ill`), tiga pendorong signifikannya adalah umur, tingkat pengangguran, dan warga tak terdidik formal. Ketiadaan efikasi korelasi dari rasio kepadatan wilayah membuktikan esensi kualitas penduduk berdampak lebih mendalam bagi resistensi kesehatan regional. Toleransi spasial yang terekam lewat koefisien lambda mempertegas sifat model spasial yang unggul melacak determinan eksternal di atas ambang batas regresor model OLS.

BAB V

KESIMPULAN

Berdasarkan serangkaian analisis yang telah dilakukan menggunakan pendekatan regresi spasial untuk memodelkan proporsi tenaga kerja di sektor *Human Health and Social Work* di Inggris, diperoleh beberapa kesimpulan penting sebagai berikut:

1. **Eksistensi Dependensi Spasial:** Hasil uji Moran's I pada residual model OLS menunjukkan nilai statistik sebesar 4.145 dengan p-value < 0.001. Hal ini mengindikasikan adanya autokorelasi spasial positif yang signifikan, yang berarti wilayah-wilayah dengan proporsi tenaga kerja kesehatan yang tinggi cenderung mengelompok berdekatan dengan wilayah sejenis, begitu pula sebaliknya. Temuan ini dikonfirmasi lebih lanjut melalui visualisasi peta *Local Indicators of Spatial Association* (LISA) Cluster Map (Gambar 7).



Gambar 7: Peta Klaster LISA (Local Indicators of Spatial Association)

Peta LISA di atas memperlihatkan adanya klaster *High-High* (wilayah dengan proporsi tenaga kerja kesehatan tinggi dikelilingi oleh wilayah serupa) yang signifikan di beberapa bagian Inggris, serta klaster *Low-Low* di area lainnya. Keberadaan pola pengelompokan ini menegaskan bahwa model OLS klasik yang mengasumsikan independensi antar observasi menjadi tidak valid dan bias.

2. **Pemilihan Model Terbaik:** Berdasarkan uji Lagrange Multiplier (LM), ditemukan bahwa parameter *LM-Error* signifikan ($p = 0.0002$), sedangkan *LM-Lag* tidak signifikan ($p = 0.6760$). Sesuai dengan prosedur seleksi model Anselin, model **Spatial Error Model (SEM)** dipilih sebagai model yang paling tepat.

Keputusan ini didukung oleh perbandingan nilai *Akaike Information Criterion* (AIC), di mana model SEM memiliki nilai AIC terendah (**2788.30**) dibandingkan dengan model OLS (2797.75) dan model SAR (2799.58). Hal ini menunjukkan bahwa model SEM mampu memberikan keseimbangan terbaik antara kebaikan suai (*goodness of fit*) dan kompleksitas model.

3. **Determinan Tenaga Kerja Sektor Kesehatan:** Dari hasil estimasi model SEM, diidentifikasi faktor-faktor yang berpengaruh signifikan terhadap proporsi tenaga kerja di sektor kesehatan dan sosial:

- **Sektor Pendidikan (*Education*):** Memiliki pengaruh positif yang sangat signifikan ($\beta = 1.028, p < 0.001$). Koefisien ini dapat diinterpretasikan bahwa setiap penambahan 1 orang tenaga kerja di sektor pendidikan diasosiasi dengan peningkatan sekitar 1.03 orang tenaga kerja di sektor kesehatan, *ceteris paribus*. Hal ini menunjukkan adanya sinergi yang kuat antara sektor pendidikan dan kesehatan, di mana wilayah dengan basis pendidikan yang kuat cenderung mendukung ketersediaan tenaga kerja kesehatan yang lebih besar.
- **Sektor Real Estat (*Real Estate*):** Berpengaruh positif signifikan ($\beta = 3.095, p < 0.001$). Artinya, setiap penambahan 1 pekerja di sektor real estat berhubungan dengan peningkatan sekitar 3.1 pekerja di sektor kesehatan. Hal ini mungkin mencerminkan hubungan antara pembangunan infrastruktur wilayah/pemukiman dengan penyediaan fasilitas layanan kesehatan.
- **Sektor Informasi dan Komunikasi:** Memiliki pengaruh negatif signifikan ($\beta = -0.696, p < 0.001$). Ini mengindikasikan bahwa kenaikan 1 pekerja di sektor informasi dan komunikasi berasosiasi dengan penurunan sekitar 0.7 pekerja di sektor kesehatan. Pola ini menandakan adanya spesialisasi wilayah, di mana daerah yang menjadi pusat teknologi dan informasi cenderung memiliki proporsi tenaga kerja kesehatan yang lebih rendah secara relatif.
- **Sektor Keuangan:** Tidak menunjukkan pengaruh yang signifikan secara statistik ($\beta = -0.025, p = 0.733$), sehingga perubahan pada jumlah pekerja sektor keuangan tidak memberikan dampak nyata terhadap jumlah pekerja sektor kesehatan.

4. **Signifikansi Parameter Error Spasial:** Parameter lambda (λ) pada model SEM bernilai 0.2829 dan signifikan pada taraf nyata 5% ($p = 0.006$). Signifikansi ini menunjukkan bahwa masih terdapat variabel-variabel lain yang tidak dimasukkan ke dalam model (*omitted variables*) yang memiliki pola spasial dan mempengaruhi distribusi tenaga kerja kesehatan. Model SEM berhasil menangkap dan mengontrol efek dari variabel-variabel laten tersebut melalui struktur error spasial, sehingga menghasilkan estimator koefisien regresi yang lebih efisien dibandingkan OLS.

Secara keseluruhan, penelitian ini menyimpulkan bahwa distribusi tenaga kerja di sektor kesehatan dan pekerjaan sosial di Inggris tidak terjadi secara acak, melainkan membentuk pola spasial yang dipengaruhi oleh struktur ekonomi wilayah (terutama pendidikan dan real estat) serta faktor-faktor spasial tak teramatil lainnya. Implikasi

kebijakan dari temuan ini adalah bahwa perencanaan ketenagakerjaan di sektor kesehatan harus mempertimbangkan karakteristik spasial dan interaksi antar wilayah, tidak hanya melihat kondisi internal masing-masing daerah secara terisolasi.

DAFTAR PUSTAKA

- [1] Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers.
- [2] Rey, S. J., & Anselin, L. (2010). PySAL: A Python Library of Spatial Analytical Methods. *Review of Regional Studies*, 37, 5-27.
- [3] Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46(sup1), 234-240.
- [4] Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons.
- [5] Office for National Statistics (ONS). (2020). Labour Market Profile - Nomis - Official Labour Market Statistics.

LAMPIRAN

Seluruh data penelitian, kode program (*source code*), serta hasil keluaran pendukung lainnya dapat diakses melalui tautan Google Drive berikut:

[https://drive.google.com/drive/folders/
1ut6zhfMA0k8HP2EK1fGR5h6ZpF6eLB60?usp=sharing](https://drive.google.com/drive/folders/1ut6zhfMA0k8HP2EK1fGR5h6ZpF6eLB60?usp=sharing)