Econometric Analysis of Poverty and Economic Welfare as Predictors of Criminal Activity in Indonesia

Kurnia Qurrotul A'yun[1)], Azzah Sumaiyah[2)], Nabilla Hannah Wardani[3)], Shinta Devi Arsy Zelika[4)], Arip Ramadan[*5)], Alhassan Sesay[6)]

[1, 2, 3, 4, 5)] School of Industrial and System Engineering, Telkom University, Surabaya Campus, Surabaya, Indonesia

[6)] Faculty of Transformative Education, the United Methodist University, Freetown, Sierra Leone

[1]kurniaqurrotulayun@student.telkomuniversity.ac.id, [2]azzahsumaiyah@student.telkomuniversity.ac.id, [3]nabillahannahwardani@student.telkomuniversity.ac.id, [4]shintadeviarsyzelika@student.telkomuniversity.ac.id, [5*]aripramadan@telkomuniversity.ac.id, [6]alhassansesay3@gmail.com

ARTICLE INFO

Article history:
Received 20 April 2026
Revised 9 May 2026
Accepted 25 May 2026
Available online 30 May 2026

Keywords:
Crime Rate
GRDP
HDI
Poverty
Regression

ABSTRACT

Criminal activity is defined as a violation of legal mandates and societal standards that triggers communal instability and economic detriment. To address this challenge, identifying the core determinants of such behavior is essential, which this study achieves through a multiple regression methodology. The empirical results demonstrate that Gross Domestic Product (GDP), poverty demographics, per capita spending, and the Human Development Index (HDI) collectively exert a significant influence on national crime rates. This integrated model accounts for 0,775 of the observed variances, as evidenced by the R^2 value. Individually, higher levels of GDP, poverty, and HDI correlate with increased criminality, whereas improved per capita expenditure serves as a restrictive factor. These outcomes underscore the necessity of prioritizing welfare enhancement and poverty alleviation within national crime prevention frameworks.

1. Introduction

Criminal acts constitute a societal anomaly characterized by the breach of established laws and norms, subsequently resulting in tangible losses, distress, and communal instability. The persistence of illegal activities inevitably diminishes public security and disrupts the rhythm of daily life [1][2][3]. The catalysts for such unlawful behaviors are multifaceted, encompassing a spectrum of biological, sociological, and economic determinants [4]. Among these variables, economic circumstances exert a profound influence on the escalation of criminality, especially in scenarios where fundamental human necessities are deprived [5][6]. The capacity of a population to fulfill its rudimentary needs is often indicated by per capita expenditure. When this financial metric is robust, it alleviates the social friction that traditionally serves as a root cause of deviance [7]. Consequently, consistent public spending mirrors the adequate provision of basic requirements, acting as a direct deterrent to delinquent behavior. Conversely, disparities in economic resource distribution frequently breed social volatility. Periods of economic recession are routinely accompanied by a surge in crime, as illegal actions are adopted as alternative survival mechanisms. Elevated financial hardship exacerbates poverty levels, and the restriction of legitimate economic avenues compels certain demographics to engage in unlawful enterprises [8]. This dynamic underscores that the frequency of crime is heavily dictated by grassroots economic fragility. A stark poverty gap is primarily driven by inequitable access to wealth. Whenever communities lack the financial capability to comprehensively satisfy their essential needs, the vulnerability to local security breaches intensifies. To measure the overarching economic prosperity and regional growth, Gross Domestic Product (GDP) is frequently utilized as a standard metric [4][9].

Beyond sheer economic output, the Human Development Index (HDI) significantly governs the trajectory of crime [10]. The influence of HDI is channeled through its three foundational pillars: educational attainment, health quality, and overall living standards. A depressed HDI signifies a systemic breakdown in the delivery of essential services, an outcome that breeds societal disillusionment and erodes collective moral restraints [9].

The presence of elevated crime rates paralyzes communal peace, hindering individuals from conducting their routine endeavors without fear. Consequently, mitigating this societal challenge necessitates rigorous and sustained intervention. While the existing body of literature examining criminality is extensive, the majority of prior investigations confine their scope to localized jurisdictions, such as specific municipalities or regencies. To address this spatial limitation, the current research is structured to evaluate how GDP, HDI, per capita expenditure, and poverty demographics collectively dictate the national crime landscape across Indonesia in the year 2024.

2. Methods

This research utilizes secondary datasets obtained from Indonesia's Central Bureau of Statistics (BPS) official records. The scope of the analysis encompasses 2024 cross-provincial statistics, integrating a diverse range of socio-economic parameters.

Table 1. Descriptive Statistics Analysis

Variable	Definition	Data Type	Measurement Scale
Crime	Number of criminal incidents in each province in 2024	Numeric	Ratio
ln_GDP	Natural logarithm of Gross Regional Domestic Product per province	Numeric	Ratio
Poor Population	Number of poor people in each province	Numeric	Ratio
Expenditure	Per capita expenditure in each province	Numeric	Ratio
HDI	Human Development Index, including education, health, and living standards	Numeric	Ratio

Data transformation was performed on the GDP variable, which was subsequently used in regression testing. The obtained data were analyzed using descriptive statistics to summarize the general trends and describe each variable, providing information on the distribution of values and the level of variation for each variable under study [11] [12]. To evaluate how the independent variables collectively impact the dependent variable, this study employs a multiple linear regression framework. The estimation process utilizes the Ordinary Least Squares (OLS) technique, which follows the preliminary descriptive analysis [13]. The multiple linear regression model is defined as:

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

where Y_i is the dependent variable (Crime), $X_{1i}, X_{2i}, \dots, X_{ki}$ are the independent variables (ln_GDP, Poor Population, Expenditure, HDI), β_0 is the intercept, $\beta_1, \beta_2, \dots, \beta_k$ are the regression coefficients, and ε_i is the error term [14]. The OLS procedure estimates the coefficients β_j by minimizing the sum of squared residuals:

$$Minimize\ \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2 \quad (2)$$

where $\hat{Y}_i$ is the predicted value from the model.

The steps of analysis are: first, compute descriptive statistics to understand the distribution of each variable; second, test classical assumptions such as normality of residuals, multicollinearity, and homoskedasticity; third, apply OLS to calculate the regression coefficients; and finally, perform F-tests and t-tests to assess simultaneous and partial significance of the independent variables [15]. Positive coefficients indicate that an increase in the corresponding independent variable raises the predicted crime level, while negative coefficients indicate an inverse relationship [16][17].

 DOI: https://doi.org/10.52435/jaiit.v8i1.773

Subsequently, a two-step significance evaluation is performed to ascertain the empirical validity of the proposed variable interactions [18]. This stage involves a bifurcated testing approach consisting of F and t statistics. The F-test is designed to evaluate the aggregate significance of the independent variables concurrently. Conversely, the t-test provides a disaggregated analysis, determining the unique significance of each predictor in explaining the target variable's variance [19].

Methodologically, both evaluations follow five identical hypothesis testing steps, with the primary distinction lying in the specific test statistic formula applied. The first step begins by explicitly establishing the null hypothesis (H_0) and the alternative hypothesis (H_1) in plain text. For the simultaneous evaluation (F-test), the null hypothesis (H_0) states that all independent variables do not concurrently have a significant effect on the dependent variable, whereas the alternative hypothesis (H_1) states that all independent variables concurrently have a significant effect on the dependent variable. Conversely, for the partial evaluation (t-test), the null hypothesis (H_0) states that an individual independent variable does not have a significant effect on the dependent variable, while the alternative hypothesis (H_1) states that an individual independent variable has a significant effect on the dependent variable. The second step involves determining the significance level (α), which is firmly set at 0.05 [20].

The fundamental difference between these two evaluations lies in the third step, which is the determination of the test statistics. The F-test utilizes a variance ratio based on the coefficient of determination (R^2), the number of predictors (k), and the sample size (n) [21], formulated as follows:

$$F = \frac{\frac{R^2}{k}}{\frac{1 - R^2}{n - k - 1}} \qquad (3)$$

Meanwhile, the t-test employs the ratio of the parameter estimate to its standard error (SE), which is formulated as:

$$t = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \qquad (4)$$

Subsequently, the fourth step establishes the rejection region for the decision criteria, where (H_0) is rejected if the significance value (p-value) < (0.05), or if the calculated statistic exceeds its respective critical table value or if the calculated statistic exceeds its respective critical table value. Finally, the fifth step concludes with the drawing of empirical conclusions to determine whether the variance of the target variable can be significantly explained either simultaneously or individually.

The validity of the regression output depends on the fulfillment of several classical assumption tests to ensure that the estimated parameters satisfy the BLUE (Best Linear Unbiased Estimator) criteria. In particular, the regression residuals are expected to fulfill the IIDN assumptions, meaning that they are independently, identically, and normally distributed, while the independent variables must remain free from multicollinearity issues. Failure to satisfy these fundamental assumptions may result in biased and unreliable estimation results [22]. Normality of the data is scrutinized as the primary classical assumption, ensuring that the variables under study do not deviate significantly from a normal distribution [23]. Evaluating this criterion involves the application of the Kolmogorov-Smirnov statistical method because it provides a reliable approach for assessing the conformity between empirical and theoretical normal distributions. Prior studies have indicated that the Kolmogorov–Smirnov method is comparatively less affected by sample size fluctuations and tends to perform consistently when skewness and kurtosis coefficients are close to zero [24]. The Kolmogorov-Smirnov test statistic (D) is mathematically expressed as [25]:

$$D = maximum|F_0(X) - S_n(X)| \qquad (5)$$

The normality test is based on two statistical hypotheses. The null hypothesis (h_0) states that the dataset is normally distributed, while the alternative hypothesis (H_1) states that the dataset is not normally distributed. Under this framework, the null hypothesis of normality is accepted when the resulting *p-value* exceeds 0.05. conversely, a *p-value* below this threshold ($p < 0.05$) indicates that the datasets is not normally distributed[26].

The Variance Inflation Factor (VIF) functions as the definitive indicator for detecting multicollinearity, which evaluates whether redundant correlations exist among the explanatory variables [27]. The VIF value is calculated using the following equation [28]:

$$VIF = \frac{1}{1 - R^2} \tag{6}$$

Under this criterion, VIF values surpassing the threshold of 5 or 10 signify that the estimated regression coefficients lack reliability, as high redundancy between independent variables can distort the statistical output [29]. The next step entails applying the Glejser method as a final classical assumption check to detect any variance inequalities within the regression model [30]. The objective is to verify whether the residuals exhibit homoscedastic properties across all data points. The Glejser regression model can be expressed as follows [31]:

$$|u_i| = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \upsilon_1 \tag{7}$$

A robust model is identified when the probability values associated with the independent variables and absolute residuals are greater than the 5% threshold. Consequently, surpassing this benchmark demonstrates that the model is devoid of heteroscedasticity [26].

After all classical assumption prerequisites have been satisfied, the analysis proceeds to the coefficient of determination (R^2) test to evaluate the model's capacity in accounting for dependent variable fluctuations [32]. This metric spans a range between 0 and 1, serving as a benchmark for explanatory quality. As the R^2value approaches unity, it signifies a robust capability of the predictors to jointly clarify the variance in the outcome variable. Conversely, a figure nearing zero suggests that the model possesses minimal explanatory strength [30]. The structural formulation for calculating this coefficient is presented below [33].

$$R^2 = 1 - \frac{\sum_{i=1}^{m} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{m} (Y_i - \bar{Y})^2} \tag{8}$$

Goodness-of-fit within this model is quantified through the R^2 value, which assesses the proportion of variance explained by the regression. This coefficient is derived from comparing the residual sum of squares against the total variance observed in the empirical data. In this context, X_i denotes the predicted outcomes produced by the regression analysis, whereas Y_i identifies the original values recorded from the observations.

3. Results and Discussions

The fundamental characteristics of the dataset, such as the average, standard deviation, variance, and the spectrum of minimum to maximum values, are outlined through descriptive statistical analysis. Table 2 provides a comprehensive summary of these statistical metrics for the variables evaluated in this research.

Table 2. Descriptive Statistics Analysis

Variable	Count	Mean	Std. Dev.	Min	Max
Crime	31	14991.13	16092.42	1593	60724
ln GDP	31	12.13	1.09	10.40	14.48
Poor Population	31	726.38	1023.35	41.11	3893.82
Expenditure	31	1460765.774	267111.3	975854	2109071
HDI	31	74.65	2.65	69.14	81.62

Based on Table 2, the analysis shows that GDP values generally have a relatively good average with moderate variation across regions. However, the wide range between the minimum and maximum

DOI: https://doi.org/10.52435/jaiit.v8i1.773

values indicates a gap in economic activity levels between provinces in the research sample. The variable for the number of poor people shows a very large variation. The high standard deviation and significant differences between the minimum and maximum values reflect inequality in poverty levels across regions. This condition indicates that poverty distribution is uneven and still concentrated in certain areas. Meanwhile, the per capita expenditure variable shows relatively more stable variation compared to the number of poor people. Although there are differences across regions, the data distribution tends to be more even, indicating a degree of uniformity in the welfare levels across most areas. The Human Development Index (HDI) has a relatively small standard deviation, suggesting that human development levels across the regions in the sample are more homogeneous. This indicates that the difference in human development quality is less pronounced compared to other economic indicators.

Following the preliminary assessment of data characteristics, an Ordinary Least Squares (OLS) approach was applied to construct a multiple linear regression framework. By minimizing the aggregate of squared errors, this technique effectively bridges the gap between the observed figures and the model's estimated values for the dependent variable. Consequently, as long as classical prerequisites are fulfilled, the resulting estimators conform to the Best Linear Unbiased Estimator (BLUE) criteria. To evaluate how crime frequencies across 31 provinces are influenced by gross domestic product, poverty demographics, individual spending, and the human development index, the aforementioned OLS regression was deployed. The execution of this methodology yielded the subsequent mathematical formulation:

$$\hat{Y} = -191.50 + 7192.02X_1 + 5{,}31X_2 - 0{,}03X_3 + 2130{,}32X_4$$

Assuming all other factors remain static, the formulation reveals that criminal activity is positively driven by GDP, poverty levels, and the human development index, whereas personal expenditure exerts an inverse influence. Quantitatively, a single-unit expansion in GDP correlates with a projected rise of 7,192.02 incidents in criminality. Similarly, when the impoverished demographic grows by one unit, crime incidents are anticipated to escalate by 5.31. In contrast, an equivalent increase in individual spending corresponds to a 0.03 reduction in criminal occurrences. Furthermore, elevating the HDI by a single point is predicted to drive up the crime metric by 2,130.32 units, provided other conditions are strictly controlled.

To validate whether the coefficients generated by the model are statistically significant, a series of significance tests needs to be conducted. The significance tests are performed using two approaches: simultaneously and partially. To determine whether the regression model as a whole is significant, a simultaneous significance test is conducted using the F-test.

Table 3. Simultaneous Significance Test Results

F-statistic	Prob F-statistic
22.34	4.29×10^{-8}

Referring to the simultaneous evaluation in Table 3, an F-statistic of 22.34 is recorded, accompanied by a probability value of 4.29 e-08. Because this figure falls drastically below the standard 0.05 threshold, the null hypothesis is consequently dismissed in favor of the alternative hypothesis. Therefore, it is confirmed that crime frequencies are collectively and significantly impacted by variations in GDP, poverty demographics, per capita spending, and HDI. After the model was proven to be significant simultaneously, the analysis continued with the partial test (t-test) to examine the individual effects of each independent variable.

Table 4. Results of the Partial Significance Test

Variable	Coeff	Std. Error	t	Sig. (p-value)
const	-1.915 e+05	5.75 e+04	-3.33	0.003
ln_GDP	7192.02	2458.61	2.93	0.007
Poor Population	5.31	2.512	2.11	0.044

Expenditure	-0.03	0.01	-3.27	0.003
HDI	2130.32	944.63	2.25	0.033

The individual significance assessments detailed in Table 4 reveal a probability score of 0.007 for the GDP metric, successfully satisfying the < 0.05 criteria for statistical significance. Consequently, economic output independently drives crime levels. The upward trajectory of crime alongside economic growth is evidenced by its positive coefficient of 7,192.02, assuming static external variables. Furthermore, the poverty indicator yields a p-value of 0.044, confirming its significant and independent role in shaping crime trends. Its coefficient of 5.31 demonstrates that a growing impoverished population directly translates into higher criminality. Regarding individual spending, the recorded probability of 0.003 falls below the strict 0.05 cutoff, proving its substantial influence on criminal frequency. The deterrent effect of heightened personal expenditure is reflected in its negative coefficient of -0.03. Lastly, human development index variations achieved statistical significance with a p-value of 0.033. Surprisingly, assuming other factors are controlled, the positive coefficient of 2,130.32 denotes that advancements in human development are correlated with an escalation in regional crime.

Although the model shows statistical significance, the validity of the estimation results still needs to be tested through classical assumption testing. Therefore, normality, multicollinearity, and heteroscedasticity tests were conducted. The normality test was performed to determine whether the research data is normally distributed or approximately normal. This test uses the Kolmogorov-Smirnov test, where data is considered normally distributed if the probability value is greater than 0.05. Otherwise, if the probability value is less than 0.05, the data is considered non-normally distributed.

Table 5. Results of the Normality Test

KS Stat	p-value
0.19	0.17

Based on Table 5, the results of the Kolmogorov-Smirnov normality test show that the KS Statistic value is 0.19, with a significance value of 0.17. Since the significance value is greater than 0.05, it can be concluded that the residuals of the regression model are normally distributed. After the normality assumption is satisfied, the analysis continued with multicollinearity testing to examine the presence of high correlations between the independent variables. This test is conducted using the Variance Inflation Factor (VIF), where a VIF value greater than 5 or 10 generally indicates strong correlations between the independent variables, which may affect the stability of the regression coefficient estimates.

Table 6. Results of the Multicollinearity Test

Variable	VIF
ln_GDP	3.19
Poor Population	2.94
Expenditure	2.64
HDI	2.79

The results of the multicollinearity test in Table 6 show that the GDP variable has a VIF value of 3.19, the number of poor people has a VIF value of 2.94, per capita expenditure has a VIF value of 2.64, and HDI has a VIF value of 2.79. All VIF values are below the common threshold of 10, indicating that there is no serious multicollinearity issue between the independent variables in the model. Therefore, the regression model meets the assumption of no high correlations among the independent variables. Next, a heteroscedasticity test was performed to ensure that the residual variance in the model remains constant across all observations. One commonly used method is the Glejser test, where if the significance value between the independent variables and the absolute residual values is greater than 0.05, the model is considered free from heteroscedasticity.

 DOI: https://doi.org/10.52435/jaiit.v8i1.773

Table 7. Results of the Heteroscedasticity Test

Variable	Coeff	Std. Error	t	Sig. (p-value)
ln_GDP	2779.10	1606.28	1.73	0.09
Poor Population	-1.53	1.63	-0.93	0.36
Expenditure	-0.002	0.004	-0.40	0.69
HDI	-331.47	258.20	-1.28	0.21

Based on the results of the heteroscedasticity test using the Glejser method in Table 7, it was found that all variables have a significance (p-value) above the 0.05 significance level. The GDP variable has a p-value of 0.09, the number of poor people has a p-value of 0.36, per capita expenditure has a p-value of 0.69, and the Human Development Index has a p-value of 0.21. These results indicate that there is no significant effect of the independent variables and the absolute residual values, so it can be concluded that the regression model does not exhibit heteroscedasticity. With all classical assumptions met, the regression model is deemed suitable for further interpretation. The analysis then continues with the coefficient of determination (R^2) test to assess how well the regression model explains the variation in the dependent variable. The R^2 value provides an indication of the overall contribution of the independent variables, where the closer it is to 1, the greater the model's ability to explain this variation. This analysis serves as an initial step in understanding the model's strength before examining the individual effects of each variable in more detail.

Table 8. Results of the Coefficient of Determination

R-squared	Adj. R-squared
0.775	0.740

Based on Table 8, the coefficient of determination (R^2) is 0.77, with an Adjusted R^2 value of 0.74. This indicates that approximately 77.5% of the variation in the dependent variable, which is the immunization rate, can be explained jointly by the independent variables included in the model. The remaining 22.5% is influenced by other factors not included in this study.

4. Conclusion

The results of the study indicate that GDP, the number of poor people, per capita expenditure, and HDI simultaneously have a significant effect on crime rates, with a coefficient of determination (R^2) of 0.775, meaning that 77.5% of the variation in crime rates can be explained by the model. Partially, GDP, the number of poor people, and HDI have a positive effect on crime rates, while per capita expenditure has a negative effect. These findings suggest that the dynamics of economic growth, poverty levels, and the quality of human development are related to the variation in crime rates across regions, while improvements in public welfare have the potential to reduce crime rates. This result implies the importance of inclusive economic development policies, poverty reduction, and welfare improvement as part of crime control strategies. However, this study has limitations due to the use of secondary cross-sectional data and the exclusion of other potential influencing variables, such as unemployment rates, income inequality, and the effectiveness of law enforcement. Therefore, future research is recommended to use panel data and include additional socio-economic and institutional variables to gain a more comprehensive understanding of the determinants of crime rates.

Acknowledgement

We would like to express our gratitude to the relevant institutions for providing the data used in this study. Our sincere thanks and appreciation go to all parties who have provided guidance and support throughout the research process.

References

[1] Badan Pusat Statistik, *Statistik Kriminal 2024/2025*, vol. 16. Jakarta: Badan Pusat Statistik, 2025. Accessed: Feb. 04, 2026. [Online]. Available: https://www.bps.go.id

[2] B. S. Pratama and G. Purwanto, “PERBANDINGAN K-MEANS DAN K-MEDOIDS DALAM PENGELOMPOKKAN TINGKAT KEJAHATAN PADA PROVINSI JAWA TENGAH,” *Idealis: Indonesia Journal Information System*, vol. 8, no. 2, pp. 295–303, Jul. 2025, [Online]. Available: http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

[3] D. Yuliyanti and Martanto, “CLUSTERING TINGKAT KEJAHATAN KRIMINAL MENGGUNAKAN METODE K-MEANS DI WILAYAH KABUPATEN CIREBON,” *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 6, pp. 3509–3514, Dec. 2023, doi: https://doi.org/10.36040/jati.v7i6.8894.

[4] W. Atthoriq, Septriani, and E. Winarni, “Analisis Pengaruh Ketimpangan Pendapatan dan Pengangguran Terhadap PDRB Melalui Kriminalitas Di Indonesia,” *Jurnal Development*, vol. 12, no. 2, pp. 192–209, Dec. 2024, doi: https://doi.org/10.53978/jd.v12i2.381.

[5] Risma and A. R. Dariah, “Pengaruh Kemiskinan, PDRB, dan Pengangguran terhadap Kriminalitas di Jakarta dan Jabar,” *Jurnal Riset Ilmu Ekonomi dan Bisnis*, vol. 4, no. 2, pp. 109–116, Dec. 2024, doi: 10.29313/jrieb.v4i2.5025.

[6] A. S. Wicaksono and Suharto, “Analisis pengaruh faktor ekonomi terhadap kriminalitas di Kabupaten/Kota Daerah Istimewa Yogyakarta,” *Jurnal Kebijakan Ekonomi dan Keuangan*, vol. 2, no. 1, pp. 50–57, 2023, doi: 10.20885/jkek.vol2.iss1.art6.

[7] A. Riyardi and R. B. Guritno, “Faktor Ekonomi Yang Mempengaruhi Penurunan Kriminalitas Di Provinsi Jawa Tengah: Analisis Mikroekonomi,” *Jurnal Ilmu Kepolisian*, vol. 16, no. 1, pp. 50–61, Apr. 2022, doi: https://doi.org/10.35879/jik.v16i1.349.

[8] R. Saputra, “Analisis Tingkat Pendidikan, Kemiskinan dan Pengangguran Terhadap Kriminalitas di Bekasi,” *(JIHHP) Jurnal Ilmu Hukum, Humaniora dan Politik*, vol. 3, no. 4, pp. 159–163, Aug. 2023, doi: 10.38035/jihhp.v3i4.

[9] Z. I. Idris, W. L. Husin, P. Mu, and M. Fuad Alamsyah, “Journal of Innovative and Creativity Dampak Pengangguran, Kemiskinan dan Indeks Pembangunan Manusia terhadap tingkat kriminalitas di Provinsi Gorontalo,” *Journal of Innovative and Creativity*, vol. 6, no. 1, p. 2026, Jan. 2026.

[10] A. S. Ardelia, “PENGARUH INDEKS PEMBANGUNAN MANUSIA TERHADAP TINGKAT KRIMINALITAS DI KALIMANTAN BARAT DALAM KERANGKA EKONOMI ISLAM,” *JMI: Jurnal Muamalat Indonesia*, vol. 3, no. 1, pp. 276–284, 2023, doi: 10.64075/jmi.v3i3.

[11] A. Ramdani, C. Rakhmat, E. S. Nurdin, and A. Kosasih, “PEMBELAJARAN KEWIRAUSAHAAN DI PERGURUAN TINGGI (STUDI ANALISIS KUANITATIF DESKRIPTIF TERHADAP PROFIL KARAKTER KINERJA MAHASISWA),” *JURNAL EKONOMI PENDIDIKAN DAN KEWIRAUSAHAAN*, vol. 11, no. 1, pp. 5–20, Apr. 2023, doi: 10.26740/jepk.v11n1.p5-20.

[12] H. M. Hamid, S. R. Mulyani, and C. Wirawan, “Pengaruh Kualitas Pelayanan Pendekatan Terra terhadap Kepuasan Pasien,” *Jurnal Ekonomi dan Statistik Indonesia*, vol. 4, no. 1, pp. 46–60, Sep. 2024, doi: 10.11594/jesi.04.01.05.

[13] Nurhaswinda, D. Poni Egistin, M. Yahdi Rauza, R. Has Ramadhan, and S. Ramadani, “Analisis regresi linier sederhana dan penerapannya,” *JURNAL CAHAYA NUSANTARA*, vol. 1, no. 2, pp. 3093–8113, Jan. 2025, [Online]. Available: https://creativecommons.org/licenses/by/4.0/

[14] N. Ratnaduhita *et al.*, “Pengaruh Pelatihan dan Pengembangan SDM Pada ABA Collection Terhadap Kinerja Pegawai Menggunakan Metode Regresi Linear Berganda,” *Journal of Advances in Information and Industrial Technology (JAIIT)*, vol. 3, no. 1, pp. 19–30, May 2021, doi: https://doi.org/10.52435/jaiit.v3i1.87.

[15] A. Ramadan, N. Chamidah, I. N. Budiantara, N. R. A. A. Siregar, and D. Aydin, "Modeling Pulmonary Tuberculosis Case Based on HIV and AIDS Cases in Indonesia Using Negative Binomial Regression Least Square Spline," *Statistics, Optimization & Information Computing*, 2026.

[16] A. Ramadan, N. Chamidah, I. N. Budiantara, B. Lestari, and D. Aydin, "Method for Modelling the Number of HIV and AIDS Cases Using Least Square Spline Biresponse Nonparametric Negative Binomial Regression," *MethodsX*, p. 103336, 2025.

[17] A. Ramadan, N. Chamidah, and I. N. Budiantara, "Modelling the number of HIV cases in Indonesia using negative binomial regression based on least square spline estimator," *Commun. Math. Biol. Neurosci.*, vol. 2024, p. Article-ID, 2024.

[18] N. A. Rabukuntari, B. P. Maulidina, and L. A. Hasanah, "Pengaruh Usia dan Jenis Kelamin Terhadap Selera Menonton Tayangan Berita Menggunakan Metode Regresi Logistik Ordinal," *Journal of Advances in Information and Industrial Technology*, vol. 4, no. 1, pp. 11–22, May 2022, doi: 10.52435/jaiit.v5i1.310.

[19] S. Junaidi, M. Devegi, and H. Kurniawan, "Pelatihan Pengolahan dan Visualisasi Data Penduduk menggunakan Python," *ADMA : Jurnal Pengabdian dan Pemberdayaan Masyarakat*, vol. 4, no. 1, pp. 151–162, Jul. 2023, doi: 10.30812/adma.v4i1.2963.

[20] T. D. N. Marwinda and D. Danardono, "Perbandingan Iuran Normal Pensiun Metode Entry Age Normal dan Projected Unit Credit dengan Suku Bunga CIR (Cox Ingersoll Ross)," *JURNAL PEMBELAJARAN DAN MATEMATIKA SIGMA (JPMS)*, vol. 10, no. 2, pp. 133–138, Nov. 2024, doi: 10.36987/jpms.v10i2.5881.

[21] D. Gujarati, *Econometrics by Example*. PALGRAVE MACMILLAN, 2011.

[22] D. R. Amalia and W. Triwacananingrum, "The disclosure effect of sustainability reporting and financial statements on investment efficiency: Evidence in Indonesia," *Indonesian Journal of Sustainability Accounting and Management*, vol. 6, no. 1, pp. 82–93, 2022.

[23] S. H. Difinubun, O. Dominggus Nara, and M. Abdin, "ANALISIS PENGARUH SUMBER DAYA MANUSIA TERHADAP ASPEK KINERJA PEKERJA PADA PROYEK PEMBANGUNAN GEDUNG LABORATORIUM TERPADU PENDUKUNG BLOK MASELA UNIVERSITAS PATTIMURA," *JOURNAL AGREGATE*, vol. 2, no. 1, pp. 76–86, Mar. 2023, doi: https://doi.org/10.31959/ja.v2i1.1252.

[24] S. DEMİR, "Comparison of Normality Tests in Terms of Sample Sizes under Different Skewness and Kurtosis Coefficients," *International Journal of Assessment Tools in Education*, vol. 9, no. 2, pp. 397–409, Jun. 2022, doi: 10.21449/ijate.1101295.

[25] Nuryadi, T. D. Astuti, E. S. Utami, and m Budiantara, *DASAR-DASAR STATISTIK PENELITIAN*. Yogyakarta: SIBUKU MEDIA, 2017. doi: 978-602-6558-04-6.

[26] I. Nurfitriani, "Pengaruh Rasio Profitabilitas terhadap Nilai Perusahaan (Studi pada Perusahaan LQ45 yang Terdaftar di Bursa Efek Indonesia Periode 2019-2023)," *EKALAYA : Jurnal Ekonomi Akuntansi*, vol. 1, no. 4, pp. 385–396, Dec. 2023, doi: 10.59966/ekalaya.v1i4.1199.

[27] Effiyaldi *et al.*, "PENERAPAN UJI MULTIKOLINIERITAS DALAM PENELITIAN MANAJEMEN SUMBER DAYA MANUSIA," *Jurnal Ilmiah Manajemen dan Kewirausahaan (JUMANAGE)*, vol. 1, no. 2, pp. 94–102, Jul. 2022, doi: 10.33998/jumanage.2022.1.2.89.

[28] N. A. Alya, Q. Almaulidiyah, B. R. Farouk, D. Rantini, A. Ramadan, and F. Othman, "Comparison of Geographically Weighted Regression (GWR) and Mixed Geographically Weighted Regression (MGWR) Models on the Poverty Levels in Central Java in 2023," *IAENG International Journal of Applied Mathematics*, vol. 54, no. 12, pp. 2746–2757, Dec. 2024.

[29] D. Rantini, A. Ramadan, A. Sesay, and M. F. A. Hillaby, "House Price Modeling in Semarang and Surabaya City using Component Regression with Frequentist and Bayesian Approach," in *AIP Conference Proceedings*, AIP Publishing, Nov. 2024, pp. 330–353. doi: 10.2991/978-94-6463-566-9_22.

[30] F. Kumayas, A. G. Kumenaung, and H. F. D. Siwu, "Pengaruh Jumlah Penduduk, Tingkat Pendidikan Dan Tingkat Pengangguran Terhadap Kemiskinan Di Kabupaten Minahasa," *Jurnal Berkala Efisiensi Ilmiah*, vol. 24, no. 4, pp. 72–89, Apr. 2024.

[31] S. M. Sholihah, N. Y. Aditiya, E. S. Evani, and S. Maghfiroh, "KONSEP UJI ASUMSI KLASIK PADA REGRESI LINIER BERGANDA," *Jurnal Riset Akuntansi Soedirman*, 2023, doi: 10.32424/1.jras.2023.2.2.10792.

[32] W. S. Tondok, J. B. Kalangi, and W. F. I. Rompas, "PENGARUH ANGKATAN KERJA DAN PENGELUARAN KONSUMSI RUMAH TANGGA TERHADAP PERTUMBUHAN EKONOMI DI KABUPATEN TANA TORAJA TAHUN 2011-2021," *Jurnal Berkala Ilmiah Efisiensi*, vol. 23, no. 5, pp. 49–60, May 2023.

[33] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, pp. 1–24, 2021, doi: 10.7717/PEERJ-CS.623.