JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 10, No. 1, January 2026, pp. 1511-1532

The Severe Stunting Cases of Children in Central Java Province Explained by Negative Binomial Regression Model

Rhendy K. P. Widiyanto[1,2*], Fatkhurokhman Fauzi[1,3], Achmad Fauzan[1,4], Anang Kurnia[1]
[1]Statistics and Data Science Study Program, IPB University, Indonesia
[2]Center for Agricultural Data and Information System, Ministry of Agriculture, Indonesia
[3]Department of Statistics, Universitas Muhammadiyah Semarang, Indonesia
[4]Department of Statistics, Universitas Islam Indonesia, Indonesia
stat.rhendy@apps.ipb.ac.id

Article History:
Received : 17-09-2025
Revised : 22-11-2025
Accepted : 02-12-2025
Online : 01-01-2026

Keywords:
Backward Elimination;
Severe Stunting;
Negative Binomial Regression;
Overdispersion;
Poisson Regression.

ABSTRACT

Severe stunting, or very short stature among children, remains a critical public health concern in Central Java Province. Robust statistical modelling is essential to identify the key factors associated with these cases and to guide targeted interventions. This study employs count regression models with an offset variable to analyze the factors influencing severe stunting cases across districts in Central Java. By using 2023 official data in districts level taken from the Ministry of Home Affairs and the Statistics Indonesia, we initially utilize a Poisson regression model in this study. However, due to evidence of overdispersion, a Negative Binomial regression model was adopted. Backward elimination was then applied to obtain the most parsimonious model. The Negative Binomial regression successfully addressed overdispersion. Five factors were identified as having a statistically significant influence on severe stunting cases: (1) the proportion of pregnant mothers with Chronic Energy Deficiency receiving nutritious food supplements, (2) the percentage of toddlers (6-23 months) receiving complementary nutritious food, (3) the proportion of households with access to good sanitation, (4) Gross Domestic Product per capita, and (5) the number of local healthcare facilities. These factors have negative relation to the stunting rates, meaning improving these factors will reduce the rates of severe stunting. The findings provide a validated statistical model for severe stunting and offer clear policy directions. To mitigate severe stunting, local governments should prioritize: enhancing nutritious food support for pregnant mothers and toddlers, improving household sanitation, stimulating local economic growth, and increasing accessibility to healthcare facilities.

Crossref
https://doi.org/10.31764/jtam.v10i1.34846

——————————— ◆ ————————————

A. INTRODUCTION

Stunting is a form of chronic growth disorder in children caused by long-term malnutrition and exposure to repeated infections, especially during the first 1,000 days of life. This disorder not only affects the height of children who become much shorter than their age standards, but also has implications for suboptimal brain development, cognitive delays, motor function problems, and reduced learning abilities. This condition not only affects children's linear growth, but also leads to suboptimal brain development, cognitive delays, reduced learning capacity, and increased risk of non-communicable diseases later in life, which in turn lowers productivity and economic potential at adulthood (Soliman et al., 2021; Taslim et al., 2023). Stunting is often associated with low socioeconomic conditions, poor access to nutritious food, inadequate sanitation, and repeated infections that worsen nutrient absorption (Santoso & Pujianto, 2024). Within this broader condition, severe stunting (very short-for-age) represents the most critical form of growth failure and therefore warrants specific analytical and policy attention rather than being treated only as part of overall stunting prevalence.

Ministry of Health, Indonesia classified stunting in children into 2 groups: stunted children and severely stunted children. Stunted children are under-5-year toddlers who have their z score of height or weight ratio by age is lower than -2 standard deviation of other children in the same age. The severe ones then those having z score lower than -3 standard deviation of other children with same age (TN2PK, 2017). Various previous studies have shown many factors influence stunting in children, yet few are focused on the severe group of stunting. One of the significant factors influence stunting is the percentage of mothers with chronic energy deficiency (CED) who have receive nutritional support which is said to be able to reduce risk of child stunting, especially through improvements in maternal nutritional status before and during pregnancy. Providing nutritious complementary foods to toddlers has also been proven to be effective in reducing the prevalence of stunting. This could be done by increasing nutritional intake during the golden period of child growth and development (Humphrey et al., 2019). Another factor is to have access to good sanitation, for example closed toilets and hygienic living. This may also be effective in reducing prevalence of stunting by minimizing the risk of disease infection that can interfere with nutrient absorption (Rah et al., 2015). However, most of this empirical evidence is still framed in terms of overall stunting, while specific evidence focusing on severe stunting as the most extreme outcome remains relatively limited.

The National Team for the Acceleration of Poverty Reduction, under Office of Vice President, Indonesia, associated stunting to the limited access of households to nutritious meals (TN2PK, 2017). Furthermore, FAO defines inability to access adequate food is linked to the level of household income and consecutively produce annual indicator to monitor food price related to people access to healthy diets (Food and Agriculture Organization, 2024). Thus, higher GDP per capita should also supports stunting reduction through increased purchasing power for healthy food and healthcare services, although its impact depends on the distribution of income and education. A high Human Development Index (HDI) reflects better access to education, health services, and income, which are significantly associated with stunting reduction (Johri et al., 2016). In addition, adequate local healthcare facilities increase immunization coverage, nutrition counselling, and community-based interventions that support optimal child growth (Regulation of the Minister of Health of the Republic of Indonesia Number 2 of 2020 on Child Anthropometry Standards, 2020). Taken together, these findings highlight that maternal nutrition, child feeding, sanitation, economic capacity, human development, and access to healthcare are key determinants that must also be examined specifically for severe stunting.

Modelling the factors that influence stunting is essential to understanding the complex interactions between health, social, and economic aspects that contribute to this condition. Stunting, especially the severe one, may not only be influenced by poor nutritional intake, but also by various determinants such as maternal CED on mothers, complementary feeding, sanitation quality, per capita income, human development index (HDI), and availability of healthcare facilities (Dewi et al., 2019; Mulyani et al., 2025; Vaivada et al., 2020). By developing a model which includes these factors, policymakers can identify the most effective intervention strategies, such as increasing maternal access to nutrition and health support, improving sanitation infrastructure, and strengthening community health services(Kalinda et al., 2024). Moreover, a well-specified model provides a quantitative basis to prioritise districts and cities with the highest burden of severe stunting and to design targeted programmes that are tailored to local needs, thereby supporting more efficient and sustainable stunting reduction efforts.

Modelling the factors that influence severe stunting cases using Poisson regression is an appropriate approach, especially when the data analysed is in the form of count data, such as the incidence of severe stunting in a certain population. Poisson regression can be used to evaluate the relationship between explanatory variables, such as maternal nutritional status, complementary feeding, good sanitation access, per capita GDP, HDI, and number of health facilities, on the incidence of severe stunting (Htet et al., 2023; Khan et al., 2024; Seifu et al., 2024). By adding offset variable (such as the population of children in a particular region), the model can produce more accurate estimates in epidemiologically and demographically relevant contexts (Partap et al., 2019). However, when the variance of the counts is substantially larger than the mean, as is often the case for health-related incidence data, naive application of the Poisson model can result in underestimated standard errors and misleading statistical inferences, so analysts must be aware of the overdispersion issue in their model-building process (Nariswari et al., 2023).

Overdispersion in Poisson regression occurs when the variance of the data is greater than the mean assumed by the Poisson model, causing bias in standard errors estimate, and thus lead to inaccurate inferences. The effects of overdispersion are often due to additional unmodeled variability, such as excess zeros or heterogeneity between observations. Research suggests that overdispersion can be addressed with alternative models such as Negative Binomial regression or quasi-likelihood approaches that accommodate the additional variability without changing the underlying structure of the model (Fernandez & Vatcheva, 2022; Mardalena et al., 2022; Tiara et al., 2023). In practice, recognising and correcting for overdispersion is a crucial step to ensure that the identified determinants of severe stunting are not artefacts of misspecified variance assumptions.

Negative Binomial regression model as solution to overdispersion issues have been widely applied in various cases over the past eight years. In a study of anaemia in women of childbearing age in Indonesia, Negative Binomial regression produced the best model compared to the Poisson ones, with a better measurement of goodness of fit and a deviation approaching 1, indicating the ability of this model to handle large variances (Tiara et al., 2023). In a study of child mortality, this approach provided accurate estimates of risk factors, making it the primary choice for data with unequal variances (Fitriyah et al., 2015). In a study of elderly health, Negative Binomial regression provided better results than Poisson regression for count data showing high variance due to inter-individual variability (Fernandez & Vatcheva, 2022). In ischemic stroke case data, this model was able to control for random effects between hospitals, providing deeper insight into the risk of hospitalization. Other studies have shown that Negative Binomial regression is superior in analysing data with strong spatial dependence, such as the geographic distribution of infectious diseases (da Silva & Rodrigues, 2014). This solution ensures more reliable analysis in various contexts by accommodating additional variability that cannot be explained by the Poisson model. Despite this extensive use, applications of Negative Binomial regression that explicitly model severe stunting rates using an offset for the under-five population, particularly at the district/city level in Central Java Province, are still scarce.

any have shown that Negative Binomial regression can overcome overdispersion in Poisson regression effectively. However, previous empirical studies on stunting in Indonesia generally focus on overall stunting prevalence, use methods that do not fully account for overdispersion in count data, or do not explicitly model severe stunting as a separate outcome of interest. There is also limited evidence on the use of a systematic model selection procedure, such as backward elimination, to obtain a parsimonious yet interpretable Negative Binomial model for severe stunting cases at the district/city level. Therefore, this study applies a Negative Binomial regression model with an offset variable to model the rate of severe stunting cases among under-five children in districts and cities of Central Java Province, and uses backward elimination to identify the most important determinants. Methodologically, the study contributes by demonstrating the usefulness of Negative Binomial regression with an offset for overdispersed severe stunting data, and by providing a clear comparison with the Poisson model. From a policy perspective, the findings offer empirical evidence on which combinations of maternal nutrition, child feeding, sanitation, economic capacity, and healthcare facilities should be prioritised to design more targeted and effective severe stunting prevention strategies in Central Java.

B. METHODS

1. Methods Used in Research

This subchapter will explain the series of methods used in this study. It also explains the theoretical aspects from a statistical perspective. The following is an explanation of the statistical methods used in this study:

a. Multicollinearities

Multicollinearity describes a condition that may arise when analysts simultaneously consider more than one explanatory variable in regression model. This condition occurs when two or more explanatory variables in a sample overlap. Because of this overlap, the analysis method cannot fully distinguish the explanatory factors from each other or isolate their independent effects. The impact of multicollinearity makes the regression coefficients unstable (even wrong), increases the variance of the regression coefficients, and causes prediction errors (Salmerón et al., 2020). Detecting the presence or absence of multicollinearity between independent variables using the Variance Inflation Factor (VIF) (Salmerón et al., 2020). If the VIF is high (with 5 or 10 as the accepted threshold), then it can be concluded that there is multicollinearity (Salmerón et al., 2018). VIF is the ratio between the original model OLS estimator's variance and the model's variance where the variables are orthogonal. Equation 1 is the VIF calculation formula (Salmerón et al., 2020). In this study, VIF is used to ensure that the relationships among predictors do not distort their estimated effects on severe stunting.

$$VIF\ (X_k) = \frac{1}{1 - {R_k}^2} \tag{1}$$

In Equation 1, X_k denotes the independent variable whose VIF is calculated, R_k^2 is the coefficient of determination of the regression of variable X_k against all other independent variables in the model.

b. Breusch–Pagan test

One of the main assumptions that must be met in regression analysis is that the variance of the residuals is constant throughout the observations, known as homoscedasticity. The opposite of homoscedasticity is heteroscedasticity, a condition when the residual variance is not constant and varies according to the value of the independent variable. The consequence of heteroscedasticity is that the coefficient estimate becomes inefficient (although unbiased). In addition, the presence of heteroscedasticity reduces the accuracy of the standard error and even has implications for testing on the t-test statistics and the F-test becoming invalid (Le Gallo et al., 2020). Methods that can be used to test heteroscedasticity include the Breusch-Pagan (BP) test. The BP test aims to detect whether there is a relationship between the residual variance and the independent variables in the model. In a spatial context, this test is applied to identify whether different variance patterns in spatial data can be caused by geographic location factors or other characteristics (Lessani & Li, 2024).

$$BP = \frac{1}{2}[f^T Z(Z^T Z)^{-1} Z^T f] \sim \chi_p^2$$
$$f_i = \left(\frac{\varepsilon_i^2}{\sigma^2} - 1\right) \tag{2}$$

where ε_i is the error for the i-th observation and σ^2 is the variance of the model; Z is a constant matrix with size $n \times (k+1)$. Heteroscedasticity occurs when there is a systematic pattern in the residual variance. The resulting test statistic follows the Chi-Square distribution. The null hypothesis (H_0) states that there is no heteroscedasticity (the residual variance is constant), while the alternative hypothesis (H_1) explains that heteroscedasticity occurs. There is heteroscedasticity if the BP value$>\chi_{\alpha,p}^2$ or p-value $<$ 0.05 (Ispriyanti et al., 2018). This step helps to keep the standard errors and significance tests in the severe stunting model statistically reliable.

c. Spatial Autocorrelation

Spatial analysis has developed to help solve various problems in various sectors by considering geographical conditions. One approach in this analysis is to use spatial autocorrelation to detect the geographical influence of several observations. According to Tobler's First Law, all things are interrelated, but things that are close together are more strongly related than those that are far apart (Zheng et al., 2023). In spatial analysis, a spatial weight matrix shows the relationship between one location and another. This matrix represents the spatial relationship between locations and can be adjusted to the analysis method (Moraga, 2024; Pebesma & Bivand, 2023).

The type of spatial weight chosen depends on the modelling needs and the conditions of the analysed area. Generally, there are two basic weighting categories, namely, distance and tangency (Moraga, 2024). Types of weighting based on distance include (1) inverse

distance (IDW), (2) k-nearest Neighbour (k-NN), and (3) Critical Cut of Neighbourhood. Types of weighting based on tangency include (1) Queen Contiguity, (2) Rock Contiguity, and (3) Bishop Contiguity. Three approaches can be used to measure spatial dependence between variables: Moran's Index, Geary's C, and Tango's Excess (Pfeiffer et al., 2008). In this study, we chose to use Moran's Index because of its simplicity in computing and its advantages in measuring global spatial dependence(Zheng et al., 2023).
The following is the procedure for evaluating spatial dependence between regions: (1) the spatial weight matrix is calculated using the IDW and queen contiguity methods, (2) the weight matrix is standardized by ensuring that the total weight in each row is one, and (3) the Moran Index is used with the standardized spatial weight matrix to test for the existence of spatial dependence, as shown in the following equation.

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\left((X_i - \bar{X})(X_j - \bar{X})\right)}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \tag{3}$$

The Moran's Index (I) ranges from [−1,1], where n represents the total number of observations (locations), X_i and X_j are the observed values at locations i and j, respectively, X is the mean of the observed values, and w_{ij} is the standardized spatial weight matrix. The null hypothesis for testing spatial dependency using Moran's Index states that there is no spatial autocorrelation between X_i and X_j. The null hypothesis is rejected if the test statistic $|Z_i| < Z_{\frac{\alpha}{2}}$ or if $p - val < \alpha$ (Mathur, 2015). The two observed variables are said to have a grouping pattern if $I > E[I]$, have a distribution pattern if $I > E[I]$, and have an uneven distribution pattern if $I = E[I]$. Here, Moran's I is used to check whether a non-spatial count regression model is adequate for the severe stunting data across districts and cities.

d. Poisson Regression

The severe stunting case counts data resembles the data for Poisson distribution, i.e. in the form of integers and non-negative. Poisson regression in the Generalized Linear Model (GLM) concept uses the log function or logarithm based on Euler numbers as a link function to connect the response variable expectation value with a linear combination of explanatory variables.

$$log\left(\frac{\hat{\mu}_i}{m_i}\right) = \sum_{j=1}^{k} \hat{\beta}_j X_{ij} \tag{4}$$

Poisson regression has an assumption, namely the assumption that the mean is equal to the variance. This assumption comes from the character of the Poisson distribution, which is the basis of the Poisson regression model (Dzupire et al., 2018). If the mean is not equal to the variance, also known as under or over dispersion, it will produce a very small p-value and results that are not meaningful to reality (Fávero et al., 2021). In the case of overdispersion, the variance is more significant than the mean, thus alternative models based on different distribution, for example Negative Binomials, can be used. In this study, Poisson regression with an offset is first specified to relate severe stunting counts to the explanatory variables.

e. Overdispersion Test

Overdispersion tests in count data analysis often use deviance and chi-square statistics to determine whether the data have more significant variability than predicted by the model. Overdispersion can lead to inference errors if not addressed. It could be indicated by comparing the deviance or Pearson Chi-Square statistic with the degrees of freedom (df). When the overdispersion ratio to df exceeds one, the data is considered overdispersion (Payne et al., 2018). The following is the equation for calculating the deviance statistic (D):

$$D = 2\sum_{i=2}^{n}\left[y_i \log\left(\frac{y_i}{\hat{\mu}_i}\right) - (y_i - \hat{\mu}_i)\right] \quad (5)$$

where y_i is the observed value, and $\hat{\mu}_i$ is the model's predicted value. In addition to using deviance, overdispersion can be identified using the Chi-Square approach. (χ^2_{n-k}) (Gelhan & Hill, 2007):

$$z_i = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i}} \quad (6)$$

$$Estimated\ Ovredispersion = \frac{1}{n-k}\sum_{i-1}^{n} z_i^2 \quad (7)$$

Here, k represents the number of independent variables, and n is the number of observations. These two statistics are compared across degrees of freedom to test whether the data have significant excess variance (Afroz, 2024). This method is critical in Generalized Linear Models (GLMs), in which the Poisson distribution assumption is often used for count data. If overdispersion is detected, other models, such as Negative Binomial regression model, are often more appropriate because it can handle overdispersion naturally. Proper handling of overdispersion ensures valid and more reliable statistical inference results (Hossain et al., 2020). The overdispersion diagnostic therefore guides the decision to retain the Poisson model or to move to the Negative Binomial model for severe stunting.

f. Negative Binomial Regression

Stunting data characterized as count data and modelled with Poisson regression, prone to have overdispersion issues due to its variance estimation larger than its mean (Fávero et al., 2021). Dobson & Barnett (2018) suggest using Negative Binomial model instead Poisson model to handle overdispersion, while shows how Negative Binomial distribution is yielded from allowing variance of Poisson distribution in Y

$\left(Y|\lambda \sim Poisson(\lambda)\right)$ varies following Gamma distribution $\left(\lambda \sim G(\mu,\gamma)\right)$ (Cameron & Trivedi, 2013). The Negative Binomial model with k explanatory variables and an offset variable (m) is expressed by:

$$log\left(\frac{\hat{\mu}_i}{m_i}\right) = \sum_{j=1}^{k} \hat{\beta}_j X_{ij} \text{, for i } = 1, 2, \ldots, \text{n} \quad (8)$$

In the case that the intercept should be in the model, the Equation 8 can be written as following:

$$\hat{\mu}_i = exp\left[\log(m_i) + \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \cdots + \hat{\beta}_k X_{ik}\right] \quad (9)$$

The central and dispersion parameter, as well as probability mass function for Negative Binomial distribution consecutively are given by (Putri Ananda et al., 2024) as follows:

$$E(Y) = \mu;$$
$$Var(Y) = \mu + \gamma\mu^2;$$
$$f(Y|\mu,\gamma) = \frac{\Gamma\left(Y+\frac{1}{\gamma}\right)}{\Gamma\left(\frac{1}{\gamma}\right)\text{Y!}}\left(\frac{1}{1+\gamma\mu}\right)^{\frac{1}{\gamma}}\left(\frac{\gamma\mu}{1+\gamma\mu}\right)^{Y} \text{ for } Y = 0, 1, 2, \ldots$$

Should we assume that the dispersion parameter of Negative Binomial distribution in Y are unknown but equal for all districts in Central Java Province, then its log-likelihood function can be written as follow:

$$\begin{aligned} \mathcal{L}(\beta,\gamma;Y) = &\sum_{i=1}^{n}\left[log\Gamma\left(Y_i + \frac{1}{\gamma}\right) - log\Gamma\left(\frac{1}{\gamma}\right) - log\Gamma(Y_i+1)\right] \\ &+ \sum_{i=1}^{n}\left[Y_i log\left(\frac{\gamma\mu_i}{1+\gamma\mu_i}\right) - \left(\frac{1}{\gamma}\right) log(1+\gamma\mu_i)\right] \end{aligned} \quad (10)$$

Using Equation 10, it can be shown that the derivatives of log-likelihood function are:

$$\begin{aligned} \frac{\partial\mathcal{L}(\beta,\gamma;Y)}{\partial\beta_j} &= \sum_{i=1}^{n}\frac{x_{ij}(Y_i-\mu_i)}{1+\gamma\mu_i}, \text{for } j = 1, 2, \ldots, k \\ \frac{\partial\mathcal{L}(\beta,\gamma;Y)}{\partial\gamma} &= \sum_{i=1}^{n}\left\{\frac{1}{\gamma^2}\left(log(1+\gamma\mu_i) - \sum_{j=0}^{Y_i-1}\frac{1}{j+\gamma^{-1}}\right) + \frac{Y_i-\mu_i}{\gamma(1+\gamma\mu_i)}\right\} \end{aligned} \quad (11)$$

while the second derivatives are given by:

$$\frac{-\partial^2\mathcal{L}(\beta,\gamma;Y)}{\partial\beta_r\partial\beta_s} = \sum_{i=1}^{n}\frac{\mu_i(1+\gamma Y_i)x_{ir}x_{is}}{(1+\gamma\mu_i)^2}, \text{for } r,s = 1,2,\dots,k$$

$$\frac{-\partial^2\mathcal{L}(\beta,\gamma;Y)}{\partial\beta_r\partial\gamma} = \sum_{i=1}^{n}\frac{\mu_i(Y_i-\mu_i)x_{ir}}{(1+\gamma\mu_i)^2}, \text{for } r = 1,2,\dots,k$$

$$\frac{-\partial^2\mathcal{L}(\beta,\gamma;Y)}{\partial\gamma^2} = \sum_{i=1}^{n}\left\{\sum_{j=0}^{Y_i-1}\left(\frac{j}{1+\gamma j}\right)^2 + \frac{2log(1+\gamma\mu_i)}{\gamma^3} - \frac{2\mu_i}{\gamma^2(1+\gamma\mu_i)} - \frac{(Y_i+\gamma^{-1})\mu_i^2}{(1+\gamma\mu_i)^2}\right\} \tag{12}$$

Using both derivatives, one can find the estimate of β_j for $j = 1,2,\dots,k$, as well as estimation of γ. This model is then adopted as the main tool to estimate the association between each determinant and the rate of severe stunting.

g. Best Model Selection

Akaike Information Criterion (AIC) is a value used in statistical model selection to assess the balance between the model's fit to the data and the complexity of the model. AIC was introduced by Hirotugu Akaike in 1973 to extend the principle of maximum likelihood into the model selection process. The AIC formula is:

$$AIC = -2\log(L) + 2k \tag{13}$$

Here, L is the value of likelihood function from a model provided with β_j for $j = 1,2,\dots,k$, and k is the number of parameters estimated in the model. The concept considers model's goodness of fit $(-2\ln(L))$ while also adhere to parsimonious model preferences to avoid overfitting due to a complex model $(2k)$ (Cavanaugh & Neath, 2019). In this study, AIC is used in a backward elimination procedure to obtain a parsimonious and interpretable model for severe stunting cases.

2. Data and Research Variable

Study of severe cases of stunting in children in Central Java Province in this paper is using the data taken from the Ministry of Home Affairs and Statistics Indonesia. These official statistics ensure consistent definitions and measurement across all districts and cities in the province. The data used for response variable (Y) is the number of stunting children with very short stature in 2023 from all districts and cities in Central Java Province. This definition follows the national standard for severe stunting based on WHO child growth charts. This data was then explored with six explanatory variables and one offset variable. The explanatory variables were (1) incidence of CED in mother who had nutritious food support, (2) percentage of 6 to 23 months old toddler who receive nutritious complementary food aside their mother's breast milk, (3) percentage of households with access to good waste and sanitation system, (4) districts and cities domestic GDP per capita, (5) Human Development Index, and (6) numbers of local health facility unit (Pusat Kesehatan Masyarakat, abbr. Puskesmas). The offset variable was numbers of toddlers under 5 years in each district and cities in Central Java Province in 2023. Unless stated otherwise, all the data was taken from year 2023 records. Information regarding the variables and its description is presented in Table 1. Together with the offset, these variables allow the modelling of severe stunting as a rate rather than only as raw counts, which is important for fair comparison across districts and cities.

This study employs a quantitative, non-experimental observational research design using cross-sectional secondary data. The data source of severe stunting cases, as well as the offset and the explanatory variables X_1, X_2, and X_3, were taken from the Directorate General of Regional Development, Ministry of Home Affairs, Indonesia (Directorate General of Village Government Development, 2024). The explanatory variable X_4, and X_5 were sourced from Statistics Indonesia (Badan Pusat Statistik Jawa Tengah, 2024), while the last explanatory variable, X_6, were available from the Directorate General of Villages Governance, Ministry of Home Affairs, Indonesia (Directorate General of Village Government Development, 2024). These institutions are mandated by the government to monitor child nutrition and socio-economic indicators, supporting the validity of the data used in this study. The cross-sectional design at the district/city level is appropriate for identifying contextual determinants of severe stunting in Central Java in 2023. All statistical analyses, including data processing, overdispersion testing, and the estimation of the Negative Binomial regression model, were carried out using R statistical software (version R 4.4.2) with appropriate packages for count data modelling, as shown in Table 1.

Table 1. Variables and data for severe stunting model in Central Java Province

	Variable	Unit	Description
Y	Severe stunting cases	Child	Numbers of children with severe stunting symptoms in district/city in Central Java Province
X_1	Percentage mother with CED and got support	%	Proportion of pregnant mother given nutritious supplement due to having CED symptoms
X_2	Percentage of toddlers with food complement	%	Percentage of aged 6-23 months toddlers received nutritious complementary food alongside mother's breast milk
X_3	Percentage of household with good access to sanitation	%	Percentage of households with good sanitation system
X_4	GDP per capita	Rupiah	Gross Domestic Product per people in certain area, calculated based on market price, in millions Rupiah
X_5	Human Development Index	Index	Value of human development progress in a district/city, considering aspect from health, education, and adequate standard of living
X_6	Numbers of local health facility	Unit	The amount of healthcare facility unit in a district/city
Offset	Numbers of children under 5 year	Child	Numbers of under-5-years old children in district/city

3. Research Steps

Study of severe stunting cases in Central Java Province using variables mentioned in previous section will be done through the Figure 1 below.

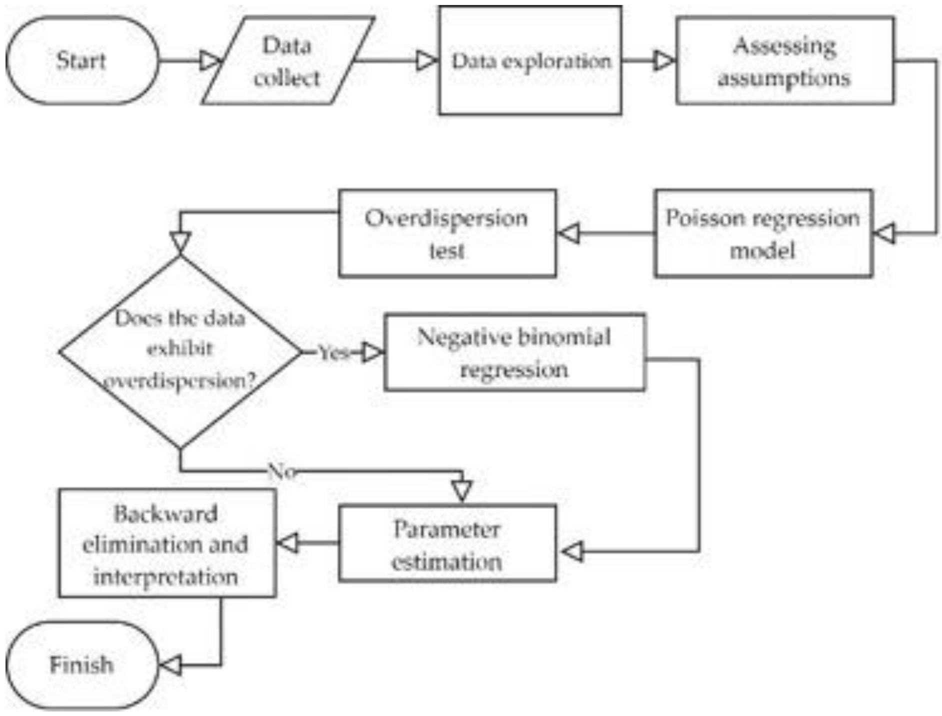

Figure 1. Research Steps Flowchart

The explanation of the flowchart based on Figure 1 is as follows.

a. Determine all the research variables (including the offset variable) and collect the associated data. The data used in the study were taken from the Ministry of Home Affairs and Statistics Indonesia with the scope only in Central Java Province and only data labeled year 2023. The Central Java Province itself has 29 districts and 6 cities. The data will then be represented by the districts and cities in Central Java Province.
b. Data taken from different sources will then be pre-processed to maintain the consistency of all districts and cities order.
c. By using thematic map representation, the response variable then explored to better understand the status of severe stunting cases in Central Java Province.
d. Equivalent with other parametric statistical model, classic model prerequisites then checked as a basis for model building given the data. The assumptions checking done to the data will be:
    1) Multicollinearity test for all explanatory variables. The test will be done by observing the Variance Inflation Factor (VIF) of all explanatory variables.
    2) The constant variance of error term (homoscedasticity) test. The Breusch-Pagan homoscedasticity test will be used to check the presence of heteroscedasticity effect in the data.
    3) The spatial independency test amongst all districts and cities will also be employed using Moran's Index of Independency test, with Inverse Distance Weighted (IDW) and Queen Contiguity algorithm.

e. Based on the conclusions of assumptions checking, the Poisson regression model then developed for the given data. Here, one should not overlook the equi-dispersion assumption of discrete variable model, as in Poisson model. Thus, the consecutive step is to check the apparent effect of overdispersion in the model.
f. The overdispersion test based on residual deviance of the model will be done through hypothesis testing. Observation of the Wald test result of each parameter estimates in the model will also be done to reinforce conclusion made by the significance test.
g. Should overdispersion indeed occur in the Poisson model, the study will switch to the Negative Binomial regression model as an effort to eliminate the overdispersion effect existed in previous model.
h. The parameter estimates and its Wald test was then examined to determine its significance in the model. Should there are any estimates which non statistically significant to explain the variability of response variable, re-model the data then should be done without the no-effect explanatory variables.
i. Backward elimination procedure then employed to select the best model for the data. The elimination criteria used will be based on AIC value. The best model should be having the lowest AIC value with all of its explanatory variables are statistically significant.
j. The last step will be interpreting the best model and make suggestions based on its conclusion.

C. RESULT AND DISCUSSION

1. Data Exploration

In regard to better understand severe stunting cases in Central Java Province, a thematic map, presented in Figure 2, was explored in ahead of model building. Based on Figure 2, the darker the colour indicates, the more cases of severe stunting there are in the province, and vice versa. Magelang City has the fewest cases of stunting, as are the surrounding districts, namely Semarang Regency, Salatiga City, and Boyolali Regency. Meanwhile, the area that includes the highest stunting category is the Tegal Regency. When viewed from Figure 2, the areas neighbouring Tegal Regency include Brebes Regency, Purbalingga Regency, Banjarnegara Regency, and Banyumas Regency. This value is linear with the number of toddlers in each region; namely, the number of toddlers in Magelang City is 4938 (prevalence 7.7%), while Tegal Regency has 100252 toddlers (prevalence 17.7%).

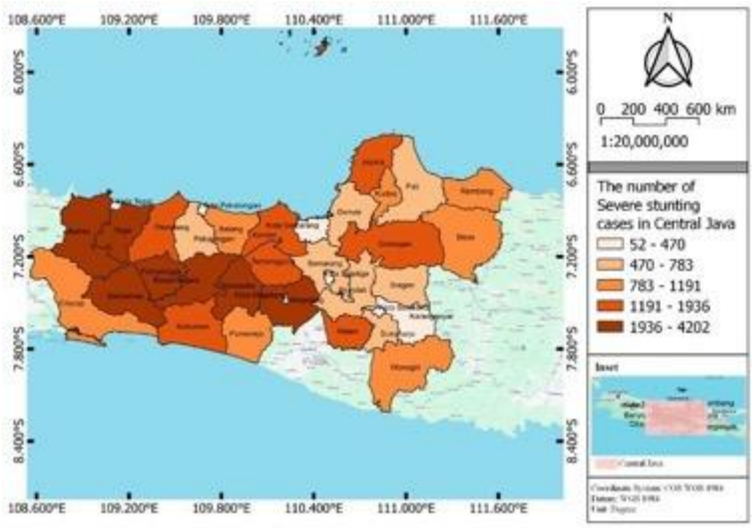

Figure 2. Distribution of severe stunting cases in Central Java Province in 2023

The large number of stunting cases is inseparable from the cases of early marriage in the Tegal Regency in 2023, which should be the minimum age for marriage of 19 years (Law of the Republic of Indonesia Number 16 of 2019 on the Amendment to Law, 2019)**.** Early marriage is a marriage that occurs at a vulnerable age whose preparation targets are said to be suboptimal both physically, mentally and materially, while what is needed in marriage is mental maturity and good mental readiness (Rohmaniah et al., 2025). In cases of early marriage, many impacts occur, one of which is the low quality of the family, both in terms of physical readiness in dealing with social and economic problems of the household, as well as the physical readiness of prospective teenage mothers in conceiving and giving birth to their children (Albab & Pratiwirum, 2022; Subramanee et al., 2022; Yoosefi Lebni et al., 2023). Based on Statistics Indonesia data of 2023, 17.51% of women aged <17 years were recorded as having early marriages, and 20% of women aged 17-18 were having early marriages (BPS-Statistics Indonesia Jawa Tengah Province, 2021). There have been many studies related to preventing early marriage (Phillips & Mbizvo, 2016; Sandøy et al., 2016; Silumbwe et al., 2020).

Although Tegal Regency has the highest number, the Tegal Regency Government has made various efforts to reduce stunting rates, including through the stunting reduction acceleration team (TPPS) at various levels, which is currently carrying out joint actions to simultaneously intervene in stunting prevention by conducting data collection, weighing, measuring, education and intervention for all prospective brides, pregnant women and toddlers on an ongoing basis **[51].** One of the results is a significant decrease in the prevalence of stunting in the district, with severe stunting figures in 2023 of 4202 children decreasing to 3311 children (as of early December 2024).

2. Multicollinearity Test

This study uses the Variance Inflation Factor (VIF) to identify multicollinearity between predictor variables. Based on **Equation (1),** the results of the VIF calculation are presented in Table 2.

Table 2. The VIF value

Variable	X_1	X_2	X_3	X_4	X_5	X_6
VIF Value	2.1875	1.1064	1.4385	3.5937	3.6386	1.8252

Based on the results in Table 2, all VIF values of the independent variables are below the threshold of 5, indicating no serious multicollinearity. The highest VIF value is obtained by variable X_5 of 3.6386, while the lowest VIF value is 1.1064 in variable X_2. Overall, the VIF values ranging from 1.106 to 3.6386 indicate that the correlation between the independent variables is still acceptable so that multicollinearity does not disturb the regression parameter estimates.

3. Breusch–Pagan test

Heteroscedasticity testing is performed using the Breusch-Pagan test to determine whether the residual variance is not constant and varies according to the value of the independent variable. Based on Equation 2, the BP test statistic value is presented in Table 3.

Table 3. Breusch-Pagan test

Breusch–Pagan test	**Degree of Freedom**	**P-Value**
8.1734	6	0.2257

Based on Table 3, the BP value is 8.1734. This value is much smaller when compared to $\chi^2_{0{,}05{,}6}$ which is 12.59. Moreover, the p-value obtained is 0.2257, which is greater than a significance level of $\alpha = 0.05$. Hence, based on the BP value and p-value, the null hypothesis should not be rejected. This indicates that there is not enough evidence to conclude that there is heteroscedasticity in the regression model. Thus, the assumption of homoscedasticity is still met in this model.

4. Spatial Autocorrelation

The next test is the spatial autocorrelation test using Moran's Index test. The hypothesis built into this test is that there is no spatial dependency amongst observations, which in this study is the districts and cities, while the alternative hypothesis is that there is spatial dependency. Based on the Literature Review section, 2 types of weights are used in calculating Moran's index, namely Queen Contiguity and IDW. Table 4 is the result of the calculation of the Moran's Index according to **Equation 3.**

Table 4. Moran`s Index test

No	**Weighting Matrix**	**Moran's I Statistics**	$Z(I)$	**P-Value**	**Conclusion**
1	IDW	0.0216	1.3115	0.1897	There is no spatial autocorrelation
2	Queen Contiguity	0.1481	1.6019	0.1092	There is no spatial autocorrelation

Following Table 4, the results show no spatial autocorrelation, meaning that observations at a location do not depend on other neighbouring locations.

5. Poisson Regression

Poisson regression is a good approach based on exploring severe stunting data in Central Java Province. The explanatory variables in the Poisson regression modelling are X_1 to X_6; these variables are proven to affect severe stunting (Y). Furthermore, parameter estimation uses the MLE approach to determine the factors that affect the stunting; the estimation results are in Table 5.

Table 5. Parameter Estimation of Poisson Regression

Parameters	Estimate	Std. Error	Wald Statistics (Z)	P-Value
β_0	3.14200	0.158000	19.89	0.000*
β_1	-0.01717	0.005001	-34.33	0.000*
β_2	-0.01016	0.004506	-22.54	0.000*
β_3	-0.01101	0.003904	-28.20	0.000*
β_4	-0.01057	0.003419	-30.91	0.000*
β_5	-0.04118	0.023480	-17.84	0.000*
β_6	-0.009308	0.029690	-31.35	0.000*

*) Significant with alpha= 0.05

Based on Table 5, all variables are statistically significant and affect severe stunting in Central Java Province. However, the Poisson regression model must be checked for overdispersion to confirm the correct estimation results. Overdispersion can cause inefficient estimation, making p-value more minor than the actual value. Table 5 shows that the p-value are very small, so it can be suspected that overdispersion has occurred.

6. Overdispersion Test

Poisson regression assumes that the mean value is equal to the variance; if this assumption is not met, then the model has overdispersion or underdispersion. Overdispersion/underdispersion can cause the model to be unaccountable for its results. Testing for overdispersion/underdispersion in the Poisson regression model can be done by looking at the results of the deviance value divided by the degrees of freedom (df); if it produces a value of more than one, then overdispersion occurs (Fávero et al., 2021). The Chi-Square approach divided by the number of observations minus the number of response variables (dispersion ratio) can also be used as an overdispersion test (Fávero et al., 2021).

Table 6. Overdispersion Test

Dispersion Ratio	Pearson's Chi-Squared	P-Value	Residual Deviance	Degrees of Freedom (df)	Residual Deviance / (df)
109.487	3065.635	0.001	3267.7	28	116.7036

The division of residual deviation by degrees of freedom yields a value of 116.7036; overdispersion is detected. In addition, the dispersion ratio value (see Table 6) also shows more than one; this confirms that the Poisson model suffers from overdispersion. The thought that the p-value (see Table 5) are minimal due to overdispersion in the Poisson model is confirmed. The results of the Poisson regression cannot be accounted for; this can lead to misleading interpretations. One solution to overcome overdispersion in Poisson regression is to model it with a Negative Binomial approach (Fávero et al., 2021).

7. Negative Binomial Regression

The severe stunting data in Central Java Province has also been modelled using the Negative Binomial model to address overdispersion problems in the Poisson model. Here the dispersion ratio is 0.875 which is closer to 1 compared to the previous model (see Table 8). The result indicates that the negative binomial model for Central Java Province's severe stunting data does not have overdispersion problem. Table 7 presents parameters model estimation and its Wald statistics for negative binomial model. By using 5% significance reference level, it can be seen that the estimate of model coefficients for X_1 (β_1), X_3 (β_3), X_4 (β_4), and X_6 (β_6) are statistically proven to be different from 0. Furthermore, should we enlarge the significance reference level to 10%, then the estimation of coefficients model for X_2 (β_2) is also statistically significant. Since the overdispersion problem had been tackled by remodel the data using the Negative Binomial model, backward elimination procedure using AIC value as elimination criteria can then be used to formulate the best model for the data.

Table 7. Parameters estimation of Negative Binomial model

Parameters	Estimate	Std. Error	Wald Statistics (Z)	P-Value
β_0	1.2252699	1.4585891	0.840	0.40089
β_1	-0.0099904	0.0048039	-2.080	0.03756*
β_2	-0.0130879	0.0067236	-1.947	0.05159*
β_3	-0.0120873	0.0046531	-2.598	0.00939*
β_4	-0.0101842	0.0033425	-3.047	0.00231*
β_5	-0.0196195	0.0239067	-0.821	0.41183
β_6	-0.0005248	0.0002635	-1.991	0.04644*

*) Significant with alpha= 0.05

Backward elimination procedure to determine the best model with all significant explanatory variables is done by finding non-significant β in the model, then omit the variable in the next model building. In the case of severe stunting in children in Central Java Province, previous Negative Binomial model has hinted that the β_5 was unable to be proven statistically significant even with a significance level of 10%. Thus, this variable is the main focus to be excluded in the backward elimination procedure. The performance of the models was then compared, and the best model will be selected based on its AIC value as well as whether all of β included in the model are statistically significant. Table 8 presented a comparison of models which was considered in the backward elimination process. Following the Table, it is apparent that the lowest AIC value is held by the Negative Binomial model without β_5.

Table 8. Backward Elimination based on Akaike's Information Criterion (AIC)

Model name	Number of parameters	Variable X included	Residual Deviance	Dispersion Ratio	AIC
Poisson Regression	7	$X_1, X_2, X_3, X_4, X_5, X_6$	3267.7	109.487	3580.9
Negative Binomial (initial model)	7	$X_1, X_2, X_3, X_4, X_5, X_6$	35.678	0.875	506.84
Negative Binomial (without β_5)	6	X_1, X_2, X_3, X_4, X_6	35.63982	0.925	505.48

Table 9 provides estimation of coefficients should the severe stunting data in Central Java Province be modelled with the best Negative Binomial model. The model provided could be termed as the best model given the data due to none of its coefficient's estimation model were statistically non-significant under the 5% significance level. The model also has the lowest AIC value amongst other considered models.

Table 9. Parameters estimation of the Best Negative Binomial model

Parameters	Estimate	Std. Error	Wald Statistics (Z)	P-Value
β_0	0.2479785	0.8133	0.305	0.76044
β_1	-0.0109246	0.0047	-2.310	0.02086
β_2	-0.0143183	0.0066	-2.165	0.03035
β_3	-0.0138514	0.0042	-3.258	0.00112
β_4	-0.0125444	0.0019	-6.465	0.00000
β_6	-0.0005492	0.0003	-2.090	0.03664

*) Significant test with alpha= 0.05

The best model can then be expressed by:

$$log\left(\frac{\hat{\mu}_i}{m_i}\right) = 0.2479785 - 0{,}0109246X_1 - 0.0143183X_2 - 0.0138514X_3 - 0.0125444X_4 - 0.0005492X_6 \quad (14)$$

By using the model, which accounted for the population size of children under 5 years in each district, we may conclude that the rates of severe stunting cases in Central Java Province have an inverse relationship with X_1, X_2, X_3, X_4, and X_6. In other terms, we may say that the cases will be increasing if any or all of the X's are decreasing. For examples regarding the X_1, here the decrease of the percentage of pregnant mothers who have chronic energy deficiency (CED) and have been given nutritious food supplements will directly increase the rates of severe stunting cases as much as 1.1% if the other X's are unchanged. This conclusion is acceptable due to num (Asna & Syah, 2023; Putri Adila et al., 2023; Putri Ananda et al., 2024; Sajalia et al., 2018; Yuliastanti et al., 2023)(Asna & Syah, 2023; Putri Adila et al., 2023; Putri Ananda et al., 2024; Sajalia et al., 2018; Yuliastanti et al., 2023)) stating that the CED on pregnant mothers is one of the risk factors of malnutrition cases in infants. Thus, providing the mothers with nutritious food supplements will reduce the risk which eventually may also reduce the cases of severe stunting. Another case is for X_2, if the proportion of toddlers receiving complementary food is decreasing by a percent, then we should expect that the severe stunting rates would be increasing by 1.4%. This change in rate is approximately the same as a unit change in the proportion of households with good sanitation (X_3). GDP per capita (X_4) on the other hand, is associated with an addition of 1.25% in the severe stunting rate. The last factor, X_6, will decrease the severe stunting rates by 0.06% should there are a new unit of Puskesmas built in Central Java. The change of rates for this factor is the lowest amongst other factors in this study. Thus, adding a unit of *Puskesmas* might not be effective in reducing the severe stunting rates. We would recommend firstly evaluating the quality of services of existing *Puskesmas* as our model only captures the quantity of *Puskesmas* instead of its quality.

D. CONCLUSION AND SUGGESTIONS

This study makes a distinct methodological and practical contribution to understanding the determinants of severe stunting in Central Java. Methodologically, we demonstrate that modelling count data like stunting cases requires careful consideration of dispersion. The successful application of the Negative Binomial regression model, which corrected for overdispersion present in a standard Poisson model, provides a more robust and reliable statistical framework for public health researchers analyzing similar prevalence data. Our empirical analysis robustly identifies five key factors significantly associated with the reduction of severe stunting cases. The strong statistical association of nutrition-specific interventions targeting pregnant women with Chronic Energy Deficiency (CED) and toddlers aged 6-23 months underscores the non-negotiable role of targeted nutritional support during the first 1000 days of life. Furthermore, the significance of underlying socio-economic and environmental drivers is clear: household sanitation, local GDP per capita (as a proxy for community wealth), and the density of local health facilities (*Puskesmas*) are foundational to creating an environment where children can thrive. These findings offer a clear, evidence-based roadmap for local policymakers. To effectively mitigate severe stunting, interventions must be multi-sectoral, namely we need to: (1) strengthen the health sector by prioritizing the distribution of nutritious food supplements for at-risk pregnant women and ensure the provision of complementary foods for toddlers, primarily through the existing network of *Puskesmas*; (2) invest in sanitation and economic development, especially by allocating resources to improve household sanitation infrastructure and implement programs that boost local economic growth and household income, recognizing their critical role in child health; and (3) and leverage healthcare access due to the significant influence of *Puskesmas* availability which could be a highly effective strategy for reaching vulnerable populations. For future research, this study establishes a validated statistical approach. Building upon this work, subsequent studies could incorporate additional variables such as local food inflation and government social spending. Furthermore, employing a hierarchical model could unravel the complex interplay between individual-level risk factors and district-level policies, offering even more nuanced guidance for targeted public health action.

REFERENCES

Afroz, F. (2024). Proposing a New Estimator of Overdispersion for Multinomial Data. *Dhaka University Journal of Science*, *72*(1), 56–62. https://doi.org/10.3329/dujs.v72i1.71247

Albab, F. U., & Pratiwirum, Y. (2022). Efektivitas Pencegahan Pernikahan Usia Dini Pada Badan Kependudukan Keluarga Berencana Nasional (Bkkbn) Kabupaten Sorong. *Muadalah : Jurnal Hukum*, *2*(2), 115–124. https://doi.org/10.47945/muadalah.v2i2.757

Asna, A. F., & Syah, Muh. N. H. (2023). Chronic Energy Malnutrition in Mothers Associated with Stunting. *Jurnal Gizi Dan Dietetik Indonesia (Indonesian Journal of Nutrition and Dietetics)*, *11*(2), 77. https://doi.org/10.21927/ijnd.2023.11(2).77-84

Badan Pusat Statistik Jawa Tengah. (2024). *PDRB Atas Dasar Harga Berlaku Menurut Kabupaten/Kota di Provinsi Jawa Tengah (Juta Rupiah), 2022-2023*. Https://Jateng.Bps.Go.Id/Id/Statistics-Table/2/MTc0MCMy/-Seri-2010--Pdrb-Atas-Dasar-Harga-Berlaku-Menurut-Kabupaten-Kota-Di-Provinsi-Jawa-Tengah--Juta-Rupiah-.Html.

BPS-Statistics Indonesia Jawa Tengah Province. (2021). *Percentage of Female Population Aged 10 Years and Over who Have Ever Been Married by Area of Residence and Age of First Marriage (Percent), 2018*. Https://Jateng.Bps.Go.Id/En/Statistics-Table/2/OTc3IzI=/Percentage-of-Female-Population-Aged-10-Years-and-over-Who-Have-Ever-Been-Married-by-Area-of-------Residence-and-Age-of-First-Marriage.Html.

Cameron, A. C., & Trivedi, P. K. (2013). Regression Analysis of Count Data. In *Econometric Society Monographs* (2nd ed.). Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9781139013567

Cavanaugh, J. E., & Neath, A. A. (2019). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *WIREs Computational Statistics*, *11*(3), e1460. https://doi.org/https://doi.org/10.1002/wics.1460

da Silva, A. R., & Rodrigues, T. C. V. (2014). Geographically Weighted Negative Binomial Regression—incorporating overdispersion. *Statistics and Computing*, *24*(5), 769–783. https://doi.org/10.1007/s11222-013-9401-9

Dewi, A., Lanti Retno dewi, Y., & Murti, B. (2019). Life Course Factors Associated with Stunting in Children Aged 2-5 Years: A Path Analysis. *Journal of Maternal and Child Health*, *4*, 348–357. https://doi.org/10.26911/thejmch.2019.04.05.09

Directorate General of Village Government Development, M. of H. A. (2024). *Posyandu Development Data*.

Dobson, A. J., & Barnett, A. G. (2018). *An Introduction to Generalized Linear Models*. Chapman and Hall/CRC. https://doi.org/https://doi.org/10.1201/9781315182780

Dzupire, N. C., Ngare, P., & Odongo, L. (2018). A Poisson-Gamma Model for Zero Inflated Rainfall Data. *Journal of Probability and Statistics*, *2018*. https://doi.org/10.1155/2018/1012647

Fávero, L. P., de Freitas Souza, R., Belfiore, P., Corrêa, H. L., & Haddad, M. F. C. (2021). Count Data Regression Analysis: Concepts, Overdispersion Detection, Zero-inflation Identification, and Applications with R. *Practical Assessment, Research and Evaluation*, *26*, 1–22. https://doi.org/10.7275/44nn-cj68

Fernandez, G. A., & Vatcheva, K. P. (2022). A comparison of statistical methods for modeling count data with an application to hospital length of stay. *BMC Medical Research Methodology*, *22*(1), 211. https://doi.org/10.1186/s12874-022-01685-8

Fitriyah, H., Kurnia, A., & Afendi, F. M. (2015). Negative Binomial Regression Methods To Analyze Factors Affecting Child Mortality Rates in West Java. *International Conference On Research, Implementation And Education Of Mathematics And Sciences*, 17–19.

Food and Agriculture Organization. (2024). *The State of Food Security and Nutrition in the World 2024*.

Gelhan, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Hossain, Z., Akter, R., Sultana, N., & Kabir, E. (2020). Modelling zero-truncated overdispersed antenatal health care count data of women in Bangladesh. *PLoS ONE*, *15*(1). https://doi.org/10.1371/journal.pone.0227824

Htet, M. K., Do, T. T., Wah, T., Zin, T., Hmone, M. P., Raihana, S., Kirkwood, E., Hlaing, L. M., & Dibley, M. J. (2023). Socio-economic and agricultural factors associated with stunting of under 5-year children:

Findings from surveys in mountains, dry zone and delta regions of rural Myanmar (2016-2017). *Public Health Nutrition*, *26*(8), 1644–1657. https://doi.org/10.1017/S1368980023001076

Humphrey, J. H., Mbuya, M. N. N., Ntozini, R., Moulton, L. H., Stoltzfus, R. J., Tavengwa, N. V, Mutasa, K., Majo, F., Mutasa, B., Mangwadu, G., Chasokela, C. M., Chigumira, A., Chasekwa, B., Smith, L. E., Tielsch, J. M., Jones, A. D., Manges, A. R., Maluccio, J. A., Prendergast, A. J., … Makoni, T. (2019). Independent and combined effects of improved water, sanitation, and hygiene, and improved complementary feeding, on child stunting and anaemia in rural Zimbabwe: a cluster-randomised trial. *The Lancet Global Health*, *7*(1), e132–e147. https://doi.org/10.1016/S2214-109X(18)30374-7

Ispriyanti, D., Prahutama, A., & Taryono, A. P. N. (2018). Modelling space of spread Dengue Hemorrhagic Fever (DHF) in Central Java use spatial durbin model. *Journal of Physics: Conference Series*, *1025*(1). https://doi.org/10.1088/1742-6596/1025/1/012112

Johri, M., Subramanian, S. V, Koné, G. K., Dudeja, S., Chandra, D., Minoyan, N., Sylvestre, M.-P., & Pahwa, S. (2016). Maternal Health Literacy Is Associated with Early Childhood Nutritional Status in India. *The Journal of Nutrition*, *146*(7), 1402–1410. https://doi.org/https://doi.org/10.3945/jn.115.226290

Kalinda, C., Qambayot, M. A., Ishimwe, S. M. C., Regnier, D., Bazimya, D., Uwizeyimana, T., Desie, S., Rudert, C., Gebremariam, A., Brennan, E., Karumba, S., Wong, R., & Bekele, A. (2024). Leveraging multisectoral approach to understand the determinants of childhood stunting in Rwanda: a systematic review and meta-analysis. *Systematic Reviews*, *13*(1), 16. https://doi.org/10.1186/s13643-023-02438-4

Khan, M. M. A., Billah, M. A., Fatima, K., Islam, M. M., Sarker, B. K., Khanam, S. J., Banke-Thomas, A., & Khan, M. N. (2024). Child undernutrition and its association with household environmental conditions in Bangladesh. *Public Health Nutrition*, *28*(1). https://doi.org/10.1017/S1368980024002325

Law of the Republic of Indonesia Number 16 of 2019 on the Amendment to Law, Pub. L. No. 16 (2019).

Le Gallo, J., López, F. A., & Chasco, C. (2020). Testing for spatial group-wise heteroskedasticity in spatial autocorrelation regression models: Lagrange multiplier scan tests. *The Annals of Regional Science*, *64*(2), 287–312. https://doi.org/10.1007/s00168-019-00919-w

Lessani, M. N., & Li, Z. (2024). SGWR: similarity and geographically weighted regression. *International Journal of Geographical Information Science*, *38*(7), 1232–1255. https://doi.org/10.1080/13658816.2024.2342319

Mardalena, S., Purhadi, P., Purnomo, J. D. T., & Prastyo, D. D. (2022). The Geographically Weighted Multivariate Poisson Inverse Gaussian Regression Model and Its Applications. *Applied Sciences (Switzerland)*, *12*(9). https://doi.org/10.3390/app12094199

Mathur, M. (2015). Spatial autocorrelation analysis in plant population: An overview. *Journal of Applied and Natural Science*, *7*(1), 501–513. https://doi.org/10.31018/jans.v7i1.639

Moraga, P. (2024). *Spatial Statistics for Data Science*. CRC Press. https://www.routledge.com/

Mulyani, A. T., Khairinisa, M. A., Khatib, A., & Chaerunisaa, A. Y. (2025). Understanding Stunting: Impact, Causes, and Strategy to Accelerate Stunting Reduction—A Narrative Review. *Nutrients*, *17*(9). https://doi.org/10.3390/nu17091493

Nariswari, R., Widhiyanthi, A. A., Arifin, S., & Yudistira, I. G. A. A. (2023, December). Zero inflated Poisson Regression: A solution of overdispersion in stunting data. In *AIP Conference Proceedings* (Vol. 2975, No. 1, p. 080007). AIP Publishing LLC. https://doi.org/10.1063/5.0181105

Partap, U., Young, E. H., Allotey, P., Sandhu, M. S., & Reidpath, D. D. (2019). Characterisation and correlates of stunting among Malaysian children and adolescents aged 6-19 years. *Global Health, Epidemiology and Genomics*, *4*. https://doi.org/10.1017/gheg.2019.1

Payne, E. H., Gebregziabher, M., Hardin, J. W., Ramakrishnan, V., & Egede, L. E. (2018). An empirical approach to determine a threshold for assessing overdispersion in Poisson and negative binomial models for count data. *Communications in Statistics: Simulation and Computation*, *47*(6), 1722–1738. https://doi.org/10.1080/03610918.2017.1323223

Pebesma, E., & Bivand, R. (2023). *Spatial Data Science: With Applications in R*. CRC Press. https://r-spatial.org/book/.

Pfeiffer, D. U., Robinson, T. P., Stevenson, M., Stevens, K. B., Rogers, D. J., & Clements, A. C. A. (2008). Spatial Analysis in Epidemiology. In *British Library Cataloguing*. Oxford University Press.

Phillips, S. J., & Mbizvo, M. T. (2016). Empowering adolescent girls in Sub-Saharan Africa to prevent unintended pregnancy and HIV: A critical research gap. *International Journal of Gynecology & Obstetrics*, *132*(1), 1–3. https://doi.org/10.1016/j.ijgo.2015.10.005

Putri Adila, W., Sri Yanti, R., & Sriyanti, R. (2023). The relationship of chronic energy deficiency (CED), exclusive breastfeeding, and economic with stunting in Nagari Aua Kuning West Pasaman. *Science Midwifery*, *10*(6), 4471-4480. www.midwifery.iocspublisher.orgJournalhomepage:www.midwifery.iocspublisher.org

Putri Ananda, E. Y., Annas, S., Ihsan, H., & Aswi, A. (2024). Negative Binomial Regression Analysis of Factors Influencing Stunting Cases in Central Lombok Regency. *Inferensi*, *7*(3), 167. https://doi.org/10.12962/j27213862.v7i3.21436

Rah, J. H., Cronin, A. A., Badgaiyan, B., Aguayo, V. M., Coates, S., & Ahmed, S. (2015). Household sanitation and personal hygiene practices are associated with child stunting in rural India: a cross-sectional analysis of surveys. *BMJ open*, *5*(2), e005180. https://doi.org/10.1136/bmjopen-2014-005180

Regulation of the Minister of Health of the Republic of Indonesia Number 2 of 2020 on Child Anthropometry Standards, Pub. L. No. 2 (2020).

Rohmaniah, E., Rahmawati, I., & A'la, M. Z. (2025). The Impact of Early Marriage: a Literature Review. *Jurnal Ners*, *9*(2), 2176–2183. https://doi.org/10.31004/jn.v9i2.42841

Sajalia, H., Dewi, Y. L. R., & Murti, B. (2018). Life Course Epidemiology on the Determinants of Stunting in Children Under Five in East Lombok, West Nusa Tenggara. *Journal of Maternal and Child Health*, *03*(04), 242–251. https://doi.org/10.26911/thejmch.2018.03.04.01

Salmerón, R., García, C. B., & García, J. (2018). Variance Inflation Factor and Condition Number in multiple linear regression. *Journal of Statistical Computation and Simulation*, *88*(12), 2365–2384. https://doi.org/10.1080/00949655.2018.1463376

Salmerón, R., García, C., & García, J. (2020). Overcoming the inconsistences of the variance inflation factor: A redefined VIF and a test to detect statistical troubling multicollinearity. *arXiv preprint arXiv:2005.02245*. http://arxiv.org/abs/2005.02245

Sandøy, I. F., Mudenda, M., Zulu, J., Munsaka, E., Blystad, A., Makasa, M. C., Mæstad, O., Tungodden, B., Jacobs, C., Kampata, L., Fylkesnes, K., Svanemyr, J., Moland, K. M., Banda, R., & Musonda, P. (2016). Effectiveness of a girls' empowerment programme on early childbearing, marriage and school dropout among adolescent girls in rural Zambia: study protocol for a cluster randomized trial. *Trials*, *17*(1), 588. https://doi.org/10.1186/s13063-016-1682-9

Santoso, P., & Pujianto, T. (2024). The Analysis of Factors that Influence Stunting. *Jurnal Ners Dan Kebidanan (Journal of Ners and Midwifery)*, *11*(2), 200–208. https://doi.org/10.26699/jnk.v11i2.ART.p200-208

Seifu, B. L., Tesema, G. A., Fentie, B. M., Yehuala, T. Z., Moloro, A. H., & Mare, K. U. (2024). Geographical variation in hotspots of stunting among under-five children in Ethiopia: A geographically weighted regression and multilevel robust Poisson regression analysis. *PLoS ONE*, *19*(5 May). https://doi.org/10.1371/journal.pone.0303071

Silumbwe, A., Nkole, T., Munakampe, M. N., Cordero, J. P., Milford, C., Zulu, J. M., & Steyn, P. S. (2020). Facilitating community participation in family planning and contraceptive services provision and uptake: community and health provider perspectives. *Reproductive Health*, *17*(1), 119. https://doi.org/10.1186/s12978-020-00968-x

Soliman, A., De Sanctis, V., Alaaraj, N., Ahmed, S., Alyafei, F., Hamed, N., & Soliman, N. (2021). Early and long-term consequences of nutritional stunting: From childhood to adulthood. *Acta Biomedica*, *92*(1). https://doi.org/10.23750/abm.v92i1.11346

Subramanee, S. D., Agho, K., Lakshmi, J., Huda, Md. N., Joshi, R., & Akombi-Inyang, B. (2022). Child Marriage in South Asia: A Systematic Review. *International Journal of Environmental Research and Public Health*, *19*(22), 15138. https://doi.org/10.3390/ijerph192215138

Taslim, N. A., Farradisya, S., Gunawan, W. Ben, Alfatihah, A., Barus, R. I. B., Ratri, L. K., Arnamalia, A., Barazani, H., Samtiya, M., Mayulu, N., Kim, B., Hardinsyah, H., Surya, E., & Nurkolis, F. (2023). The interlink between chrono-nutrition and stunting: current insights and future perspectives. *Frontiers in Nutrition*, *10*. https://doi.org/10.3389/fnut.2023.1303969

Tiara, Y., Aidi, M. N., Erfiani, E., & Rachmawati, R. (2023). Overdispersion Handling In Poisson Regression Model By Applying Negative Binomial Regression. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, *17*(1), 0417–0426. https://doi.org/10.30598/barekengvol17iss1pp0417-0426

TN2PK. (2017). *100 Kabupaten/Kota Prioritas Untuk Intervensi Anak Kerdil (Stunting)*.

Vaivada, T., Akseer, N., Akseer, S., Somaskandan, A., Stefopulos, M., & Bhutta, Z. A. (2020). Stunting in childhood: an overview of global burden, trends, determinants, and drivers of decline. *The American Journal of Clinical Nutrition*, *112*, 777S-791S. https://doi.org/https://doi.org/10.1093/ajcn/nqaa159

Yoosefi Lebni, J., Solhi, M., Ebadi Fard Azar, F., Khalajabadi Farahani, F., & Irandoost, S. F. (2023). Exploring the Consequences of Early Marriage: A Conventional Content Analysis. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, *60*. https://doi.org/10.1177/00469580231159963

Yuliastanti, T., Ambarwati, W. N., Sulastri, S., & Rahmawati, A. (2023). History of Chronic Energy Deficiency (CED) of Pregnant Women and Stunting in Toddlers. *International Journal of Nursing Information*, *2*(2), 7–12. https://doi.org/10.58418/ijni.v2i2.45

Zheng, B., Lin, X., Yin, D., & Qi, X. (2023). Does Tobler's first law of geography apply to internet attention? A case study of the Asian elephant northern migration event. *PLoS ONE*, *18*(3), 1–17. https://doi.org/10.1371/journal.pone.0282474