



## Analisis Faktor-Faktor yang Mempengaruhi Angka Harapan Hidup (AHH) di Papua Menggunakan Regresi Logistik Biner

Nur Aida Rahmasari<sup>1</sup>, Indah Rahma Abdillah<sup>2</sup>, Fretty Novelyna Ramadhani<sup>3</sup>, Aviolla Terza Damaliana<sup>4</sup>, Shindi Shella May Wara<sup>5</sup>

<sup>1,2,3,4,5</sup>Sains Data, Universitas Pembangunan Nasional “Veteran” Jawa Timur

<sup>1</sup>[23083010024@student.upnjatim.ac.id](mailto:23083010024@student.upnjatim.ac.id)

<sup>2</sup>[23083010002@student.upnjatim.ac.id](mailto:23083010002@student.upnjatim.ac.id)

<sup>3</sup>[23083010005@student.upnjatim.ac.id](mailto:23083010005@student.upnjatim.ac.id)

<sup>4</sup>[aviolla.terza.sada@upnjatim.ac.id](mailto:aviolla.terza.sada@upnjatim.ac.id)

<sup>5</sup>[shindi.shella.fasilkom@upnjatim.ac.id](mailto:shindi.shella.fasilkom@upnjatim.ac.id)

Corresponding author email: [23083010024@student.upnjatim.ac.id](mailto:23083010024@student.upnjatim.ac.id)

**Abstract:** Life Expectancy (LE) is one of the key indicators used to assess the level of health and well-being in a region. In Papua Province, the LE remains below the national average, indicating specific challenges in efforts to improve the quality of life of its population. This study analyzes factors influencing LE using a binary logistic regression model to classify LE into high or low categories. The analysis results show that variables such as the percentage of people living in poverty, complete immunization coverage, number of medical personnel, number of healthcare facilities, and Gross Regional Domestic Product (GRDP) per capita have a significant influence on Life Expectancy (LE) in Papua, while the illiteracy rate does not have a significant effect in the model. These findings can serve as a basis for policy formulation to improve the health status and welfare of the people in Papua Province.

**Keywords:** LE, Binary Logistic Regression, Papua

**Abstrak:** Angka Harapan Hidup (AHH) menjadi salah satu indikator utama untuk menilai tingkat kesehatan dan kesejahteraan suatu daerah. Di Provinsi Papua, nilai AHH masih berada di bawah rata-rata nasional, yang mengindikasikan adanya tantangan tersendiri dalam upaya meningkatkan kualitas hidup penduduknya. Penelitian ini menganalisis faktor-faktor yang mempengaruhi AHH menggunakan model regresi logistik biner untuk mengklasifikasikan kategori AHH tinggi atau rendah. Hasil analisis menunjukkan bahwa variabel seperti persentase penduduk miskin, cakupan imunisasi lengkap, jumlah tenaga medis, jumlah layanan kesehatan, serta Produk Domestik Regional Bruto (PDRB) per kapita berpengaruh signifikan terhadap Angka Harapan Hidup (AHH) di Papua, sedangkan tingkat buta huruf tidak memiliki pengaruh signifikan dalam model. Temuan ini dapat menjadi dasar bagi perumusan kebijakan untuk meningkatkan derajat kesehatan dan kesejahteraan masyarakat di Provinsi Papua.

**Kata kunci:** AHH, Regresi Logistik Biner, Papua

### 1. PENDAHULUAN

Tingginya tingkat kesehatan masyarakat mencerminkan keberhasilan pelaksanaan program kesehatan serta pembangunan sosial ekonomi, yang secara tidak langsung berdampak pada peningkatan Angka Harapan Hidup (AHH). Menurut Badan Pusat Statistik [1], AHH saat lahir (*life expectancy at birth*) merupakan perkiraan rata-rata usia hidup yang akan dicapai oleh bayi yang lahir pada tahun tertentu. Pada tahun 2023, AHH di Papua tercatat sebesar 64,53 tahun untuk laki-laki dan 68,45 tahun untuk perempuan, yang masih jauh tertinggal dibandingkan nasional yaitu 71,9 tahun untuk laki-laki dan 75,5 tahun untuk perempuan [2]. Kondisi tersebut menandakan adanya ketimpangan dalam kualitas layanan kesehatan dan faktor sosial ekonomi di Papua.

Berbagai penelitian terdahulu telah mengidentifikasi sejumlah faktor yang secara signifikan mempengaruhi Angka Harapan Hidup (AHH). Septianingsih [3], melalui pemodelan data panel dengan *Random Effect Model*, mengungkapkan bahwa persentase penduduk miskin memiliki hubungan signifikan dengan AHH di Indonesia. Kemiskinan berdampak negatif terhadap akses layanan kesehatan, pola hidup sehat, ketersediaan gizi yang memadai, serta kualitas lingkungan hidup. Di sisi lain, Wardhana



& Kharisma [4], dalam penelitiannya di Jawa Barat menggunakan model *Generalized Least Squares* (GLS), menunjukkan bahwa Produk Domestik Regional Bruto (PDRB) berkorelasi positif dengan AHH di Jawa Barat, karena kesejahteraan ekonomi meningkatkan akses terhadap pelayanan kesehatan dan kualitas hidup. Selain itu, Amalia & Mahmudah [5], menemukan bahwa angka buta huruf berdampak negatif terhadap AHH di Jawa Timur, mengingat pendidikan yang lebih baik meningkatkan kesadaran dan perilaku hidup sehat, serta kemampuan mengakses informasi kesehatan.

Selain faktor ekonomi dan pendidikan, faktor kesehatan masyarakat juga berperan sangat penting dalam menentukan AHH. Sugiantari & Budiantara [6], dalam penelitiannya menggunakan regresi semiparametrik spline, menemukan bahwa cakupan imunisasi lengkap merupakan determinan penting dalam meningkatkan AHH di Jawa Timur, karena berperan dalam pencegahan penyakit menular yang dapat menyebabkan kematian dini. Sejalan dengan itu, Pratiwi & Budyanra [7], melalui analisis determinan dengan regresi data panel di Maluku, menemukan bahwa rasio puskesmas per kecamatan sebagai representasi ketersediaan tenaga medis dan fasilitas kesehatan berpengaruh signifikan terhadap AHH. Semakin baik distribusi tenaga medis dan fasilitas kesehatan, semakin tinggi kualitas layanan yang diterima masyarakat, sehingga mendukung peningkatan harapan hidup.

Dengan berbagai kompleksitas yang dihadapi, penting untuk mengkaji faktor-faktor yang mempengaruhi rendahnya Angka Harapan Hidup (AHH) di Papua. Regresi logistik biner dipilih karena sesuai untuk variabel dependen yang bersifat kategorik, serta mampu menunjukkan pengaruh masing-masing variabel independen. Metode ini turut digunakan oleh Chumairah [8] dalam penelitiannya untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap AHH di Indonesia. Penelitian ini bertujuan untuk mengisi kesenjangan kajian sebelumnya dengan memfokuskan pada wilayah Papua serta mengeksplorasi variabel-variabel yang belum diteliti, seperti angka buta huruf, cakupan imunisasi, jumlah tenaga medis, dan PDRB, untuk menganalisis faktor-faktor yang memengaruhi Angka Harapan Hidup.

Melalui regresi logistik biner, hubungan antara variabel-variabel sosial, ekonomi, kesehatan, dan pendidikan dapat dianalisis secara simultan. Hasil analisis diharapkan memberikan gambaran mengenai faktor-faktor yang mempengaruhi AHH di Papua, serta menjadi dasar bagi perumusan kebijakan pembangunan yang lebih terarah. Dengan memahami faktor utama yang berperan, intervensi yang lebih efektif dan tepat sasaran dapat dirancang untuk mendorong peningkatan kesejahteraan dan kualitas hidup masyarakat di Papua.

## 2. METODE PENELITIAN

### 2.1. Data dan Sumber Data

Penelitian ini menggunakan data sekunder yang bersumber dari Badan Pusat Statistik (BPS) tahun 2023, mencakup 29 kabupaten/kota di Provinsi Papua. Penelitian ini melibatkan satu variabel dependen dan enam variabel independen, sebagaimana ditampilkan dalam Tabel 1.

Tabel 1. Variabel Data

Kode	Variabel	Satuan/Pengukuran	Jenis Data
$x_1$	Persentase Penduduk Miskin	Persen (%)	Rasio
$x_2$	Cakupan Imunisasi Lengkap	Persen (%)	Rasio
$x_3$	Jumlah Tenaga Medis	Orang	Rasio
$x_4$	Jumlah Layanan Kesehatan	Unit	Rasio
$x_5$	Angka Buta Huruf	Persen (%)	Rasio



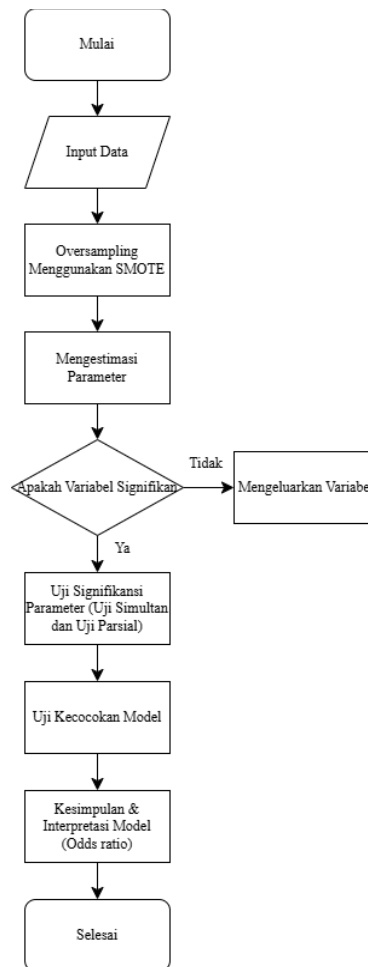
$x_6$	Produk Domestik Regional Bruto (PDRB)	Rupiah (Rp) Per Kapita	Rasio
Y	Angka Harapan Hidup (AHH)	0 = Rendah, 1 = Tinggi	Nominal

Variabel dependen dalam penelitian ini adalah Angka Harapan Hidup (AHH) yang telah diklasifikasikan ke dalam dua kategori, yakni “Rendah” dan “Tinggi”. Penetapan kategori ini mengacu pada metode yang digunakan oleh Chumairah [8], yang merujuk pada penelitian Angraini & Listyaningsih [9]. Aturan kategorisasi adalah sebagai berikut:

- Kategori AHH (y) rendah (kode 0), apabila nilai AHH < rata-rata AHH.
- Kategori AHH (y) tinggi (kode 1), apabila nilai AHH  $\geq$  rata-rata AHH.

### 2.2. Metode yang Digunakan

Penelitian ini dilakukan dengan memanfaatkan perangkat lunak statistik Python sebagai alat bantu dalam proses pengolahan dan analisis data. Adapun tahapan penelitian ini dijabarkan sebagai berikut:



**Gambar 1** Alur Proses Analisis Faktor-Faktor yang Mempengaruhi AHH di Papua



Berdasarkan diagram alur tersebut, pendekatan yang digunakan untuk memodelkan hubungan antar variabel adalah regresi logistik biner dengan langkah-langkah sebagai berikut:

- Regresi Logistik Biner

Regresi logistik biner merupakan salah satu metode pemodelan matematis yang digunakan untuk menganalisis hubungan antara satu atau lebih variabel independen dengan variabel dependen yang bersifat dikotomis atau hanya memiliki dua kategori, seperti  $Y=1$  untuk menyatakan “berhasil” dan  $Y=0$  untuk “tidak berhasil” [10]. Tujuan dari metode ini adalah untuk mengetahui faktor-faktor sosial, ekonomi, dan kesehatan yang paling berpengaruh terhadap kemungkinan suatu kabupaten/kota masuk dalam kategori AHH yang tinggi. Secara umum, persamaan model regresi logistik biner dapat dituliskan sebagai berikut:

$$P(y | x) = \pi(x) \quad [11] \quad (1)$$

Di mana:

$P(y | x)$  = Probabilitas terjadinya  $Y$  dengan kondisi variabel  $x$  tertentu

$y$  = Variabel dependen

$x$  = Variabel independen

$\pi(x)$  = Notasi singkatan dari probabilitas bahwa  $Y=1$  jika diberi masukan  $x$

Dengan:

$$\pi(x) = \frac{\exp(\beta + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta + \beta_1 x_1 + \dots + \beta_p x_p)} \quad [11] \quad (2)$$

Di mana:

$\pi(x)$  = Probabilitas bahwa  $Y=1$  (kejadian yang diharapkan)

$\exp()$  = notasi singkatan dari eksponen (pangkat natural)

$\beta_0$  = Konstanta (intercept) nilai awal model saat semua  $x_i = 0$

$\beta_1, \beta_2, \dots, \beta_p$  = Koefisien regresi yang mewakili pengaruh masing-masing variabel independen

$x_1, x_2, \dots, x_p$  = Nilai dari tiap variabel independen

Karena model logit bersifat nonlinier terhadap parameter-parameternya, maka diperlukan transformasi logit agar model tersebut dapat diubah ke dalam bentuk linear. Melalui Transformasi ini, diperoleh model regresi logistik yang linear terhadap fungsi logit. Sebagaimana ditunjukkan berikut:

$$\hat{g}(x) = \ln \frac{\pi(x)}{1-\pi(x)} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad [11] \quad (3)$$

Di mana:

$\hat{g}(x)$  = Fungsi logit, yaitu log odds dari probabilitas keberhasilan

$\pi(x)$  = Notasi singkatan dari probabilitas bahwa  $Y=1$  jika diberi masukan  $x$

$\frac{\pi(x)}{1-\pi(x)}$  = Rasio peluang (odds) dari keberhasilan terhadap kegagalan

$\beta_0$  = Konstanta (intercept)

$\beta_1, \beta_2, \dots, \beta_p$  = koefisien regresi untuk masing-masing variabel independen



- **Oversampling Menggunakan SMOTE**  
SMOTE (*Synthetic Minority Oversampling Technique*) adalah teknik yang digunakan untuk menangani masalah ketidakseimbangan kelas (*Class Imbalance Problem*). Metode ini bekerja dengan menyeimbangkan dataset yang timpang melalui pembuatan data sintetis baru dari kelas yang jumlahnya lebih sedikit [12].

- **Mengestimasi Parameter**  
Metode *Maximum Likelihood* digunakan untuk mengestimasi parameter dalam regresi logistik dengan memaksimalkan probabilitas data yang diamati. Proses ini melibatkan pembentukan *likelihood function*, yaitu fungsi yang menyatakan peluang data berdasarkan parameter yang belum diketahui, bentuk umum dari fungsi *likelihood* dalam regresi logistik biner dapat dirumuskan sebagai berikut:

$$\pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \quad [13] \quad (4)$$

Di mana:

$\pi(x_i)$  = Probabilitas bahwa  $y_1 = 1$  untuk input  $x_i$

$1 - \pi(x_i)$  = Probabilitas bahwa  $y_1 = 0$  untuk input  $x_i$

$x_i$  = Nilai-nilai variabel independen untuk observasi ke- $i$

$y_i$  = Nilai variabel dependen untuk observasi ke- $i$  (0 atau 1)

- **Uji Signifikansi Parameter**
  - **Uji Simultan**  
Hipotesis statistik berikut diajukan untuk menguji apakah terdapat pengaruh signifikan secara simultan dari seluruh variabel independen terhadap variabel dependen dalam model regresi logistik biner:

$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$  (Seluruh variabel independen secara simultan tidak menunjukkan pengaruh yang bermakna)

$H_1$ : Paling tidak ada satu  $\beta_j \neq 0$  (Terdapat minimal satu variabel independen yang memiliki pengaruh bermakna)

Nilai *G-test* mengikuti distribusi *chi-square* dengan derajat kebebasan sebesar jumlah variabel prediktor ( $k$ ). Apabila nilai *p-value*  $< \alpha$  (0.05), maka  $H_0$  ditolak, yang berarti model signifikan secara simultan. Adapun rumus statistik *G* sebagai berikut:

$$G = -2 \ln \left[ \frac{L_0}{L_p} \right] \quad [14] \quad (5)$$

Di mana:

$L_0$  = *Likelihood* dari model nol (hanya konstanta/intercept)

$L_p$  = *Likelihood* dari model lengkap (dengan variabel independen)

- **Uji Parsial**  
Melalui uji parsial, dapat diketahui apakah setiap variabel independen secara sendiri-sendiri memberikan pengaruh yang bermakna terhadap variabel dependen. Berikut perumusan hipotesisnya:



$H_0: \beta_i = 0$  (Variabel independen ke- $i$  tidak memiliki pengaruh yang signifikan terhadap variabel dependen)

$H_1: \beta_i \neq 0$  (Variabel independen ke- $i$  memiliki pengaruh yang signifikan terhadap variabel dependen)

Uji statistik dilakukan menggunakan uji Wald yang mengikuti distribusi chi-square dengan 1 derajat kebebasan. Keputusan pengujian diambil dengan membandingkan nilai statistik Wald dengan nilai kritis chi-square pada tingkat signifikansi  $\alpha$  (0.05). Jika  $W > X^2_{(1,\alpha)}$ , Maka  $H_0$  ditolak, yang berarti variabel independen ke- $i$  berpengaruh signifikan secara parsial terhadap variabel dependen. Adapun rumus uji Wald adalah sebagai berikut:

$$W = \left( \frac{\beta_i}{SE(\hat{\beta}_i)} \right)^2 \quad [14] \quad (6)$$

Di mana:

$\beta_i$  = Estimasi koefisien variabel independen ke- $i$

$SE(\hat{\beta}_i)$  = Standar error koefisien variabel bebas ke- $i$

- Uji Kecocokan Model

Uji kecocokan model bertujuan menilai sejauh mana model sesuai dengan data, yakni apakah nilai observasi mendekati nilai yang diperkirakan oleh model. Goodness of Fit (GoF) terpenuhi ketika prediksi model sesuai dengan data aktual. Dalam regresi logistik, pengujian ini dilakukan dengan uji *chi-square* melalui uji *Hosmer and Lemeshow* pada tingkat signifikansi 5%. Hipotesis statistik untuk uji kecocokan model adalah sebagai berikut:

$H_0$ : Model yang diuji sesuai (fit) dengan data

$H_1$ : Model yang diuji tidak sesuai (tidak fit) dengan data

Keputusan diambil berdasarkan p-value:

Jika p-value  $> 0.05$ , maka  $H_0$  gagal ditolak, model dianggap sesuai (fit) dengan data

Jika p-value  $\leq 0.05$ , maka  $H_0$  ditolak, model tidak sesuai (tidak fit) dengan data

Adapun rumus statistik *Hosmer and Lemeshow* (H-L) sebagai berikut:

$$C = \sum_{g=1}^G \frac{(O_g - E_g)^2}{E_g(1 - P_g)} \quad [15] \quad (7)$$

Di mana:

$O_g$  = Observasi aktual pada grup ke- $g$

$E_g$  = Nilai yang diharapkan (ekspektasi) pada grup ke- $g$ , yaitu hasil prediksi model

$P_g$  = Probabilitas keberhasilan pada grup ke- $g$

$G$  = Jumlah grup atau kategori yang diamati

- Odds Ratio



Nilai peluang logit dalam model regresi dengan dua kategori, yaitu 0 dan 1, menunjukkan perbandingan peluang terjadinya kategori 1 terhadap kategori 0. Rasio peluang ini disebut *odds ratio* (OR), yang merupakan perbandingan antara kemungkinan suatu kejadian terjadi ( $x = 1$ ) dengan kemungkinan tidak terjadi ( $x = 0$ ), dan dinyatakan dalam bentuk persamaan berikut:

$$OR = \frac{\pi(1)/1-\pi(1)}{\pi(0)/1-\pi(0)} [16] \quad (8)$$

Di mana:

$\pi(1)$  = Probabilitas keberhasilan ketika variabel independen bernilai 1

$\pi(0)$  = Probabilitas keberhasilan ketika variabel independen bernilai 0

$\pi(1)/1 - \pi(1)$  = Odds dari kejadian saat  $X = 1$

$\pi(0)/1 - \pi(0)$  = Odds dari kejadian saat  $X = 0$

### 3. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk menganalisis faktor-faktor yang mempengaruhi Angka Harapan Hidup (AHH) di Papua dengan menggunakan regresi logistik biner. Data yang digunakan berasal dari 29 kabupaten/kota di Provinsi Papua, dengan 7 variabel, terdiri 1 variabel dependen (AHH) dan 6 variabel independen. Kategori AHH diklasifikasikan menjadi dua kelas, yaitu Tinggi dan Rendah, yang kemudian dikonversi menjadi variabel biner untuk keperluan pemodelan.

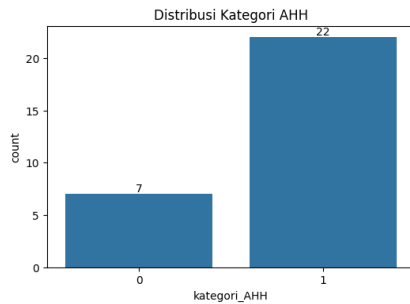
#### 3.1. Statistik Deskriptif

Analisis deskriptif digunakan untuk melihat pola umum serta distribusi data pada masing-masing variabel. Rangkuman hasil statistik deskriptif disajikan pada Tabel 1.

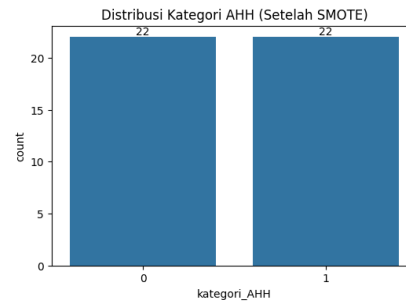
Tabel 2. Statistik Deskriptif

Variabel	Mean	Min	Max	Std. Dev
$x_1$	27.61	10.01	40.01	9.43
$x_2$	62.43	0.93	96.19	32.90
$x_3$	331.41	1.00	1517	360.41
$x_4$	19.75	5.00	40.00	10.96
$x_5$	8.58	0.00	25.81	8.54
$x_6$	48781	9598	256580	50384

Dari Tabel 2 dapat dilihat bahwa terdapat variasi cukup besar dalam jumlah tenaga medis dan PDRB antar kabupaten di Papua. Ini menunjukkan adanya ketimpangan dalam hal sumber daya kesehatan dan ekonomi yang berpotensi berpengaruh terhadap AHH.



**Gambar 1a**



**Gambar 1b**

**Gambar 1.** Distribusi Kategori AHH Sebelum *Oversampling* (1a) dan Setelah *Oversampling* (1b)

Dari gambar 1a menunjukkan distribusi kategori AHH yang tidak seimbang, dengan jumlah kategori tinggi (1) jauh lebih banyak dibandingkan kategori rendah (0). Ketidakeimbangan ini dapat mempengaruhi hasil analisis, sehingga digunakan metode SMOTE untuk menyeimbangkan data melalui oversampling pada kelas minoritas.

### 3.2. Estimasi Parameter

Setelah dilakukan penyeimbangan data menggunakan SMOTE, regresi logistik biner dilakukan untuk memodelkan kategori AHH berdasarkan variabel-variabel prediktor. Estimasi parameter awal dilakukan terhadap semua variabel.

**Tabel 3.** Estimasi Parameter

Variabel	Koefisien	P-value
Intercept	88.07	0.032
$x_1$	-1.43	0.034
$x_2$	-0.30	0.035
$x_3$	0.08	0.030
$x_4$	-1.03	0.032
$x_5$	-0.67	0.051
$x_6$	-0.0005	0.035

Berdasarkan Tabel 3, Sebagian besar variabel independen menunjukkan nilai *p-value* di bawah 0.05, kecuali variabel angka buta huruf ( $x_3$ ). Hal ini menunjukkan bahwa angka buta huruf ( $x_3$ ) tidak signifikan dan dengan demikian tidak digunakan dalam model.

### 3.3. Uji Signifikansi Model

Untuk menilai sejauh mana variabel independen dalam model mempengaruhi variabel dependen, pengujian simultan dan parsial dilakukan.

**Tabel 4.** Uji Simultan

Statistik	Nilai
$L_p$	-15.029
$L_o$	-30.498
G-test	0.000009627





Pengujian simultan dilakukan untuk menilai apakah kelima variabel independen secara kolektif memiliki pengaruh yang signifikan terhadap variabel dependen AHH. Dari hasil analisis, nilai *log likelihood* untuk model nol ( $L_0$ ) sebesar -30.498 dan untuk model lengkap ( $L_p$ ) sebesar -15.029. Nilai *G-test* yang dihasilkan memiliki *p-value* sebesar 0.000009627 yang menunjukkan pengaruh yang signifikan secara keseluruhan.

Karena nilai *p-value* yang diperoleh jauh di bawah batas signifikansi 0.05, maka hipotesis nol ( $H_0$ ) ditolak. Artinya, secara simultan variabel-variabel independen dalam model memiliki pengaruh yang signifikan terhadap AHH. Dengan demikian, model yang menyertakan prediktor lebih unggul dibandingkan model nol yang hanya memuat intercept.

Tabel 5. Uji Parsial

Variabel	Koefisien	<i>P-value</i>
$x_1$	-0.5948	0.010
$x_2$	-0.0909	0.025
$x_3$	0.0323	0.013
$x_4$	-0.4899	0.010
$x_6$	-0.0002	0.014

Berdasarkan Tabel 5, dilihat bahwa seluruh variabel independen yang digunakan dalam model akhir menunjukkan pengaruh yang signifikan secara parsial terhadap kategori AHH. Hal ini dibuktikan dengan nilai *p-value* masing-masing variabel yang kurang dari tingkat signifikansi 0.05.

- Persentase Penduduk Miskin mempunyai koefisien regresi negatif sebesar -0.5948 dengan *p-value* 0.010, menunjukkan bahwa peningkatan persentase penduduk miskin berkorelasi dengan penurunan peluang suatu wilayah untuk masuk dalam kategori AHH tinggi.
- Cakupan Imunisasi Lengkap juga berpengaruh negatif signifikan dengan koefisien -0.0909, *p-value* 0.025 ini mengindikasikan bahwa cakupan imunisasi yang rendah dapat berkaitan dengan kemungkinan lebih rendah untuk mencapai AHH.
- Jumlah Tenaga Medis memiliki koefisien positif sebesar 0.0323 dan *p-value* 0.013, menunjukkan bahwa semakin banyak tenaga medis yang tersedia di suatu wilayah, semakin besar pula kemungkinan wilayah itu termasuk dalam kategori AHH tinggi.
- Jumlah layanan kesehatan memiliki koefisien negatif sebesar -0.4899 dengan *p-value* 0.010, ini mengindikasikan bahwa peningkatan jumlah tenaga medis di suatu wilayah akan berkaitan dengan kemungkinan yang lebih rendah bagi wilayah itu untuk masuk ke dalam AHH tinggi.
- PDRB memiliki pengaruh negatif dengan koefisien -0.0002 dan *p-value* 0.014, meski pengaruhnya relatif kecil, hasil ini tetap signifikan, yang menunjukkan bahwa peningkatan PDRB tanpa pemerataan atau keadilan ekonomi mungkin tidak serta-merta mendorong peningkatan AHH.

### 3.4. Uji Kecocokan Model

Uji kecocokan model ini menggunakan uji *Hosmer-Lemeshow* digunakan menguji apakah model sesuai dengan data.

Tabel 6. Uji Kecocokan Model



<i>Hosmer-Lemeshow Test Statistic</i>	<i>P-value</i>
8.85	0.3550

Hasil uji Hosmer-Lemeshow menunjukkan nilai statistik sebesar 8.85 dengan *p-value* 0.3550. Karena nilai *p-value* lebih tinggi dari tingkat signifikansi 0.05, jadi hipotesis nol ( $H_0$ ) gagal ditolak. Hal ini menunjukkan bahwa model regresi logistik dapat dikatakan sesuai dengan data yang diamati.

### 3.5. Interpretasi Odds Ratio dan Confidence Interval

Untuk memahami pengaruh masing-masing variabel terhadap peluang suatu wilayah memiliki angka harapan hidup tinggi, maka dilakukan interpretasi berdasarkan nilai odds ratio dan hasil regresi logistik. Nilai odds ratio menunjukkan seberapa besar pengaruh perubahan satu unit dalam variabel independen terhadap kemungkinan perubahan AHH, dengan catatan bahwa semua variabel lain diasumsikan tidak berubah.

**Tabel 7.** Odds Ratio dan Confidence Interval

Variabel	Odds Ratio	CI 2.5%	CI 97.5%
$x_1$	0.55	0.35	0.87
$x_2$	0.91	0.84	0.99
$x_3$	1.03	1.02	1.06
$x_4$	0.61	0.42	0.89
$x_6$	1.00	0.9997	0.99997

Dari Tabel 7 diperoleh bahwa:

- Persentase Penduduk Miskin memiliki odds ratio sebesar 0.55 mengindikasikan bahwa peningkatan 1% pada persentase penduduk miskin akan menurunkan peluang suatu daerah masuk dalam AHH sebesar 45%. Hal ini menunjukkan bahwa kemiskinan memiliki dampak negatif terhadap kesehatan masyarakat secara umum. Confidence interval 97.5% untuk estimasi ini berada antara 0.35 hingga 0.87.
- Cakupan Imunisasi Lengkap menunjukkan odds ratio sebesar 0.91 mengindikasikan bahwa setiap peningkatan 1% cakupan imunisasi lengkap akan menurunkan peluang masuk AHH tinggi sebesar 9%. Meskipun terkesan kontra intuitif, hal ini bisa mencerminkan bahwa wilayah dengan cakupan imunisasi tinggi justru sedang mengejar ketertinggalan kesehatan atau menjadi prioritas program intervensi. Rentang kepercayaannya berada antara 0,84 hingga 0,99.
- Jumlah Tenaga Medis memiliki odds ratio sebesar 1,03 mengindikasikan bahwa setiap penambahan satu orang tenaga medis akan meningkatkan peluang suatu wilayah memiliki angka harapan hidup tinggi sebesar 3%. Rentang kepercayaan berada antara 1,01 hingga 1,06, menandakan hubungan ini cukup konsisten,
- Jumlah Layanan Kesehatan memiliki odds ratio sebesar 0,61 mengindikasikan bahwa menambahnya satu unit layanan kesehatan justru menurunkan peluang masuk kategori AHH tinggi sebesar 39%. Hasil ini dapat disebabkan oleh kemungkinan bahwa layanan kesehatan lebih banyak tersebar di daerah dengan tingkat kesehatan masyarakat yang rendah, sehingga keberadaannya merupakan respon terhadap masalah, bukan penanda kondisi baik.



- PDRB memiliki odds ratio sebesar 1,00 dengan confidence interval sangat sempit yaitu 0,9997 hingga 0,99997. Artinya, peningkatan satu satuan PDRB tidak memberikan perubahan peluang yang signifikan secara praktis, meskipun secara statistik tetap menunjukkan pengaruh.

Secara keseluruhan, odds ratio menunjukkan bahwa variabel-variabel sosial ekonomi dan kesehatan memiliki kontribusi penting dalam menjelaskan kemungkinan suatu daerah termasuk dalam AHH. Beberapa variabel seperti jumlah tenaga medis menunjukkan pengaruh positif, sedangkan kemiskinan dan cakupan imunisasi menunjukkan hubungan negatif yang mungkin dipengaruhi oleh konteks wilayah masing-masing.

Sehingga diperoleh persamaan akhir model regresi logistik:

$$\hat{g}(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$
$$\hat{g}(x) = 31.0689 - 0.5948x_1 - 0.0909x_2 + 0.0323x_3 - 0.4899x_4 - 0.0002x_6$$

Persamaan akhir model regresi logistik menunjukkan bahwa faktor sosial ekonomi dan kesehatan berpengaruh terhadap kemungkinan suatu wilayah memiliki AHH tinggi. Nilai intercept sebesar 31.0689 menjadi dasar perhitungan, meskipun tidak bermakna secara praktis.

Koefisien negatif pada persentase penduduk miskin dan cakupan imunisasi lengkap menunjukkan bahwa kenaikan pada kedua variabel tersebut cenderung menurunkan peluang suatu daerah memiliki AHH tinggi. Hal ini mengindikasikan bahwa kemiskinan dan fokus intervensi kesehatan di wilayah tertentu masih menjadi tantangan.

Jumlah tenaga medis menunjukkan pengaruh positif, artinya semakin banyak tenaga medis, semakin besar kemungkinan AHH tinggi. Sebaliknya, jumlah layanan kesehatan justru menunjukkan pengaruh negatif, yang mungkin mencerminkan respons terhadap kondisi kesehatan yang buruk di suatu wilayah.

PDRB memiliki pengaruh negatif yang kecil namun signifikan, menunjukkan bahwa pertumbuhan ekonomi belum tentu sejalan dengan peningkatan kualitas hidup atau pemerataan akses kesehatan. Secara keseluruhan, model ini menggambarkan bahwa peningkatan AHH tidak hanya bergantung pada satu faktor, tetapi merupakan hasil dari interaksi kompleks berbagai kondisi sosial dan kesehatan.

#### 4. KESIMPULAN

Dari proses analisis data dan pembahasan dalam penelitian ini diperoleh beberapa kesimpulan sebagai berikut:

1. Model regresi logistik biner yang diperoleh untuk memodelkan Angka Harapan Hidup (AHH) di Provinsi Papua tahun 2023 adalah sebagai berikut:

$$\hat{g}(x) = 31.0689 - 0.5948x_1 - 0.0909x_2 + 0.0323x_3 - 0.4899x_4 - 0.0002x_6$$

2. Beberapa variabel independen yang terbukti secara statistik mempunyai pengaruh yang signifikan terhadap Angka Harapan Hidup di Provinsi Papua Tahun 2023 antara lain:

- Persentase Penduduk Miskin ( $x_1$ ) berpengaruh negatif terhadap AHH.
- Cakupan Imunisasi Lengkap ( $x_2$ ) berpengaruh negatif terhadap AHH.
- Jumlah Tenaga Medis ( $x_3$ ) berpengaruh positif terhadap AHH.
- Jumlah Layanan Kesehatan ( $x_4$ ) berpengaruh negatif terhadap AHH.
- Produk Domestik Regional Bruto ( $x_6$ ) berpengaruh negatif terhadap AHH.

Sementara itu, Angka Buta Huruf ( $x_5$ ) merupakan satu-satunya variabel yang tidak mempunyai pengaruh signifikan terhadap hasil pemodelan, sehingga tidak dimasukkan dalam model akhir.



Secara umum, variabel-variabel yang berpengaruh positif seperti Jumlah Tenaga Medis ( $x_3$ ), menunjukkan bahwa peningkatan nilai tersebut meningkatkan peluang suatu wilayah untuk memilih Angka Harapan Hidup (AHH) yang tinggi. Sebaliknya, variabel-variabel yang berpengaruh negatif seperti Persentase Penduduk Miskin ( $x_1$ ), Cakupan Imunisasi Lengkap ( $x_2$ ), Jumlah Layanan Kesehatan ( $x_4$ ), Produk Domestik Regional Bruto ( $x_6$ ) menunjukkan bahwa peningkatan nilai pada variabel-variabel itu berkaitan dengan menurunkan peluang suatu wilayah untuk memiliki AHH tinggi. Arah pengaruh ini menunjukkan bahwa tantangan sosial ekonomi dan distribusi layanan kesehatan di Papua masih menjadi faktor penting dalam pencapaian kualitas hidup masyarakat.

## REFERENSI

- [1] Badan Pusat Statistik Kabupaten Temanggung, “Angka Harapan Hidup (AHH),” BPS Kabupaten Temanggung. [Online]. Available: <https://temanggungkab.bps.go.id/id/news/2025/01/22/564/angka-harapan-hidup--ahh-.html>
- [2] Badan Pusat Statistik, “Angka Harapan Hidup (AHH) Menurut Provinsi dan Jenis Kelamin,” Badan Pusat Statistik. [Online]. Available: <https://www.bps.go.id/id/statistics-table/2/NTAxIzI=/angka-harapan-hidup-ahh-menurut-provinsi-dan-jenis-kelamin.html>
- [3] A. Septianingsih, “Pemodelan Data Panel Menggunakan Random Effect Model Untuk Mengetahui Faktor Yang Mempengaruhi Umur Harapan Hidup Di Indonesia,” *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 3, no. 3, pp. 525–536, 2022, doi: 10.46306/lb.v3i3.163.
- [4] A. Wardhana and B. Kharisma, “Angka Harapan Hidup Di Era Desentralisasi (Kabupaten/Kota Jawa Barat),” *E-Jurnal Ekon. dan Bisnis Univ. Udayana*, no. June, p. 405, 2020, doi: 10.24843/eeb.2020.v09.i05.p02.
- [5] N. Amalia and Mahmudah, “Faktor Yang Mempengaruhi Angka Harapan Hidup Di Provinsi Jawa Timur Tahun 2014 Dengan Melihat Nilai Statistik Cp Mallows Factors That Affecting Life Expectancy In East Java 2014 With Cp Mallows Statistics,” *J. Wiyata*, vol. 2014, no. 1, pp. 13–19, 2019.
- [6] A. P. Sugiantari and I. N. Budiantara, “Analisis faktor-faktor yang mempengaruhi angka harapan hidup di Jawa Timur menggunakan regresi semiparametrik spline,” *J. Sains dan Seni Pomits*, vol. 2, no. 1, pp. D37–D41, 2013.
- [7] F. Pratiwi and B. Budyanra, “Analisis Determinan Angka Harapan Hidup Di Provinsi Maluku Tahun 2015-2017,” *Semin. Nas. Off. Stat.*, vol. 2019, no. 1, pp. 602–607, 2020, doi: 10.34123/semnasoffstat.v2019i1.25.
- [8] A. Chumairah, “ANALISIS REGRESI LOGISTIK BINER UNTUK MENENTUKAN FAKTOR-FAKTOR YANG MEMPENGARUHI ANGKA HARAPAN HIDUP DI INDONESIA,” 2022.
- [9] E. Anggraini and U. Lisyarningsih, “Disparitas Spasial Angka Harapan Hidup Di Indonesia Tahun 2010,” *J. Bumi Indones.*, vol. 2, pp. 71–80, 2013.
- [10] W. Alwi, E. Ermawati, and S. Husain, “Analisis Regresi Logistik Biner Untuk Memprediksi Kepuasan Pengunjung Pada Rumah Sakit Umum Daerah Majene,” *J. MSA (Matematika dan Stat. serta Apl.)*, vol. 6, no. 1, p. 20, 2018, doi: 10.24252/msa.v6i1.4783.
- [11] N. A. N. Roosyidah and P. K. Supriyatna, “Pemodelan Regresi Logistik untuk Diagnosis Dini Infeksi Covid-19 di Indonesia,” *Jambura J. Math.*, vol. 4, no. 2, pp. 232–246, 2022, doi: 10.34312/jjom.v4i2.12653.
- [12] L. D. Perwara, F. A. Bachtiar, and Indriati, “Penerapan Algoritma Decision Tree C4. 5 Untuk Deteksi Fraud Pada Kartu Kredit dengan Oversampling Synthetic Minority Technique (SMOTE),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 8, pp. 2664–2669, 2020, [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/download/7745/3649>
- [13] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. New York: John Wiley & Sons, Inc., 2013.



# SENADA

Seminar Nasional Sains Data 2025 (SENADA 2025)

E-ISSN 2808-5841

UPN “Veteran” Jawa Timur

P-ISSN 2808-7283

- [14] E. Roflin, F. Riana, E. Munarsih, Pariyana, and I. A. Liberty, *Regresi Logistik Biner dan Multinomial*. NEM, 2023.
- [15] M. K. Kotimah and S. P. Wulandari, “Model Regresi Logistik Biner Stratifikasi pada Partisipasi Ekonomi Perempuan di Provinsi Jawa Timur,” *J. Sains Dan Seni Pomits*, vol. 3, no. 1, pp. D1–D6, 2014.
- [16] R. S. Susanti, S. D. Anugrawati, Fitrah, J. Usman, and Yusrianto, “Analisis Faktor Risiko Penyebab Diabetes Melitus dengan Menggunakan Regresi Logistik Biner,” *J. MSA (Mat. dan Stat. serta Apl.*, vol. 11, no. 2, pp. 37–45, 2023, doi: 10.24252/msa.v11i2.41051.