

Factor Analysis on Stunting in East Java 2021 Using The Principal Component Analysis (PCA) Method

Sella Nasrania¹, Muhammad Zulfadhli², Dzakki Damar Rabbani³, Abghaza Bayu Kusuma Wardana Putra⁴

^{1,3,4}Departement of Business Statistics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

²Departement of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

Article Info

Article history:

Received June 3, 2026

Revised August 4, 2026

Accepted September 25, 2026

Keywords:

Factor Analysis

Dependent

Data Adequacy

Multivariate Normal

Principal Component Analysis

Stunting

ABSTRACT

Stunting is a condition of failure to grow in children due to chronic malnutrition that has an impact on physical and cognitive health. To understand the factors that contribute to the occurrence of stunting, indicators are needed that cover various aspects such as health, access to sanitation, education, and socioeconomic status. This study aims to analyze the factors that influence the incidence of stunting in East Java in 2021 using the Principal Component Analysis (PCA). The PCA method is applied to reduce data consisting of many interrelated variables into smaller main components while retaining the main information from the original data. This study uses 9 variables with a total of 38 data that affect the stunting in East Java in 2021. Secondary data was obtained from the Central Statistics Agency of East Java Province. After analysis, it was found that the main variables that affected stunting can be explained through two main factors that account for 69.488% of the total variance. The first factor includes public welfare and health, which includes sanitation, immunization, health index, education index, and economic variables. The second factor is basic health and access to sanitation, which includes breastfeeding and access to decent drinking water. These results are expected to be the basis for policymakers in determining more effective interventions to reduce stunting rates in East Java.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Muhammad Zulfadhli

Departement of Statistics

Institut Teknologi Sepuluh Nopember

Surabaya 60111, Indonesia

Email: muhammadzulfadhli23@gmail.com

1. INTRODUCTION

Stunting is a condition that reflects chronic nutritional problems that affect children's growth and development and is one of the important indicators in assessing the quality of public health, especially children in Indonesia [1]. Stunting is defined as a condition when a child's height is shorter than the standard height of a child of his age, caused by a lack of nutritional intake in the long term. This problem is closely related to various factors, such as nutritional status, access to sanitation, immunization, and household socio-economic conditions [2]. Identification of factors that contribute to stunting is important to do this, so that governments and stakeholders can understand the most influential factors and design effective interventions to reduce the number of stunting in vulnerable areas, including East Java [3].

Identify the influencing factors *stunting*. It is important to know the role of each factor in the mitigation efforts *stunting*. Therefore, statistical methods such as *Principal Component Analysis* (PCA) can help reduce many related variables and identify the dominant factors affecting the level of *stunting* in East Java in 2021 [4]. Based on this, this study will conduct a factor analysis of the *stunting* in East Java to

simplify the double-variable data and identify the key variables that have the most impact on the incidence *stunting* [5]. *Principal Component Analysis* (PCA) is a multivariate analysis technique that aims to reduce several correlated variables into a new, smaller set of variables, called main components, that are independent of each other and can represent the information contained in the origin variable [6]. By converting the original variables into independent components, PCA helps reduce data complexity, minimize redundancy, and facilitate data interpretation.

The data in this study comes from secondary data obtained from the Central Statistics Agency (BPS) of East Java Province for 2021 [7]. The data collected includes several indicators related to household conditions and children's health, such as the percentage of child weight, access to proper sanitation, complete immunization, access to proper drinking water, health and education index, and the percentage of poor people and women who marry early [8]. This data is then analyzed using *Principal Component Analysis* (PCA) to assess key factors related to *stunting* in East Java. By using PCA, this study is expected to provide a deeper understanding of the aspects that affect the incidence *stunting*, so that it can be a solid basis for policymakers in setting development priorities and health interventions to reduce the number of *stunting* in East Java.

2. METHOD

2.1. Data Source

The data used in this study is secondary data obtained through *Copyright* © The Central Statistics Agency (BPS) of East Java Province related to the factors that affect *stunting* in East Java in 2021 with the number of data for each variable as many as 38 data.

2.2. Research Variables

The research variables used in this study are shown in Table 1.

Table 1. Research Variables

Variable	Remarks	Data Scale	Units
X1	Percentage of child weight	Interval	Presses
X2	Percentage of households that have access to proper sanitation	Interval	Presses
X3	Percentage of complete immunization	Interval	Presses
X4	The percentage of children aged 0-12 months has been breastfed	Interval	Presses
X5	Health index	Interval	
X6	Percentage of Households That Have Access to Decent Drinking Water Sources	Interval	Presses
X7	Education index	Interval	
X8	Percentage of poor population	Interval	Presses
X9	Percentage of the female population of early marriage	Interval	Presses

2.3. Analysis Step

The analysis steps used in this study are as follows.

1. Collecting data on factors affecting *stunting* in East Java in 2021 through *the bps.go.id website*.
2. Describe the characteristics of the data on factors affecting *stunting* in East Java in 2021.
3. Conducting an examination and testing of the assumptions of factor analysis consisting of four tests and examinations as follows
 - a. Multivariate normal distribution examination using a correlation test between the anobis mahalanobis distance and the quartile value
 - b. Dependency correlation test using *the barlett test*
 - c. Test data sufficiency using SME inspection
 - d. Correlation check between variables using *anti image correlation*
4. Conducting a factor analysis using the analysis of the main components of the data on factors affecting *stunting* in East Java in 2021.
5. Interpreting results
6. Draw conclusions and suggestions.

2.4. Descriptive Statistics

Descriptive analysis is a method of data analysis that aims to assess the generalization of research results based on one specific sample [9]. This analysis is done through descriptive hypothesis testing, which helps identify the basic characteristics of the data without making inferences that go beyond the observed sample. Descriptive statistics is a branch of statistics that focuses on collecting, presenting, and managing data to describe the basic characteristics of a data set [10]. This technique includes measures of data concentration *mean*, median, and mode, the size of the spread (such as *Range*, variance, and standard deviation), as well as the use of tables, graphs, and charts [11].

2.4.1 Average

How to calculate average, namely by adding up all data frequencies and dividing by the amount of data. *Red* This is very close to quantitative data, usually the value of this quantitative data with the symbol that is $x_1, x_2 \dots$ to x_n (if the data contains n values). For the symbol n , it is usually used to express a size in a sample, namely the amount of data, which will be studied, which is then used as a sample. N is also used to express the size of the population, which is the number of members as a whole in the data. Average or *red* is the ratio of the total value of observations to the number of observations [12].

The single data is as follows.

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

Description:

x_i = The value of the variable to $-i$

i = 1,2,3,... n

n = The amount of data of the value x

The group data is as follows [7].

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

Description:

x_i = Midpoint of the i -value

f_i = Frequency of class i

n = The amount of data of the value x

2.4.2 Median

Median is one of the measures of data traceability, i.e. if a group of data is sorted from smallest to largest or largest to smallest, the value of the observation is right in the middle if the amount of data is odd, or the average of the two observations in the middle if the number of observations is even [13].

Formula *median* For a single data with n odd numbers is as follows.

$$Me = \frac{1}{2} x_{n+1} \quad (3)$$

Formula *median* For a single data with n even numbers is as follows.

$$Me = \frac{1}{2} x_{\frac{n}{2}} + x_{\frac{n}{2}} \quad (4)$$

Description:

Me = *Median*

x = Data value

n = Amount of data

Formula *median* for the group data is as follows.

$$Me = Tb \left[\frac{\frac{n}{2} - F}{f} \right] p \quad (5)$$

Description:

Tb = Bottom edge of the class *median*

F = Cumulative frequency before class *median*

f = Class frequency *median*

n = Lots of data

p = Class length

2.4.3 Variance

Variance is the sum of the squares of all the deviations of individual values from the mean of the group. *Variance* can describe how a quantitative data is scattered [14]. *Variance* (S_2) a single data is expressed as follows.

$$S_2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \quad (6)$$

The formula has an explanation, namely i is the first observation and x_1 is the first observation, then x_2 is the second observation and so on. $\bar{x}$ is the average and n is the number of observations. As for the number of observations. $\bar{x}$ *variance* (S_2) a group data is stated as follows [15].

$$S_2 = \frac{\sum_{i=1}^n f_i (x_i - \bar{x})^2}{\sum_{i=1}^n f_i} \quad (7)$$

Description:

f_i = Frequency of class i
 x_i = Middle grade of the i class
 i = 1, 2, 3, ..., n
 $\bar{x}$ = Average
 n = Amount of data

2.4.4 Standard Deviation

Standard deviation is the sum of the variation or spread in a range of values. *Standard deviation* low indicates that the value tends to be close to average. From this data, *Standard Deviation* indicates that the value of the data is getting farther or farther away from the average [16]. The formula used to calculate the value *Standard Deviation* shown as follows.

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (8)$$

Description:

s = *Standard Deviation*
 $\bar{x}$ = average
 x_i = value to- x_i
 i = 1, 2, 3, ..., n
 n = a lot of data

2.5 Q-Q Plot

Q-Q Plot (*Quantile-Quantile Plot*) is a graphical method used in statistics to compare the distribution of empirical data with a theoretical distribution or to compare two sets of data distributions [17]. In a Q-Q Plot, the quantiles of the data distribution are compared with the quantiles of the reference distribution, e.g. a normal distribution, to see if the data follows the expected distribution pattern. If the data is normally distributed, the dots on the Q-Q Plot will form a straight line. This plot is often used in statistical analysis to check the assumption of normality before performing further statistical tests

2.6 Scree Plot

Scree Plot is a graph used in factor analysis or analysis of the main components of the PCA (*Principal Component Analysis*) to specify the exact number of factors or components to extract [18]. This graph displays the value of own (*eigenvalues*) on the y-axis and the number of components or factors on the x-axis. On *Scree Plot*, the dots show *Self-esteem* of each component, and this graph forms a descending line. Components that are before the point "Elbow" or "square point" (the place where the graph starts to flat) is usually chosen as the main factor, as they are considered to contribute the greatest to the variability of the data. Examples *Scree Plot* as follows. (Source : <https://support.minitab.com/>)

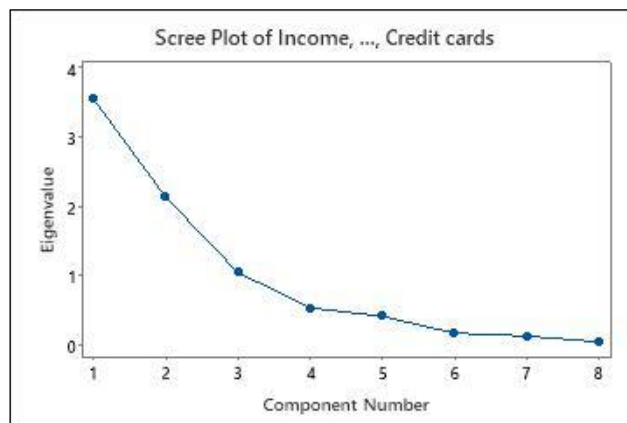


Figure 1. Scree Plot

2.7 Assumption of Factor Analysis

Assumptions in factor analysis include several requirements that need to be met to ensure that the results of the analysis can be interpreted validly and reliably [19]. These assumptions include normality tests, which ensure that data follow normal distributions; independence tests or tests *Bartlett*, which tests for a significant correlation between variables; the adequacy of the data with *Kaiser-Meyer-Olkin (SME) Test* to assess whether the data is adequate for factor analysis; and correlation examination between variables through *Anti Image Correlation*, which determines whether each variable has an adequate correlation to be included in a given factor [20].

2.7.1 Assumption of Multivariate Normal Distribution

The assumption of a multivariate normal distribution is one of the important requirements in multivariate statistical analysis which states that the variables in a data follow a normal distribution simultaneously or collectively [21]. This assumption is important in various multivariate analysis techniques such as factor analysis, multiple regression, and key component analysis (PCA) because it affects the validity of statistical test results. Multivariate normal distribution assumes that any linear combination of those variables also follows a normal distribution. In research, the multivariate normal distribution is often tested through normality tests on each variable and on combined data to ensure that sampling errors and parameter estimation results are minimal [22]. Testing of the assumptions of multivariate normal distributions in this study will be done through a t-proportion test, where the data is considered to be multivariate normally distributed if it has a proportion value between 45% to 55%. The hypothesis for this test is formulated as follows.

Hipotesis:

H0: Multivariate normal distributed data

H1: Multivariate normal undistributed data

Critical Area: Minus H0 if $45\% \leq T$ the proportion $\leq 55\%$ and the P-value $< \alpha$

Test Statistics:

$$d_j^2 = (x_j - \bar{x})' S^{-1} (x_j - \bar{x}) \quad (9)$$

Description:

j = Number of observations, 1, 2, ..., n

x_j = Object of observation to $-j$

$\bar{x}$ = Average observation to $-j$

d_j^2 = The squared value of the j

S^{-1} = Invers matriks varians kovarian

2.7.2 Assumption of Independence (Bartlett Test)

The assumption of independence in factor analysis refers to the need that the variables used in the analysis must be significantly correlated in order for factor analysis to be applied effectively. To test this assumption, Test *Bartlett's Test of Sphericity* is commonly used. The Bartlett test the null hypothesis that the correlation matrix is an identity matrix, which means that there is no significant correlation between the variables [23]. If the test results (Suryani, 2017) *Bartlett* shows a low significance value ($p < 0.05$), then the independence assumption is met, which indicates that there is sufficient correlation to proceed with the factor analysis.

Hipotesis:

H0 : $\rho = I$ (Correlation matrix is identical to identity matrix or between independent variables)
H1 : $\rho \neq I$ (Correlation matrix is not identical to identity matrix or between dependent variables)

Slowest Meaning: α

Rejection Area: Subtract H0 if χ^2 counts $> \chi^2_{\alpha, df}$ or p-value $< \alpha$

Test Statistics:

$$X^2 = -\left\{n - 1 \frac{2p+5}{6}\right\} \ln|R| \quad (10)$$

Description:

n = Amount of data

$|R|$ = Determinant value

p = Number of variables used

2.7.3 Assumption of the Adequacy of the Kaiser Mayer Olkin Test (KMO)

Test *Kaiser-Meyer-Olkin* (KMO) is a test used to assess the adequacy of data in factor analysis, i.e. to determine whether the data has sufficient correlation between variables to be suitable for use in factor analysis [24]. The KMO test produces values that range from 0 to 1; the higher the SME value, the more adequate the data is for factor analysis. If the SME value is low, it indicates that the data does not have adequate correlation between variables, so factor analysis may not be appropriate for use in the data. There are six criteria used to assess suitability in the use of factor analysis, which are as follows.

Table 2. SME Assessment Criteria

SME Value	Recommendations
$0.9 \leq \text{SMEs} \leq 1$	Excellent
$0.8 \leq \text{SMEs} \leq 0.89$	Useful
$0.7 \leq \text{SMEs} \leq 0.79$	Ordinary
$0.6 \leq \text{SME} \leq 0.69$	Enough
$0.5 \leq \text{SME} \leq 0.59$	Bad
$\text{SME} < 0.5$	Not Accepted

2.7.4 Correlation Examination Between Variables (*Korelasi Anti Image Correlation*)

Checking the correlation between variables through the Anti Image Correlation method is one step in factor analysis to ensure that the variables being analyzed have an adequate correlation and are suitable for inclusion in the factor model [25]. Anti Image Correlation measures the degree of correlation between variables outside of the main component, helping to identify whether a variable has an adequate correlation with other variables. If the Anti Image correlation value of a variable is low (e.g., less than 0.5), it may not be suitable for inclusion in factor analysis due to the lack of strong relationships with other variables

2.8 Analysis of the Main Components

Major Component Analysis is a statistical technique used to reduce the dimensions of data consisting of many correlated variables, without losing important information [26]. This analysis aims to identify a number of key components, i.e. linear combinations of origin variables, that can explain the greatest variation in the data. The first component accounts for the largest variance, followed by the second component that is orthogonal to the first component, and so on. Each component has an eigenvalue that indicates how much variation the component describes. (Pratama, 2020)

$$Z_1 = e'_1 X = e_{i1} X_1 + e_{i2} X_2 + \dots + e_{ip} X_p \quad (11)$$

Description:

i = 1, 2, ..., p

Z_1 = first PC with the first largest variance

Z_2 = The second PC with the second largest variance

Z_p = PC to- p with the largest variance top

X_1 = First variable

X_2 = Second variable

X_p = Variable ke- p

e_p = Custom vector data to- p

2.8.1 Own Value

Own value (eigenvalue) is a scalar value in linear algebra and matrix analysis that shows the magnitude of the linear transformation that occurs in a vector, also known as an eigenvector. In the context of

factor analysis or *Principal Component Analysis* (PCA), *Own value* indicates the contribution or variation described by each factor or component in the data. Large indicates that the factor or component has a significant role in describing the variability of the data, while the smaller eigenvalue is usually ignored due to its minimal contribution. *Own value* calculated from a covariance or correlation matrix and is often displayed in *Scree Plot* to determine the optimal number of factors.

2.8.2 Commonality Test

The commonality test in factor analysis is a method for determining how much variance of each variable can be explained by the extracted factor. The commonality value indicates the proportion of variance of the variables that can be explained by the overall factor in the model. This value ranges between 0 and 1, where a value close to 1 indicates that the variable is strongly represented by the extracted factor. In contrast, a low value indicates that the variable is not sufficiently well represented by the factor and may be less relevant in the model.

2.8.3 Total Variance Test

Total Variance Explained or total variance test is a method in factor analysis or major component analysis (PCA) used to assess how large *variance* of the original data that can be explained by the factors or components extracted in the model. This test is important to determine the percentage (Pratama, 2019) *variance* which is explained by several factors or components, which allows us to assess the feasibility of the built model. Typically, the total *variance* described by the selected factors must be large enough (ideally more than 50%) for the model to be considered representative and accurate in describing the original data. If the total percentage *variance* described low, the model may be less able to describe the basic structure of the data

2.8.4 Component Matrix

Component Matrix is a table in the main component analysis (PCA) or factor analysis that displays the weights or *Loading* of each variable on each component or factor extracted. This matrix shows the contribution of each variable to a particular factor or component, with a value *Loading*. The higher indicates that the variable has a stronger influence on the component. *Component Matrix* It is used to identify the relationship between the variable and the extracted component, thus aiding in the interpretation of the main factor or component in the analysis. In general, in factor analysis, a variable with a load above 0.5 is considered significant and relevant in describing a particular factor or component.

2.9 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a statistical technique used to reduce the dimensions of data that are made up of many variables that are correlated with each other, without losing the main information [27]. PCA aims to identify the main component, that is, the linear combination of the origin variable, which can explain the greatest variation in the data. Each component generated has *self-esteem* which shows how much variance of data can be explained by the component. By maintaining the components with *Self-esteem* largest, PCA allows for simplification of data for further analysis while still retaining most of the information. In general, the i-I PC model can be stated as follows.

$$Z_i = e'_i X \quad (12)$$

Thus, the main linear combinations can be formulated as follows.

$$\begin{aligned} Z_1 &= e'_1 X = e_{11}X_1 + e_{12}X_2 + \dots + e_{1p}X_p \\ Z_2 &= e'_2 X = e_{21}X_1 + e_{22}X_2 + \dots + e_{2p}X_p \\ &\vdots \\ Z_p &= e'_p X = e_{p1}X_1 + e_{p2}X_2 + \dots + e_{pp}X_p \end{aligned} \quad (13)$$

Description:

$i = 1, 2, \dots, p$

Z_1 = first PC with the first largest variance

Z_2 = The second PC with the second largest variance

Z_p = PC to- p with the largest variance top

X_1 = First variable

X_2 = Second variable

X_p = Variable ke- p

e_p = Custom vector data to- p

2.10 Factor Analysis

Factor analysis is a multivariate statistical technique used to reduce data by grouping highly correlated variables into one or more factors [28]. Each factor results reflect a set of interrelated variables and can be interpreted as a simpler concept or construct. The main goal of factor analysis is to simplify complex data, help identify the underlying structures in the data, and allow for a more meaningful interpretation of the relationships between variables. The model of factor analysis is shown as follows. (Pratama, 2018)

$$X_p - \mu_p = \ell_{p1}F_1 + \ell_{p2}F_2 + \dots + \ell_{pm}F_m + \varepsilon_p \quad (14)$$

Description:

μ_1 = Average of thei

ε_1 = Specific factors to-*i*

F_j = Common factor ke-*j*

ℓ_{ij} = Loading of variables Towards the Factor of Being *ij*

The variance in the fourth variable is explained by *im* A common factor is referred to as the communality of -, which is calculated as the square of *i* Loading variable to- *jm* general factors, with the formula below.

$$h_j^2 = \ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2 \quad (15)$$

2.11 Stunting

Stunting is a condition of failure to grow in children that occurs due to chronic malnutrition, especially in the first thousand days of life, from fetuses to two-year-olds. Children who experience stunting have a shorter height than the high standards of children of their age, which reflects the presence of obstacles in their physical growth and brain development. This condition not only affects physical aspects but also has an impact on cognitive development, so children who experience stunting tend to have lower learning abilities compared to children who grow normally.

Stunting is often caused by long-term inadequate nutritional intake, recurrent infections, poor sanitation, and inadequate access to health services. In Indonesia, stunting is a significant public health problem because it can affect the quality of human resources in the future. Efforts to reduce stunting rates include nutrition interventions, improved maternal and child health services, and improved access to sanitation and clean water.

2.12 Percentage of Child's Weight

A child's body weight percentage is a measure that shows the proportion of a child's weight compared to the child's ideal weight standard based on age and sex. This indicator is important for determining a child's nutritional status, especially in detecting malnutrition, including malnutrition or overweight. A child's body weight percentage is usually measured based on WHO growth standards, where a child's weight is assessed against the percentile or z-score that represents the normal distribution of weight in the population

2.13 The percentage of households that have access to proper sanitation

The percentage of households that have access to adequate sanitation is an important indicator that measures the proportion of households in a population that have access to adequate sanitation facilities, such as safe toilets and are connected to appropriate sewage systems. This indicator is used to assess the quality of the health environment and the standard of living of the community, as proper sanitation contributes to preventing infectious diseases and improving the quality of health, especially in children who are susceptible to diseases due to poor sanitation

2.14 Percentage of Complete Immunization

The percentage of complete immunization is a health indicator that shows the proportion of children who have received all doses of basic vaccination on schedule, as recommended by the Ministry of Health and WHO. Complete immunization includes essential vaccines, such as BCG, polio, hepatitis B, DPT, and measles, which must be administered within a certain period of time to ensure effective protection against various infectious diseases. This indicator is used to monitor the success of immunization programs and ensure that the majority of children in the population have protection against preventable diseases.

2.15 Percentage of children aged 0-12 months have been breastfed

The percentage of children aged 0-12 months who have been breastfed is a health indicator that measures the proportion of babies in this age group who receive breast milk (breast milk) at least once from birth. This indicator is important in the assessment of child nutrition and health, as breast milk is the ideal

primary source of nutrition for babies, especially in the first six months of life. Breast milk not only provides the necessary nutrients but also provides important immunological protection, so it is highly recommended by the WHO and the Ministry of Health as an exclusive food for babies up to six months of age.

2.16 Health Index

The Health Index is a composite indicator that measures various aspects of a population's health to provide an overview of the health status of a community. This index typically includes several health indicators such as life expectancy, infant mortality rates, morbidity rates, access to health facilities, and the prevalence of certain diseases. By combining various components, the Health Index allows for a more thorough assessment of the quality of health in a region or country.

2.17 The percentage of households that have access to a decent source of drinking water

The percentage of households that have access to a viable source of drinking water is an important indicator that describes the proportion of households that obtain drinking water from sources that meet health standards, such as water that is protected from contamination and suitable for consumption without health risks. Sources of decent drinking water typically include water from piping systems, protected wells, and protected springs, which are assessed accordingly by WHO and local government standards

2.18 Education Index

The Education Index is a composite measure that describes the level of education of a society based on several key indicators, such as literacy rates, average length of schooling, and school participation rates. This index is used to assess educational attainment in a region, and is often part of the calculation of the Human Development Index (HDI) published by the Central Statistics Agency (BPS) and international organizations

2.19 Percentage of Poor Population

The percentage of the poor population is an indicator that describes the proportion of individuals in a population who live below the poverty line. This indicator refers to the minimum economic standards required to meet basic needs, such as food, clothing, shelter, education, and health. In Indonesia, the Central Statistics Agency (BPS) defines the poor based on the poverty line calculated from various components of expenditure, including basic and non-basic needs.

2.20 Percentage of Female Population Married Early

The percentage of the female population of early marriage is an indicator that shows the proportion of women in each age group under 18 years of age, who have married. This indicator is used to assess the prevalence of early marriage in a region and reflects the various social, economic, and cultural aspects that influence marriage decisions at a young age. Early marriage can have a significant impact on women's lives, including limited access to education, health risks in early pregnancy, and increased intergenerational poverty.

2.21 Previous Research

Research related to modeling socio-economic determinants, health status, and regional development in Indonesia has demonstrated the importance of multivariable and nonparametric statistical approaches in capturing the complexity of regional data, such as unmet need status [29],[30], nonparametric binary logistic regression inference frameworks for the status of disadvantaged areas [31], and some studies using classification models such as [32] as well as modeling of open unemployment status and public health [33],[34],[35].

3. RESULTS AND DISCUSSION

3.1. Data Characteristics of Factor Stunting Based on City/Regency in East Java Province in 2021

A In this sub-chapter, the characteristics of data related to factors will be described *stunting* in cities/districts of East Java Province in 2021 which include the percentage of child weight, the percentage of households that have access to proper sanitation, the percentage of complete immunization, the percentage of children aged 0-12 months who have been breastfed, the health index, the percentage of households that have access to proper drinking water sources, the education index, the percentage of the poor population and the percentage of the female population of early marriage. The characteristics of the data are explained in the following table.

Table 3. Characteristics of S Factor *stunting* Cities/Regencies in East Java in 2021

Variable	Mean	Varians	Min	City/Regency	Max	City/Regency
X1	12.225	18.230	3.490	Mojokerto Regency	20.370	Gresik
X2	81.971	176.585	39.440	Bangkalan	97.310	Madiun
X3	63.178	281.206	17.180	São Paulo	83	Trenggalek
X4	34.297	45.077	22.770	Mojokerto Regency	50.400	Mojokerto City
X5	0.832	0.000	0.810	Bondowoso, Gresik and Bangkalan	0.850	Kediri, Mojokerto and Surabaya Cities
X6	95.183	35.180	75.780	Trenggalek	100	Stone
X7	0.641	0.005	0.510	Sampang	0.780	Malang and Madiun
X8	11.322	22.280	4.090	Stone	23.760	Sampang
X9	18.954	95.739	5.590	Madiun	44.520	Bondowoso

Table 3 shows that the variables of the percentage of child weight, the percentage of households that have access to proper sanitation, the percentage of complete immunization, the percentage of children aged 0-12 months who have been breastfed, the percentage of the poor population and the percentage of the female population who marry early have a percentage variation that is greater than the average value, which indicates that the distribution of these factors is uneven in East Java cities/districts in 2021. Gresik Regency is the district with the highest percentage of child weight, while Mojokerto Regency is the lowest. The highest percentage of households with access to proper sanitation is found in Madiun Regency and Bangkalan Regency with the lowest. Trenggalek Regency is the district with the highest percentage of complete immunization, while Sumenep Regency is the district with the lowest complete immunization. The percentage of children aged 0-12 months has been breastfed, Mojokerto City has the highest percentage, while Mojokerto Regency has the lowest percentage. Sampang Regency is the district with the highest percentage of poor people, while Batu City is the city with the lowest percentage of poor people. In addition, the highest percentage of early marriage women is in Bondowoso Regency, while the lowest is in Madiun City.

Health index factors, the percentage of households that have access to decent drinking water sources and the education index have an even distribution in each city/district of East Java Province in 2021 because the value variation is smaller than the average value. The cities of Kediri, Mojokerto and Surabaya have the highest health index, while Bondowoso, Gresik and Bangkalan Regencies have the lowest health index. Batu City is the city with the highest percentage of households that have access to the highest source of decent drinking water and the lowest Trenggalek Regency. In addition, Malang and Madiun cities are the cities with the highest education index in East Java, while Sampang Regency is the district with the lowest education index in East Java in 2021.

3.2 Examination of Factor Analysis Assumptions on Influencing Factors Data *Stunting* in East Java in 2021

The examination of the assumption of factor analysis was carried out including testing the assumption of normal distribution, dependency test, data adequacy test and correlation examination between variables which are described as follows.

3.2.1 Assumption of Multivariate Normal Distribution Assumption of Factors Influence *Stunting* in East Java in 2021

Multivariate normal distribution testing is performed to find out if the data Factors that seduce *stunting* in East Java in 2021 has normal distribution or not. The multivariate normal distribution test is performed through visual examination and statistical testing. The results of the multivariate normal distribution test are visually presented in Figure 2.

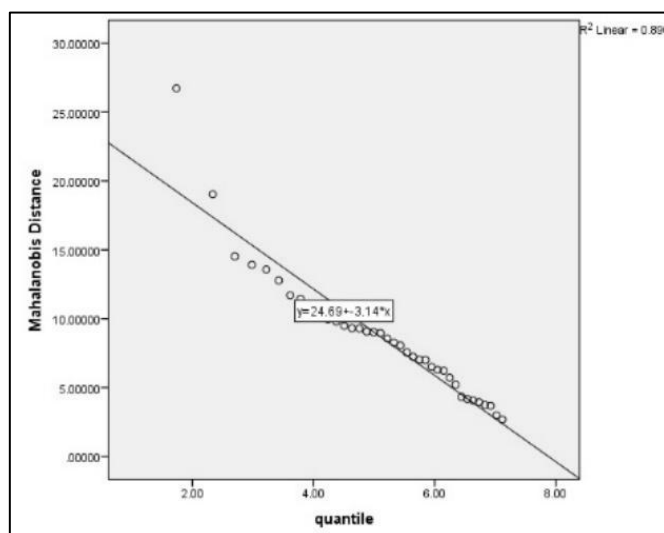


Figure 2. Q-Q Multivariate Normal Distribution Plot Factors *Stunting* in East Java in 2021

Figure 2 shows that the plots are generally scattered close to the normal line, so visually it can be said that the data of the influencing factors *stunting* in East Java in 2021 with a multivariate normal distribution. Then to ensure that the data is correctly distributed normally multivariate, pmultivariate normal distribution test using values on the data of the factors that $T_{proporsi}$ seduce *stunting* in East Java in 2021, which is explained as follows.

Hipotesis:

H_0 : Data factors that seduce *stunting* in East Java in 2021 with a multivariate normal distribution

H_1 : Data factors that seduce *stunting* in East Java in 2021 there is no normal distribution of multivariate

The multivariate normal distribution test uses a test $T_{proporsi}$ with a significant level of α of 0.05, the rejection area is obtained, namely H_0 is rejected if $T_{proporsi}$ it is outside the boundary. Based on statistical calculations, the results of the multivariate normal distribution test results in a value of 50%, which is within the interval. So that a failed rejection decision was obtained, which means that the data on factors affecting $45\% \leq T_{proporsi} \leq 55\%$ H_0 *stunting* in East Java in 2021 are normally multivariate distribution.

3.2.2 Test Barlett Influencing Factors *Stunting* in East Java in 2021

The *Barlett* test was carried out to find out whether the data on factors affecting *stunting* in East Java in 2021 were independent or not. The results of the *Barlett* test on the data on factors affecting *stunting* in East Java in 2021 are as follows.

Hipotesis :

H_0 : $\rho = I$ (correlation matrix between *stunting* effect data is independent)

H_1 : $\rho \neq I$ (correlation matrix between data on the effect of *stunting* is dependent)

By using a significant level of α of 0.05, the rejection criterion is obtained, namely H_0 is rejected if the $>$ or P-value $\chi^2_{hitung} \chi^2_{(0,05;df)} < 0.05$. The results of the statistical calculation are presented in table 3.

Table 3 Test Statistics <i>Barlett</i>		
χ^2_{hitung}	$\chi^2_{(0,05;28)}$	P-value
177.583	41.337	0.001

Based on Table 3 it is known that the *barlett* resulting in a value of 177,583 greater than $\chi^2_{hitung} \chi^2_{(0,05;28)}$ of 41,337 and strengthened by 0.001 which is smaller than $P_{value} \alpha$ is 0.05. So it was decided to reject H_0 which means the correlation between the data of the factors that affect *stunting* in East Java in 2021 is dependent.

3.2.3 Kaiser Meyer Olkin (KMO) Criteria Examination Influencing Factors *Stunting* in East Java in 2021

The KMO examination is used to find out whether a certain data sample used to measure the adequacy of the data is able to meet the requirements or not. A data is said to be sufficient if the KMO value is above 0.5. In this study, the criteria examination resulted in a KMO value of 0.785, which means that the data on factors affecting *stunting* in East Java in 2021 is sufficient to be factored in.

3.2.4 Correlation *Anti-Image* Influencing Factors *Stunting* in East Java in 2021

Anti-Image *correlation* is carried out using MSA values to explain the relationship to each variable. The condition for accepting Anti-Image correlation is that if the MSA value is more than or equal to 0.5, then the variable will be accepted and analyzed further. Meanwhile, if the MSA value is less than 0.5, then the variable cannot be analyzed so it needs to be eliminated. The *Anti-Image correlation test* on the data on the influence of factors affecting *stunting* in East Java in 2021 is shown in Table 4.

Table 4 Correlation *Anti-Image*

Variable	MSA Value
X1	0.491
X2	0.913
X3	0.690
X4	0.611
X5	0.813
X6	0.643
X7	0.806
X8	0.757
X9	0.787

Based on Table 4, it is known that X1 shows the result of the *Anti-Image correlation test* of 0.491 which is smaller than 0.5, so in this study the variable X₁ cannot be analyzed and needs to be eliminated. Meanwhile, the variables X₂ to X₉ produce an *Anti-Image correlation test value* for each variable that is larger than 0.5. So that by eliminating the X₁ variable, the variable influence of factors affecting *stunting* in East Java in 2021 can be predicted and analyzed further.

3.3 Factor Analysis with *Principal Component Analysis* on Factors Data *Stunting* Based on City/Regency in East Java Province in 2021

Principal Component Analysis is carried out to determine the number of variables that can represent the diversity conditions of the variables tested, so that it can be estimated that fewer new factors will be formed. One way to determine the number of *principal components* is to look for an *eigenvalue* that has a value greater than one and a cumulative percentage value of variance of more than 70% and can be visually seen using a scree plot. The *eigenvalue* of the *stunting* factor data by city/district in East Java Province in 2021 is shown in Table 5.

Table 5. Value *Own Values* Data Factor Factor *Stunting* Based on City/Regency in East Java in 2021

Components	<i>Own Value</i>		
	Total	Varians (%)	Cumulative (%)
1	4,349	54,359	54,359
2	1,210	15,129	69,488
3	0,892	11,146	80,634
4	0,667	8,332	88,966
5	0,368	4,606	93,572
6	0,245	3,061	96,632
7	0,157	1,958	98,590
8	0,113	1,410	100,000

Table 5 shows that there are the first 2 main components that have *Own value* more than one. Component 1 has *Own value* of 4.349 with a percentage of diversity of 54.359% which means that component 1 is able to explain the variable of 54.359%. Component 2 has *Own value* of 1.210 with a percentage of diversity of 15.129% which means that component 2 is able to explain the variable of 15.129%. Overall the factors formed can explain 69.488% of the diversity possessed by the variables in the factors *stunting* based on the city/district of East Java Province in 2021. This shows that the two components are good enough to describe the complex data structure related to the factors *stunting*. As for visually, it can be visualized using *scree plot* presented in Figure 3.

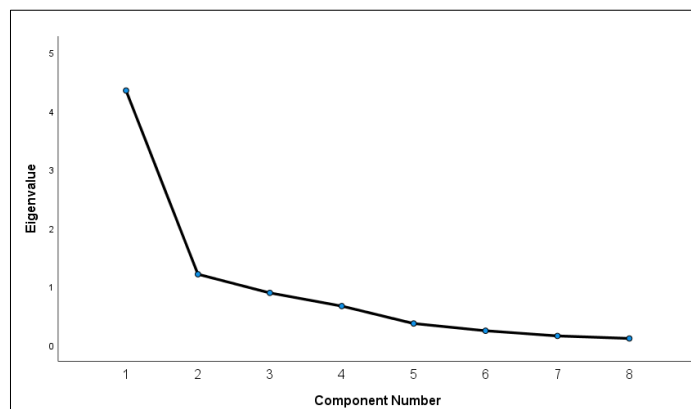


Figure 3 Scree Plot Multivariate Normal Distribution of Factors Data *Stunting* Based on City/Regency of East Java Province in 2021

Figure 3 visually shows a drastic or steep decline from component point 1 to component point 2. However, after passing the 2nd component point, the curve line has begun to slope, until point 8 the decline that occurs is more sloping. This is supported by points 1 to 2 having an eigenvalue of more than 1, while other factors have an eigenvalue of less than 1. The cumulative proportion of variability explained by the 2 factors is close to 70,000%, which is 69.488%. This shows that the first two main components are sufficient to explain most of the data variability, while the additional components contribute much less to the total variance. Therefore, the selection of the two main components is quite representative in summarizing information from the stunting factor data based on cities/districts of East Java Province in 2021.

Factor analysis was carried out to determine the grouping of variables into factors. Based on the previous main components, 2 main components were produced, so that in factor analysis, 8 variables were grouped into 2 factors. Before grouping, a community value analysis was carried out. The community value of stunting factors based on cities/districts of East Java Province in 2021 is shown in Table 6.

Table 6. Value of Data Commonality Factors *Stunting* Based on City/Regency of East Java Province in 2021

Variable	Initial	Extraction
X2	1,000	0,799
X3	1,000	0,806
X4	1,000	0,475
X5	1,000	0,803
X6	1,000	0,488
X7	1,000	0,845
X8	1,000	0,749
X9	1,000	0,593

Table 6 shows that the extraction results from the PCA analysis of variables related to *stunting* factors based on cities/districts in East Java in 2021. The *extraction* value describes the proportion of variance of each variable that is successfully explained by the two main components selected. The table shows that most of the variables have *extraction* values above 0.5; such as X2, X3, X5, X7, X8 and X9 have extraction values of 0.799; 0.806; 0.803; 0.845; 0.749 and 0.593. This shows that the main component is able to explain most of the variability of these variables. However, there are two variables, namely X4 and X6, that have *extraction* values below 0.5, namely 0.475 and 0.488. This shows that the variability of the two variables is poorly represented by the selected main component, so that these variables become less relevant in explaining the structure of the analyzed factors or require consideration to be further processed in the analysis. In addition, the highest *extraction value* was also obtained in the variable of the education index (X7) of 84.5% which means that the percentage of variability of the two factors formed can explain the educational index (X7) of 84.5% and the lowest extraction value is in the variable of the percentage of children aged 0-12 months who have been breastfed (X4) of 47.5% which means that the percentage of variability of the two factors formed can explain the percentage of children aged 0-12 months who have been breastfed (X4) of 47.5%.

The grouping of components in the data on factors affecting *stunting* based on cities/districts in East Java Province in 2021 based on the *Rotated Component Matrix* formed is shown in Table 7.

Table 7. Rotated Component Matrix on Factors Data *Stunting* Based on City/Regency in East Java in 2021

Variable	Components	
	1	2
X2	0,862	0,238
X3	0,861	-0,254

X4	0,089	0,684
X5	0,895	0,051
X6	0,071	0,695
X7	0,777	0,492
X8	-0,760	-0,413
X9	-0,636	-0,434

Table 7 shows the grouping of each variable into a specific factor based on the largest value of the matrix component of each variable. The table visualizes that factor or component 1 consists of variable X₂, X₃, X₅, X₇, X₈, and X₉. Factor 2 consists of variables X₄ and X₆. In addition, a summary of each of the factors formed can be seen in Table 8.

Table 8. New Factors in Data Factors *Stunting* Based on City/Regency in East Java in 2021

Factor	Variable	Variable Description
Public Welfare and Health	X2	Percentage of households that have access to proper sanitation
	X3	Percentage of complete immunization
	X5	Health index
	X7	Education index
	X8	Percentage of poor population
	X9	Percentage of the female population of early marriage
Basic Health and Sanitation Access	X4	The percentage of children aged 0-12 months has been breastfed
	X6	The percentage of households that have access to a decent source of drinking water

Table 8 shows that the naming of factors is carried out based on the variables included in the new factors. The first factor is called welfare and public health because the variables included in it explain the percentage of households with access to proper sanitation, the percentage of complete immunization, the health index, and the education index indicating the quality of health and education in an area. Meanwhile, the percentage of poor population and the percentage of women who marry early also contribute to the picture of social and economic welfare. This grouping reflects aspects of general well-being related to health, education, and socioeconomic conditions. The second factor is called basic health and sanitation access because the percentage of children aged 0-12 months who have been breastfed reflects the quality of basic health care and nutritional fulfillment, while the percentage of households with access to decent drinking water sources shows the quality of access to basic sanitation needs. Both are important aspects in determining the basic health conditions of a population.

The matrix transformation component serves to prove the magnitude of the correlation value of the factors formed in the data on *stunting* factors based on cities/districts in East Java Province in 2021. The following is the matrix transformation component in the variable data of *stunting* factors based on cities/districts in East Java Province in 2021 shown in Table 9.

Table 9. Transformation Components of the Stunting Factors Data Matrix by City/Regency of East Java Province in 2021

Components	1	2
1	0,923	0,385
2	-0,385	0,923

Table 9 shows that the correlation values in components 1 and 2 have values of 0.923 and 0.385 respectively. Component 1 has a value of 0.923 where the value is greater than 0.5 so it can be said that the factor 1 formed is said to be correct in summarizing the variables used in the data on *stunting* factors based on cities/districts in East Java Province in 2021. Component 2 has a value of 0.385 where the value is less than 0.5 so it can be said that the factor 2 formed is said to be inappropriate in summarizing the variables used in the data on *stunting* factors based on cities/districts in East Java Province in 2021. Next, modeling will be carried out on each variable based on the factors that have been formed as follows.

$$\begin{aligned}
 Z_2 &= 0,862F_1 + 0,238F_2 \\
 Z_3 &= 0,861F_1 - 0,254F_2 \\
 Z_4 &= 0,089F_1 + 0,684F_2 \\
 Z_5 &= 0,895F_1 + 0,051F_2 \\
 Z_6 &= 0,071F_1 + 0,695F_2 \\
 Z_7 &= 0,777F_1 + 0,492F_2 \\
 Z_8 &= -0,760F_1 - 0,413F_2 \\
 Z_9 &= -0,636F_1 - 0,434F_2
 \end{aligned}$$

4. CONCLUSION

Based on the results of the factor-based analysis *stunting* Based on cities/regencies in East Java Province in 2021, several conclusions were obtained as follows.

1. The variables of the percentage of child weight, the percentage of households that have access to proper sanitation, the percentage of complete immunization, the percentage of children aged 0-12 months who have been breastfed, the percentage of the poor population and the percentage of the female population who marry early have a percentage variation that is greater than the average value, which indicates that the distribution of these factors is uneven in East Java cities/districts in 2021. Meanwhile, the health index variables, the percentage of households that have access to proper drinking water sources and the education index have an even distribution in each city/district in East Java Province in 2021 because the variation value is smaller than the average value.
2. The results of the test of the analysis assumptions of factors affecting *stunting* in East Java in 2021 meet the assumptions of normal, dependent, both factored and predictable and analyzed.
3. The results of *the principal component analysis* produced 2 main components or factors that could explain the origin variable of 69.488% which is close to 70%. The results of the factor analysis produced two factors, namely welfare and public health factors with variables of the percentage of households with access to proper sanitation, the percentage of complete immunization, the health index, the education index, the percentage of the poor population and the percentage of women who married early. The second factor that was formed was the basic health factor and access to sanitation which consisted of the variables of the percentage of children aged 0-12 months who had been breastfed and the percentage of households with access to proper drinking water sources.

ACKNOWLEDGMENTS

Thank you to all parties involved in the compilation of this article.

FUNDING INFORMATION

There is no funding in this article.

AUTHOR CONTRIBUTIONS STATEMENT

Specify the contributions of each author and the role of using limited assistance from Artificial Intelligence (AI) as follows:

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Sella Nasrania	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Muhammad Zulfadhli		✓				✓		✓	✓	✓	✓	✓		
Dzakki Damar Rabbani	✓		✓	✓			✓			✓	✓		✓	✓
Abghaza Bayu K. W. P.		✓				✓		✓	✓	✓	✓			

C : Conceptualization

M : Methodology

So : Software

I : Investigation

R : Resources

D : Data Curation

Vi : Visualization

Su : Supervision

P : Project administration

Va : **V**alidation
Fo : **F**ormal analysis

O : Writing - **O**riginal Draft
E : Writing - Review & **E**ding

Fu : **F**unding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

- [1] W. C. Black, B. J. Babin, J. F. Hair, and R. E. Anderson, *Multivariate Data Analysis*, 7th ed. Upper Saddle River: Prentice Hall, 2010.
- [2] J. Frost, "Statistics by Jim: Mean, Median, and Mode – Measures of Central Tendency," 2022.
- [3] R. Hidayat and S. D. S. D., "Evaluation of Variable Correlation with Anti Image Correlation in Factor Analysis," *Indonesian Journal of Statistics and Applied Methods*, 2017.
- [4] S. Indrawati and H. R. H. R., "Analysis of Data Variability with Variance as an Indicator of Inhomogeneity in Social Research Data," *Journal of Social Sciences and Humanities*, 2019.
- [5] L. M. Nasution, "Descriptive Statistics," *Journal of Wisdom*, 2017.
- [6] D. Pratama and D. M. D. M., "Factor Analysis and Communality Test: Application in Social Data," *Journal of Indonesian Applied Statistical Methods*, 2017.
- [7] D. Pratama and K. Y. K. Y., "Application of Main Component Analysis in Social and Economic Research," *Indonesian Journal of Economics and Statistics*, 2020.
- [8] H. Pratama and K. R. K. R., "Component Matrix Analysis and Rotation in Principal Component Analysis," *Indonesian Journal of Science and Technology*, 2019.
- [9] T. Pratama and H. L. H. L., "Socioeconomic Influence on Complete Immunization Coverage in Children Under Five," *Indonesian Journal of Nutrition and Health*, 2019.
- [10] Y. Pratama and L. A. L. A., "Factor Analysis in Social and Economic Research: Concepts and Applications," *Indonesian Journal of Social and Economic Sciences*, 2018.
- [11] Y. Pratama and N. A. N. A., "Factor Analysis and Total Variance Explained Test in Social Research," *Indonesian Journal of Applied Social Statistics*, 2019.
- [12] A. Putra and S. R. S. R., "Determinant Analysis of Poverty in Indonesia: A Case Study in Central Java Province," *Journal of Economics and Public Policy*, 2019.
- [13] M. Putra and R. H. R. H., "The Role of Eigenvalue in Factor Selection in Key Component Analysis," *Indonesian Journal of Applied Mathematics and Statistics*, 2018.
- [14] R. Putra and S. D. S. D., "Basic Assumptions in Factor Analysis and Their Application to Social Data," *Indonesian Journal of Social Statistics*, 2016.
- [15] N. Rahayu and S. D. S. D., "Early Marriage and Socio-Economic Factors Affecting It in Indonesia," *Indonesian Journal of Gender Studies*, 2017.
- [16] D. Rahmadani and N. S. N. S., "The Effect of Access to Decent Drinking Water on Public Health in Indonesia," *Indonesian Journal of Public Health*, 2017.
- [17] D. Rahmawati and L. I. L. L., "Nutritional Status of Children Under Five in Indonesia Based on Weight Percentage," *Indonesian Journal of Nutrition and Health*, 2018.
- [18] L. Rahmawati and P. Y. P. Y., "Stunting in Toddlers: Challenges and Countermeasures in Indonesia," *Indonesian Journal of Nutrition and Health Development*, 2019.
- [19] S. Rahmawati and K. D. K. D., "Factors Influencing Exclusive Breastfeeding in Indonesia," *Indonesian Journal of Public Health*, 2018.
- [20] T. Santoso and N. P. N. P., "Application of Principal Component Analysis in Reducing Social Data Dimensions," *Indonesian Journal of Statistics and Applications*, 2015.
- [21] I. Sari and P. H. P. H., "Application of Q-Q Plot in Testing Normality Assumptions on Social Research Data," *Indonesian Journal of Social Statistics*, 2016.
- [22] N. Sari and N. D. N. D., "Evaluation of Health Index for Prioritization of Health Programs in Indonesia," *Indonesian Journal of Public Health*, 2018.
- [23] B. Setiawan and R. T. R. T., "Statistical Analysis at Maximum Values: A Case Study of Water Quality in Indonesia," *Indonesian Environmental Journal*, 2018.
- [24] A. Siregar and L. W. L. W., "Application of the Kaiser-Meyer-Olkin Test (KMO) for Factor Analysis in Social Research," *Journal of Indonesian Applied Statistical Methods*, 2016.
- [25] T. Sukmana and W. R. W. R., "Evaluation of the Education Index in Indonesia and the Factors Influencing It," *Indonesian Journal of Education*, 2017.
- [26] A. Sumedi and U. F. U. F., "Risk Factors for Stunting in Children Under Five in Indonesia," *Indonesian Journal of Public Health*, 2017.
- [27] D. Suryani and R. I. R. I., "The Use of the Bartlett Test in Testing Data Suitability for Factor Analysis," *Indonesian Journal of Statistics*, 2017.
- [28] D. Susanti and R. T. R. T., "Analysis of Minimum and Maximum Values on Extreme Weather Observation Data in Indonesia," *Indonesian Journal of Meteorology and Climatology*, 2016.
- [29] M. Zulfadhli, I. N. Budiantara, V. Ratnasari, and A. S. Suriaslan, "Fourier series nonparametric regression estimator for modeling status of unmet need in East Java Province in 2023," *Commun. Math. Biol. Neurosci.*, vol. 2025, Art. ID 2025.

- [30] A. S. Suriaslan, I. N. Budiantara, V. Ratnasari, M. Zulfadhli, and R. C. Dhewy, "An Empirical Study on Unmet Need: A Statistical Inference Framework for Truncated Spline Nonparametric Binary Logistic Regression (TSNLBR)," *J. Cultural Anal. Social Change*, pp. 868–877, 2025.
- [31] M. Zulfadhli, I. N. Budiantara, V. Ratnasari, A. S. Suriaslan, and R. C. Dhewy, "A statistical inference framework for FSNBLR: Modeling underdeveloped regional status in Eastern Indonesia," *MethodsX*, p. 103746, 2025.
- [32] M. Zulfadhli, I. N. Budiantara, V. Ratnasari, and A. S. Suriaslan, "Semiparametric Regression Estimator of Fourier Series For Categorical Data," *IAENG Int. J. Appl. Math.*, vol. 55, no. 7, pp. 2157–2164, 2025.
- [33] B. W. Otok, M. Zulfadhli, R. D. Pangesti, M. I. Kurniawan, A. E. P. Haryanto, and I. Kurniawan, "Comparison of Binary Probit Regression and Fourier Series Nonparametric Logistic Regression in Modeling Diabetes Status at Hajj General Hospital Surabaya," *Barekeng*, vol. 20, no. 1, pp. 255–270, 2026.
- [34] A. M. Achmad, M. Zulfadhli, and N. Nurfaudzan, "Binary Logistic Regression on the Status of Unemployment Rate in West Java Province in 2023," *Likelihood: J. Stat. Its Appl.*, vol. 1, no. 1, pp. 30–34, 2025.
- [35] M. Zulfadhli, I. N. Budiantara, and V. Ratnasari, "Nonparametric regression estimator of multivariable Fourier series for categorical data," *MethodsX*, vol. 13, p. 102983, 2024.