

## The 2<sup>nd</sup> FORSTAT International Conference



# The Effect of Socioeconomic Indicators on Environmental Quality in Indonesia Using Logistic Regression

*Melani Yusi Aryanda*<sup>a,\*</sup>, *Luthfiya Zuhura Syifa Fuadah*<sup>b</sup>, *Annisa Zahrotu Firda Asfari*<sup>c</sup>,  
*Rizky Kusumawardani*<sup>d</sup>

<sup>a, b, c, d</sup> Raden Mas Said State Islamic University Surakarta, Indonesia

\*Corresponding author: [melaniyusiaryanda@gmail.com](mailto:melaniyusiaryanda@gmail.com)

### ABSTRACT

Environmental degradation in Indonesia is becoming increasingly evident as development activities and the exploitation of natural resources intensify. Deforestation resulting from the expansion of plantations, mining, and industrial development leads to the loss of forest cover as well as a decline in air quality and biodiversity. These conditions indicate that rapid economic growth often places significant pressure on environmental sustainability if not balanced by sustainable management. This study aims to analyze the relationship between socioeconomic factors and the risk of environmental degradation in Indonesia in 2024. The study employs a logistic regression model using secondary data from 38 provinces sourced from the Central Statistics Agency (BPS) and the Ministry of Environment and Forestry (KLHK). Independent variables include per capita GRDP, poverty rate, Human Development Index, and population density, while the dependent variable is environmental quality measured through the Environmental Quality Index classified based on the national median. The analysis results indicate that the model is statistically significant simultaneously and can moderately explain variations in environmental quality. Partially, the poverty rate variable shows a significant negative effect on environmental quality at the 10% significance level. Meanwhile, the per capita GRDP, Human Development Index, and population density variables do not show a significant effect. These findings indicate that the risk of environmental quality decline is not solely determined by the level of prosperity but is influenced by structural dynamics and inter-regional development pressures.

**Keywords:** Environmental Degradation; Logistic Regression; Environmental Quality Index (IKLH); Sustainable Development; Socio-Economic Inequality

## 1. Introduction

Environmental quality is a key indicator in assessing a nation's development success. In Indonesia, environmental pressures are intensifying alongside accelerated economic growth, the expansion of industrial and mining activities, and massive land conversion. Deforestation, air and water pollution, and the decline in biodiversity have become increasingly alarming environmental issues. The Central Statistics Agency (BPS) and the Ministry of Environment and Forestry (KLHK) note that the Environmental Quality Index (IKLH) values across provinces in Indonesia show significant variation. Based on 2024 data, EQI values range from 56.39 to 83.75, indicating a clear disparity in environmental conditions across various regions of Indonesia [1][2].



One factor suspected of influencing environmental quality is the socioeconomic condition of the community. The Environmental Kuznets Curve (EKC) hypothesis states that in the early stages of development, economic growth tends to worsen environmental conditions; however, after passing a certain point, rising income actually drives improvements in environmental quality. Nevertheless, the relevance of this hypothesis at the regional level in Indonesia still requires more in-depth empirical research, given the high heterogeneity of economic and environmental conditions across provinces [3].

Socioeconomic variables such as Gross Regional Domestic Product (GRDP) per capita, poverty rates, the Human Development Index (HDI), and population density are believed to be associated with the environmental conditions of a region [4]. Previous studies have examined the relationship between these variables and environmental quality, but the results have been inconsistent and remain limited to specific contexts. A comprehensive study is needed that employs a statistical approach appropriate to the data characteristics, particularly when the response variable is categorical [5].

This study aims to analyze the influence of socioeconomic indicators—including per capita GRDP, poverty rates, HDI, and population density—on environmental quality in Indonesia using a binary logistic regression model. The data used covers 38 provinces in Indonesia in 2024, sourced from the Central Statistics Agency (BPS) and the Ministry of Environment and Forestry (KLHK). The research results are expected to provide empirical contributions to the formulation of environmentally conscious sustainable development policies in Indonesia.

## 2. Theoretical Framework

### 2.1. Environmental Kuznets Curve (EKC)

The Environmental Kuznets Curve (EKC) is a hypothesis describing an inverted U-shaped relationship between economic growth and environmental degradation [6]. In the early stages of development, increases in per capita income tend to be accompanied by increased pressure on the environment due to industrial expansion and energy use. However, once income reaches a certain threshold, society has greater capacity and awareness to invest in clean technologies and environmental protection. This hypothesis serves as the theoretical foundation for examining the relationship between per capita GRDP and environmental quality at the provincial level in Indonesia.

### 2.2. Environmental Quality Index (IKLH)

The Environmental Quality Index (EQI) is a composite indicator used by the Indonesian government to measure environmental conditions in a given region. The EQI is composed of three main components: the Water Quality Index (WQI), the Air Quality Index (AQI), and the Land Cover Index (LCI) [7]. The EQI ranges from 0 to 100, with higher values reflecting better environmental conditions. In this study, the EQI is used as a dependent variable categorized binarily based on the national median value of 75.04.

### 2.3. Binary Logistic Regression

Binary logistic regression is a statistical analysis method used to model the relationship between a dichotomous dependent variable and one or more independent variables [8]. This model uses the logit function to estimate the probability of an event and does not assume a linear relationship between the dependent and independent variables, nor does it require normality of the residuals. The binary logistic regression model is mathematically expressed as:

$$\ln(\pi/(1-\pi)) = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4$$

where  $\pi$  is the probability that environmental quality is below the national median ( $y=1$ ),  $\beta_0$  is the constant,  $\beta_1$  through  $\beta_4$  are the regression coefficients for each independent variable,  $X_1$  is per capita GRDP,  $X_2$  is the poverty rate,  $X_3$  is the HDI, and  $X_4$  is population density. The model coefficients are estimated using the Maximum Likelihood Estimation (MLE) method.

### 2.4. Socioeconomic Factors and Environmental Quality

Several socioeconomic variables have been extensively studied for their relationship with environmental quality. Per capita GRDP reflects a region's level of economic prosperity, which is linked to its capacity for investment and environmental management [9]. The poverty rate is closely associated

with the pressure of natural resource exploitation by low-income communities who depend on the environment as a source of livelihood [10]. The Human Development Index (HDI), as a measure of human development, reflects levels of education and health that influence behavior and environmental awareness [11]. Meanwhile, population density reflects the intensity of pressure on a region's environmental carrying capacity [12][13].

### 3. Methods

#### 3.1. Types and Sources of Data

This study uses secondary data with a descriptive quantitative approach. The data used consists of cross-sectional data from 38 provinces in Indonesia for the year 2024. The data sources come from the Central Statistics Agency (BPS) for per capita GRDP, poverty rates, HDI, and population density, as well as Environmental Quality Index (IKLH) data, which includes the Water Quality Index (IKA), Air Quality Index (IKU), and Land Quality Index (ITL)[14][15][16][17][1][2].

#### 3.2. Research Variables

The dependent variable in this study is environmental quality (Y), measured via the EQI and categorized binarily based on the national median (75.04):  $Y = 1$  if a province's EQI value is below or equal to the median (low/poor environmental quality), and  $Y = 0$  if the EQI value is above the median (good environmental quality). The categorization results in a balanced distribution, with 19 provinces in each category. The independent variables consist of four socioeconomic indicators, namely: (1) GRDP per capita ( $X_1$ ) in thousands of rupiah, representing the region's level of economic prosperity; (2) poverty rate ( $X_2$ ) in percent, reflecting the proportion of the poor population; (3) Human Development Index ( $X_3$ ) on a scale of 0–100, which measures human development achievements including health, education, and decent living standards; and (4) population density ( $X_4$ ) in people per  $\text{km}^2$ , reflecting the intensity of land use.

The independent variables used in this study are socioeconomic indicators believed to influence environmental quality. Per capita GRDP is used to represent the level of regional economic prosperity. Increased economic activity can place pressure on the environment due to rising production, consumption, and exploitation of natural resources, thereby potentially affecting environmental quality [18]. Poverty levels are used to describe community welfare conditions because communities with low economic status tend to rely more heavily on the use of natural resources to meet their basic needs, thereby potentially increasing pressure on the environment [19]. The Human Development Index (HDI) reflects the quality of human development, encompassing the dimensions of health, education, and a decent standard of living, thereby influencing public awareness regarding the sustainable management and conservation of the environment [20]. Meanwhile, population density is used to describe the pressure of human activities on the environment; the higher the population density, the greater the potential for environmental degradation due to increased demand for land, energy, and natural resources [21].

#### 3.3. Analysis Techniques

Data analysis in this study employs a binary logistic regression model using Python software with the statsmodels library. Logistic regression was chosen because this method is suitable for analyzing the relationship between independent variables and a categorical dependent variable—one classified into two categories. In this study, the dependent variable—the Environmental Quality Index (IKLH)—is categorized into two groups, making the logistic regression model the appropriate method for identifying the influence of socioeconomic indicators on environmental quality [22].

Logistic regression is widely used in social and economic research because it can estimate the probability of an event based on explanatory variables and does not require the assumption of normality in the independent variables, as in linear regression [23]. Additionally, this model can explain the effect of each independent variable on the probability of an event occurring through the interpretation of the odds ratio, thereby helping researchers understand the direction and magnitude of the variables' effects [24].

Data analysis was conducted using a binary logistic regression model with the assistance of Python software (statsmodels library). The analysis steps included: (1) descriptive statistical analysis to describe the characteristics of each variable, (2) a simultaneous test using the Likelihood Ratio Test (LRT) to

assess the overall significance of the model, (3) partial tests using the Wald (z) test statistic to assess the significance of each independent variable, (4) model fit assessment using the Pseudo R-squared (McFadden) value, and (5) interpretation of the odds ratio to determine the direction and magnitude of the effect of each variable. The significance level used was  $\alpha = 10\%$ .

## 4. Results and Discussion

### 4.1. Descriptive Statistics

Descriptive statistical analysis was conducted to describe the characteristics of each research variable from 38 provinces in Indonesia in 2024. The complete results are presented in Table 1.

**Table 1.** Descriptive Statistics of Research Variables

Variable	Min	Max	Mean	Standard Deviation
IKLH	56.39	83.75	75.11	5.56
GRDP per capita (thousand IDR)	18,105	344,350	84,344	62,197
Poverty Rate (%)	4.00	32.97	11.15	6.75
HDI	53.42	83.08	72.38	5.15
Population Density (people/km <sup>2</sup> )	5	16,165	678	2,609.45

Based on Table 1, the IKLH values across provinces in Indonesia range from 56.39 to 83.75, with an average of 75.11 and a standard deviation of 5.56. The lowest IKLH value is found in DKI Jakarta (56.39), while the highest is in West Papua (83.75). This range of values reflects significant variations in environmental conditions across regions. Per capita GRDP shows a very high disparity, with the lowest value at Rp18,105,000 (Mountainous Papua) and the highest at Rp344,350,000 (DKI Jakarta), indicating significant economic inequality among provinces.

Poverty rates range from 4.00% (Bali) to 32.97% (Mountainous Papua) with an average of 11.15%, reflecting high socioeconomic heterogeneity. Population density shows the most extreme variation, ranging from 5 people/km<sup>2</sup> (South Papua) to 16,165 people/km<sup>2</sup> (DKI Jakarta), which starkly reflects the differences in regional characteristics between Java and non-Java regions.

### 4.2. Distribution of the Dependent Variable

The dependent variable was categorized binarily based on the national IKLH median value of 75.04. The categorization results are presented in Table 2.

**Table 2.** Distribution of Dependent Variable Categories (Y)

Category (Y)	Description	Number of Provinces	Percentage
0	Good Environmental Quality (IKLH > 75.04)	19	50.00
1	Poor Environmental Quality (IKLH ≤ 75.04)	19	50.00
<b>Total</b>		<b>38</b>	<b>100.00</b>

Table 2 shows that the distribution of the dependent variable is balanced, with 19 provinces each in the good environmental quality category (Y=0) and the poor environmental quality category (Y=1). Provinces with good environmental quality (IKLH above the median) are generally located in the regions of Sulawesi, Maluku, and Papua, which still have relatively extensive forest cover. Conversely, provinces with poor environmental quality are mostly located on the islands of Java, Sumatra, and Bali, which face higher development pressures and population density.

### 4.3. Results of the Simultaneous Test (Likelihood Ratio Test)

A simultaneous test was conducted to evaluate whether all independent variables collectively have a significant effect on the dependent variable. The hypothesis tested was  $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$  (no variable has a significant effect). The test results are presented in Table 4.

**Table 4.** Measures of Goodness of Fit for the Logistic Regression Model

Measure	Value
Number of Observations	38
Model Log-Likelihood	-16.568
Log-Likelihood of the Null Hypothesis (LL-Null)	-26.340
Pseudo R-squared (McFadden)	0.3710
LLR p-value (Simultaneous Test)	0.0006143
Convergence Iteration	11

Based on Table 4, the LLR p-value of 0.0006143 is far below the 10% significance level ( $\alpha = 0.10$ ), so  $H_0$  is rejected. This means that, simultaneously, there is at least one independent variable that has a significant effect on the quality of the environment in Indonesia. The Pseudo R-squared (McFadden) value of 0.3710 indicates that the model explains approximately 37.10% of the variation in environmental quality. This value is classified as moderate-to-good for a logistic regression model, suggesting that the four socioeconomic variables used are sufficiently relevant in explaining differences in environmental quality across provinces.

#### 4.4. Results of Partial Tests and Interpretation of Coefficients

Partial testing was conducted using the Wald (z) test statistic to assess the significance of each independent variable individually. The complete results are presented in Table 3.

**Table 3.** Estimation Results of the Binary Logistic Regression Model

Variable	Coefficient ( $\beta$ )	Std. Error	z-Statistic	P >  z	Odds Ratio
Constant ( $\beta_0$ )	2.9735	2.771	1.073	0.283	19.561
GRDP per capita ( $X_1$ )	-0.9489	0.792	-1.199	0.231	0.387
Poverty Rate ( $X_2$ )	-1.3418	0.806	-1.664	0.096	0.261
IPM ( $X_3$ )	-0.5313	1.071	-0.496	0.620	0.588
Population Density ( $X_4$ )	16.2917	12.434	1.310	0.190	11,895,390

The per capita GRDP variable ( $X_1$ ) has a coefficient of -0.9489 with a z-value of -1.199 and a p-value of 0.231, which exceeds the level 10% significance. Therefore, per capita GRDP does not have a significant effect on environmental quality. An odds ratio of 0.387 indicates that an increase in per capita GRDP tends to reduce the likelihood of a province having poor environmental quality, although the effect is not statistically significant. This finding does not fully support the EKC hypothesis that linear economic prosperity correlates with environmental improvement at the inter-provincial scale in Indonesia [6][25]

The poverty rate variable ( $X_2$ ) has a coefficient of -1.3418 with a z-value of -1.664 and a p-value of 0.096. Since the p-value (0.096) is smaller than the 10% significance level ( $\alpha=0.10$ ), the poverty rate variable is proven to have a negative and statistically significant partial effect on environmental quality. An odds ratio of 0.261 indicates that provinces with higher poverty levels have a lower likelihood of being categorized as having poor environmental quality, an interesting finding. This is likely due to the fact that provinces with high poverty, such as the Papua region, still have extensive forest cover and minimal industrial pressure, unlike wealthier provinces that face more intensive development pressures [10].

The HDI variable ( $X_3$ ) has a coefficient of -0.5313 with a z-value of -0.496 and a p-value of 0.620, far from significant at the 10% level. An odds ratio of 0.588 indicates a negative but statistically insignificant relationship between the HDI and the likelihood of poor environmental quality. This suggests that human development achievements at the provincial level have not yet been directly reflected in improvements in environmental quality, likely because the HDI more closely reflects aspects of health and education, whose impacts on the environment are long-term and indirect.

The population density variable ( $X_4$ ) has the largest coefficient, namely 16.2917, with a z-value of 1.310 and a p-value of 0.190. Although the coefficient is positive (increasing the likelihood of poor environmental quality), its effect is not statistically significant. The extremely large odds ratio (11,895,390) occurs because the independent variables underwent standardization (Z-score) during

modeling. This means that an increase in population density of 1 standard deviation (approximately 2,609 people/km<sup>2</sup>) will exponentially multiply the probability that a province has poor environmental quality. This makes perfect sense given the extreme variability in population density across provinces in Indonesia.

Overall, the findings of this study indicate that at the 10% significance level, the poverty rate is the only socioeconomic variable that has been individually proven to have a significant effect on environmental quality across provinces in Indonesia. However, the model proved significant when analyzed simultaneously, and the poverty rate showed the strongest influence compared to other variables. This confirms that the dynamics of environmental quality in Indonesia cannot be simply explained through socioeconomic indicators alone but are also influenced by geographical factors, natural resource characteristics, and varying regional development patterns [9][12].

A situation where the simultaneous test (Likelihood Ratio Test) is significant while the partial test (Wald test) yields only one significant variable is common in logistic regression. This can occur due to correlation among independent variables (multicollinearity) or *shared variance*, such that these variables collectively have a significant effect on the model but are not significant individually. Additionally, the Wald test tends to be less robust at small sample sizes compared to the Likelihood Ratio test, which evaluates the model as a whole [22]. In the context of this study, the poverty rate accounts for the largest proportion of variance because this variable directly represents the condition of regional development lag, such as in the Papua region, which has extensive forest cover and relatively low industrial pressure compared to regions with high levels of prosperity.

## 5. Conclusion

This study examines the influence of socioeconomic indicators on environmental quality in Indonesia using a binary logistic regression model with data from 38 provinces in 2024. The dependent variable was categorized based on the national IKLH median value of 75.04, resulting in a balanced distribution of 19 provinces in each category.

Simultaneous test results indicate that the resulting logistic regression model is statistically significant (LLR p-value = 0.0006), explaining 37.10% of the variation in environmental quality based on the McFadden Pseudo R-squared value. However, at the 10% significance level, the poverty rate (p=0.096) was found to have a significant negative effect. Meanwhile, the other independent variables—namely per capita GRDP (p=0.231), HDI (p=0.620), and population density (p=0.190)—were not found to be significant.

These findings confirm that although the poverty rate is a significant predictor, the quality of the living environment in Indonesia cannot be fully explained solely through basic indicators such as economic prosperity, human resource quality, or population density alone. Structural dynamics and geographical characteristics across regions are factors that interact and are often more determinative.

## REFERENCES

- [1] Badan Pusat Statistik, **Statistik Lingkungan Hidup Indonesia 2024**. Jakarta: BPS, 2024.
- [2] Kementerian Lingkungan Hidup dan Kehutanan, **Laporan Kinerja Kementerian Lingkungan Hidup dan Kehutanan 2023**. Jakarta: KLHK, 2024.
- [3] A. Setiawan and M. Anwar, “The Environmental Kuznets Curve Hypothesis: An Empirical Evidence in Indonesia,” **JEJAK: Jurnal Ekonomi dan Kebijakan**, vol. 15, no. 1, pp. 207–220, 2022. [Online]. Available: <https://doi.org/10.15294/jejak.v15i1.38084>
- [4] A. B. Pertiwi, A. H. Juwita, and S. Suryanto, “Effects of Poverty, Income Inequality and Economic Growth to Environmental Quality Index (EQI) in 33 Province in Indonesia 2014–2019,” **Ekuilibrum: Jurnal Ilmiah Bidang Ilmu Ekonomi**, vol. 16, no. 2, pp. 154–163, 2021. [Online]. Available: <https://doi.org/10.24269/ekuilibrum.v16i2.2021.pp154-163>
- [5] H. Jonatan and R. Salam, “Application of Panel Data Regression: Determinants of Environmental Quality in Indonesia, 2019–2023,” **Seminar Nasional Official Statistics**, vol. 2025, no. 1, pp. 571–580, 2025. [Online]. Available: <https://doi.org/10.34123/semnasoffstat.v2025i1.2471>

- [6] G. M. Grossman and A. B. Krueger, "Economic Growth and the Environment," **The Quarterly Journal of Economics**, vol. 110, no. 2, pp. 353–377, 1995. [Online]. Available: <https://doi.org/10.2307/2118443>
- [7] Kementerian Lingkungan Hidup dan Kehutanan, **Peraturan Menteri Lingkungan Hidup dan Kehutanan Nomor 27 Tahun 2021 tentang Indeks Kualitas Lingkungan Hidup**. Jakarta: KLHK, 2021.
- [8] D. Dey, M. S. Haque, M. M. Islam, et al., "The proper application of logistic regression model in complex survey data: A systematic review," **BMC Medical Research Methodology**, vol. 25, 2025. [Online]. Available: <https://doi.org/10.1186/s12874-024-02454-5>
- [9] W. S. I. Soeparno, Y. Anggia, R. L. Sari, and F. A. Purba, "Economic Growth, Population Density, and Human Development: Key Drivers of Environmental Quality in Indonesia," **International Journal on Economics, Finance and Sustainable Development**, vol. 6, no. 3, pp. 394–401, 2024. [Online]. Available: <https://doi.org/10.31149/ijefsd.v6i7.5310>
- [10] O. Shanty, W. P. Dita, Firmansyah, and F. X. Sugiyanto, "The Relationship between Environmental Degradation, Poverty and Human Quality in Indonesia," **E3S Web of Conferences**, vol. 73, Art. no. 10020, 2018. [Online]. Available: <https://doi.org/10.1051/e3sconf/20187310020>
- [11] S. S. Ramadhantie, M. J. Ramadhan, and M. A. Hasibuan, "Pengaruh Indeks Pembangunan Manusia terhadap Indeks Kualitas Lingkungan Hidup di Indonesia Menggunakan Regresi Data Panel," **Ekologia: Jurnal Ilmiah Ilmu Dasar dan Lingkungan Hidup**, vol. 21, no. 1, pp. 35–43, 2021. [Online]. Available: <https://doi.org/10.33751/ekologia.v21i1.2111>
- [12] H. A. Zahro and T. Tutik, "Determinan Indeks Kualitas Lingkungan Hidup: Studi Kasus Delapan Provinsi di Indonesia," **Journal of Economics and Management Sciences**, vol. 7, no. 4, pp. 678–687, 2025. [Online]. Available: <https://doi.org/10.37034/jems.v7i4.206>
- [13] N. Finanda and T. Gunarto, "Analisis Pengaruh Pertumbuhan Ekonomi, Pertumbuhan Penduduk, serta Tingkat Kemiskinan terhadap Indeks Kualitas Lingkungan Hidup," **Jurnal Sosial dan Sains**, vol. 2, no. 1, pp. 193–202, 2022. [Online]. Available: <https://doi.org/10.36418/sosains.v2i1.324>
- [14] Badan Pusat Statistik, "Produk Domestik Regional Bruto per Kapita Menurut Provinsi, 2024," [Online]. Available: <https://www.bps.go.id>
- [15] Badan Pusat Statistik, "Jumlah dan Persentase Penduduk Miskin Menurut Provinsi, 2024," [Online]. Available: <https://www.bps.go.id>
- [16] Badan Pusat Statistik, "Metode Baru Indeks Pembangunan Manusia (IPM) menurut Provinsi, 2024," [Online]. Available: <https://www.bps.go.id>
- [17] Badan Pusat Statistik, "Penduduk, Laju Pertumbuhan Penduduk, Distribusi Persentase Penduduk, Kepadatan Penduduk, Rasio Jenis Kelamin Penduduk Menurut Provinsi, 2024," [Online]. Available: <https://www.bps.go.id>
- [18] Y. Navila and E. Cerya, "Pengaruh pertumbuhan ekonomi, kepadatan penduduk dan kemiskinan terhadap kualitas lingkungan hidup di Indonesia," **Jurnal Pendidikan Lingkungan dan Pembangunan**, vol. 26, no. 1, pp. 15–26, 2025. [Online]. Available: <https://doi.org/10.21009/plpb.v26i01.59043>
- [19] A. Yani, R. Restiatun, and N. Nuratika, "Indeks kualitas lingkungan hidup dan determinannya: Studi kasus di Indonesia," **Jurnal Ekonomi Pembangunan**, vol. 12, no. 3, 2023. [Online]. Available: <https://doi.org/10.23960/jep.v12i3.2132>
- [20] Badan Pusat Statistik, **Statistik Lingkungan Hidup Indonesia 2021**. Jakarta: BPS, 2021. [Online]. Available: <https://www.bps.go.id/id/publication/2021/11/30/2639657be1e8bd2548469f0f/statistik-lingkungan-hidup-indonesia-2021.html>

- [21] Y. B. Kondolele and B. Mustari, “Faktor penentu kualitas lingkungan hidup pada pusat populasi Indonesia,” **Jurnal Ekonomika dan Dinamika Sosial**, 2024. [Online]. Available: <https://journal.unhas.ac.id/index.php/jeds/article/view/27503>
- [22] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, **Applied Logistic Regression**, 3rd ed. Hoboken, NJ, USA: Wiley, 2022. [Online]. Available: <https://doi.org/10.1002/9781118548387>
- [23] G. James, D. Witten, T. Hastie, and R. Tibshirani, **An Introduction to Statistical Learning with Applications in Python**. New York, NY, USA: Springer, 2021. [Online]. Available: <https://www.statlearning.com>
- [24] S. Seabold and J. Perktold, “Statsmodels: Econometric and Statistical Modeling with Python,” 2021. [Online]. Available: <https://www.statsmodels.org/stable/index.html>
- [25] D. Adila, N. Nuryartono, and M. Oak, “The Environmental Kuznets Curve for Deforestation in Indonesia,” **Economics and Finance in Indonesia**, vol. 67, no. 2, 2021. [Online]. Available: <https://scholarhub.ui.ac.id/cgi/viewcontent.cgi?article=1018&context=efi>