

FACTOR ANALYSIS OF DENGUE HEMORRHAGIC FEVER

Nisriina Nazhiifah¹ , Bagus Sumargo2*, Fareka Erdien³ , Siti Julpia Kirana⁴ , Widyanti Rahayu⁵ , Mulyono⁶

*1,2,3,4,5Department of Statistics, Faculty of Mathematics and Natural Sciences, Jakarta State University Rawamangun Muka St., East Jakarta 13320, Indonesia ⁶Department of Computer Science, Faculty of Mathematics and Natural Sciences, Jakarta State University Rawamangun Muka St., East Jakarta 13320, Indonesia Corresponding author's e-mail: * bagussumargo@unj.ac.id*

ABSTRACT

Received: 18 May 2024 Revised: 6 June 2024 Accepted: 13 June 2024 Published: 30 June 2024 Available online.

Keywords:

Dengue Hemorrhagic Fever, Principal Component Analysis, Environmental Factors

Article History: **Dengue hemorrhagic fever (DHF) is a disease caused by the dengue virus. This study aimed to** determine the factors that cause Dengue Hemorrhagic Fever (DHF) in the province of Central Java because cases of DHF in the region become a serious case every year. The method used for this research is Principal Component Analysis and Factor Analysis using secondary data from the Central Java Provincial Health Office in 2018. There are seven variables that are considered relevant, other variables that may also have an influence are not included in this analysis. The analysis results indicate that three factors, such as population participation in health, sanitation, and clean drinking water, are the primary causes of DHF. Therefore, it is important to increase population awareness in terms of healthy living and local government intervention is needed in environmental health projects, such as sanitation and clean drinking water. This approach offers a fresh and distinct perspective on understanding the dynamics of dengue fever and the factors that contribute to its reduction.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-ShareAlike 4.0 International License.](http://creativecommons.org/licenses/by-sa/4.0/)

How to cite this article: N. Nazhiifah, B. Sumargo, F. Erdien, S.J. Kirana, W. Rahayu, Mulyono, "FACTOR ANALYSIS OF DENGUE HEMORRHAGIC FEVER", Jurnal Statistika dan Aplikasinya, vol. 8, iss. 1, pp. 75 – 86, June 2024

Copyright © 2024 Author(s) Journal homepage: *<https://journal.unj.ac.id/unj/index.php/statistika>* Journal e-mail: *jsa@unj.ac.id* Research Article ∙ Open Access

1. INTRODUCTION

Many factors have contributed to the increase in the dengue hemorrhagic fever virus, including inadequate water distribution and waste management, uncontrolled urbanization, globalization, and increasing plastic waste [1]. The results of previous studies also show that there are factors that influence the increase in dengue cases in the province of Central Java, namely the number of health centers per 1,000 people, the number of village maternity huts per 1,000 people, population density, the percentage of the population that has access to sustainable drinking water, the percentage of clean water, the quality of bacteria, fungi, and chemicals, and the number of protected spring facilities [2].

Previous research on dengue hemorrhagic fever in Central Java has also been conducted using Geographically Weighted Negative Binomial Regression (GWNBR). The results of the study concluded that the number of protected spring facilities influenced the increasing dengue cases in all areas of Central Java. Households with clean and healthy living behaviors and healthy houses were also factors in the high number of dengue cases in Central Java [3].

The health condition of the people of Central Java through the variables generated from the Susenas Kor (Survei Sosial Ekonomi Nasional Inti) according to social and economic characteristics is seen as a picture of health in Central Java. DHF cases have always been a serious problem every year in Central Java [3]. From 2016 to 2018, the number of Dengue hemorrhagic fever (DHF) cases in Central Java decreased, but in 2019 it increased compared to the previous year. The number of cases of DHF in 2018 was 3,519 cases and 37 deaths, of which mortality reached 1.05%, still higher than the national target of less than 1% [4].

Dengue hemorrhagic fever (DHF) is a disease caused by the dengue virus. The flavivirus genus and Flaviviridae family include the dengue virus. There are four types of viruses that have been defined, namely DENV-1, DENV-2, DENV-3, and DENV-4 [1]. Dengue fever can be identified by several symptoms such as fever, joint pain, muscle pain, bleeding diathesis, decreased pulse pressure, hypotension, and cyanosis, up to severe shock [5].

One of the factors that causes increasing DHF is the high rainfall and humidity in Indonesia. Rainfall is the amount of rainwater that gathers in flat places, does not seep, evaporates, and flows. In addition, humidity is the total amount of water vapor in the air [6]. The increasing number of dengue cases is not only due to increased rainfall but also due to the earth's temperature rising during the rainy season. This triggers the breeding of A. aegypti mosquitoes [7].

The purpose of this study was to identify the main factors causing a lower case of DHF in 2018, especially in Central Java Province. The year 2018 became the focus of our research because in that year there was a decrease in cases from several previous years. This research can be useful for the community as an early prevention strategy because the main factors can be known from the results of factor analysis research. This research could be useful in finding the main factor for the small number of dengue cases in 2018 in Central Java province.

2. METHODS

The data used in this study is secondary data obtained from the Central Java Province BPS in 2018. The response variable used in this study is the number of dengue cases in 35 districts/cities of Central Java Province in 2018. The following are the variables used in the study: [3].

The methods used in this research are principal components and factor analysis [8]. The principal components are used in the factor analysis method. The principal component will form a linear combination of the independent variables and produce new factors. This method considers the total variance in the cluster data.

Let covariance matrix Σ with eigenvalue and eigenvector pairs $(\lambda_1, e_1), (\lambda_1, e_2), ..., (\lambda_p, e_p)$ where $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_p \geq 0$. Then, by considering the linear combination, the principal component model form is as follows.

$$
Y_1 = e_1'X = e_{11}X_1 + e_{12}X_2 + \dots + e_{1p}X_p Y_2 = e_2'X
$$

= $e_{21}X_1 + e_{22}X_2 + \dots + e_{2p}X_p$: $Y_p = e_p'X$
= $e_{p1}X_1 + e_{p2}X_2 + \dots + e_{pp}X_p$ (1)

Then,

$$
Var(Y_i) = e_i' \sum e_i = \lambda_i
$$
 (2)

$$
Cov(Y_i, Y_k) = e_i' \sum e_k = 0 \tag{3}
$$

where

 $e_1, e_2, \dots, e_p = \text{eigenvector with } i = 1, 2, \dots, p \text{ and } i \neq k$

 λ_i = eigenvalue

Factor analysis aims to reduce the dimensions of the data by expressing the original variable as a linear combination of common factors and specific factors.

The model of factor analysis is

$$
X_1 - \mu_1 = \ell_{11}F_1 + \ell_{12}F_2 + \dots + \ell_{1m}F_m + \varepsilon_1
$$

\n
$$
X_2 - \mu_2 = \ell_{21}F_1 + \ell_{22}F_2 + \dots + \ell_{2m}F_m + \varepsilon_2
$$

\n
$$
\vdots \qquad \vdots
$$

\n
$$
X_p - \mu_p = \ell_{p1}F_1 + \ell_{p2}F_2 + \dots + \ell_{pm}F_m + \varepsilon_p
$$

Or in the form of matrix notation, it is as follows.

$$
X - \mu (p \times 1) = L (p \times m) F (m \times 1) + \varepsilon (p \times 1)
$$
 (4)

where

p = The number of variables

 μ = Mean of p_{th} variable

- l_{ij} = Loading of the i_{th} variable on the j_{th} factor
L = Factor loading matrix
- $=$ Factor loading matrix
- ε_i = Specific factors
 F_m = Common factor
- = Common factors
- $i = 1, 2, ..., p$
- $j = 1, 2, ..., q$

Data Standardization

Standardization of data is conducted if there are differences in the size of the data unit, so the data in this research has unit uniformity.

Kaiser-Meyer-Olkin (KMO) Assumption Test

The KMO test determines how suitable the data is for factor analysis. The data for factor analysis is accessible if the KMO value is greater than 0.5. The KMO value is given by the following equation:

$$
KMO = \frac{\sum_{i}^{p} \sum_{j \neq i}^{p} r_{ij}^{2}}{\sum_{i}^{p} \sum_{j \neq i}^{p} r_{ij}^{2} + \sum_{i}^{p} \sum_{j \neq i}^{p} a_{ij}^{2}}
$$
(5)

where r_i^2 correlation coefficient i_{th} variable and j_{th} variable, a_{ij}^2 partial correlation coefficient j_{th} variable and j_{th} variable, with $i = 1, 2, ..., p$ and $j = 1, 2, ..., q$

Bartlett Assumption Test

The Bartlett test aims to determine the correlation between the independent variables. The decision from the Bartlett test is to reject hypothesis H0 (the null hypothesis) if $X_{hit}^2 > X_{\alpha, \frac{p(p-1)}{2}}^2$ or

 $p - value < \alpha$ which means there is a correlation between the independent variables, and vice versa. The test statistic is as follows.

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

where p is the number of variables, *|R|* is the determinant of the correlation matrix, and *n* is the number of samples.

Measure of Sampling Adequacy (MSA) Assumption Test

The MSA test aims to determine the feasibility of the variables studied. The MSA value can be calculated using the following equation:

$$
MSA = \frac{\sum_{i}^{p} \sum_{j \neq i}^{p} r_{ij}^{2}}{\sum_{i}^{p} \sum_{j \neq i}^{p} a_{ij}^{2}}
$$
(7)

Communality

The value of communality indicates the magnitude of the diversity or variance of variables that can be explained by new factors. The communality value requirement is met when the value is greater than 0.5. If the communality value is less than 0.5, it's considered unimportant and can be excluded from the analysis and refactored. The value of communality can be calculated using the following formula:

$$
h_i^2 = l_{i1}^2 + \dots + l_{im}^2 \tag{8}
$$

Determining the number of factors with Principal Component Analysis

In determining the number of factors, the Principal Component Analysis (PCA) method is used to reduce the number of variables to new factors. The determination of the number of factors is conducted using the criteria of eigenvalues and the total variety of main components. With PCA, each variable contributes a value of 1 to the total eigenvalue, so it is said to be significant if the eigenvalue is greater than 1. The formula for finding the eigenvalue is as follows:

$$
|R - \lambda I| = 0 \tag{9}
$$

where R = correlation matrix, λ = the eigenvalue, and *I* = identity matrix.

The total variety of key components states that the proportion of the diversity of the primary elements represents the diversity of the data, and the total of the main components is considered excellent when the proportion of diversity is 80-90% [8]. Then, the total population diversity contributed to the primary components is:

$$
\frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \tag{10}
$$

Rotation Factor

Factor rotation is used to assist with interpreting factors. The approach used in factor rotation is orthogonal rotation, which means rotating along a perpendicular axis so that each factor does not correlate with the others. The varimax orthogonal rotation method is used in orthogonal rotation because it provides a high loading factor on only one factor while another loading factor approaches zero.

The estimated value of the factor loading of the primary components is:

$$
\ell_{ij} = \sqrt{\lambda_i e_{ij}} \tag{11}
$$

The diagonal matrix $S - LL'$ is a matrix for a specific variance.

$$
\psi = \begin{pmatrix} \psi_1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_p \\ \end{pmatrix}_{\text{with } \psi_i = 1 - h_i^2}
$$
 (12)

After rotating factors and knowing the number of factors formed, the naming and interpretation of each factor are conducted.

3. RESULTS

Description of Dengue Hemorrhagic Fever Disease

Dengue hemorrhagic fever (DHF) is an acute infection caused by an arbovirus (arthropod-borne virus) and transmitted by Aedes mosquitoes (Aedes albopictus and Aedes aegypti) [5]. DHF is a disease that is widely found in large parts of the tropics and subtropics, especially Southeast Asia, Central America, America, and the Caribbean. The natural host of DHF is humans. The agent is a dengue virus belonging to the family Flaviviridae and the genus Flavivirus, consisting of 4 serotypes, namely DENV-1, DENV-2, DENV-3, and DENV-4, transmitted to humans through the bites of infected mosquitoes, in particular the Aedes aegypti and Aedes albopictus that are found in almost all corners of Indonesia [1].

Symptoms of dengue hemorrhagic fever disease begin with: 1) sudden high fever of 2-7 days (38°C-40°C); 2) manifestations of bleeding, with a positive tourniquet test, purpura, conjunctival bleeding, epistaxis, melena; 3) hepatomegaly; 4) shock, pulse pressure decreases to 20 mmHg or less, systolic pressure reaches 80 mmHg or less; 5) thrombocytopenia, from days 3-7 a decrease in the platelets up to 100,000/mm; 6) hemoconcentration, increased hematocrit values; 7) other clinical symptoms that can accompany anorexia, nausea, vomiting, weakness, abdominal pain, convulsive diarrhea, and headache; 8) and pain in the muscles and joints [5]. Risk factors that can influence dengue fever include home environment, biological environment, and social environment.

Previous research on dengue hemorrhagic fever in Central Java has also been conducted using Geographically Weighted Negative Binomial Regression (GWNBR). The results of the study concluded that the number of protected spring facilities has an influence on increasing dengue cases in all regions of Central Java; households with clean living behaviors and healthy homes are also factors in the high number of dengue cases in Central Java [3].

Region Description

Figure 1 shows the spread of dengue cases in each regency/city in Central Java Province in 2018, covering 29 regencies and 6 cities with a total of 3,519 cases. The most cases are in Grobogan Regency, with the number of cases reaching 460, and the lowest cases are in Tegal City, with 5 cases. It turns out that the fastest spread with a high risk of developing dengue disease is in the eastern part of Central Java province.

Figure 1. Spread of Dengue Cases in Central Java Province 2018

Assumption Test

Data on dengue cases in 35 regencies/cities of Central Java Province in 2018 and seven attributes will be used for this analysis. The variables to be analyzed must correlate with each other. Data on variables need to be formed in the form of a correlation matrix. Then continued with the-Kaiser Meyer-Olkin Test (KMO) after a correlation matrix was formed. The KMO test is conducted to find out whether the factor analysis process can be used in interfacing existing data.

From the results of the SPSS output calculation, a KMO value of 0.565 is obtained, which means it is greater than 0.5. Furthermore, the Bartlett Test is conducted. In the Bartlett Test, it can be determined whether the variable correlates or not by looking at the value of the level of significance obtained. The results of the SPSS output calculation were obtained with a Bartlett Test significance value of 0.000, meaning that it is smaller than 0.05. The correlation between variables is high, and the factor analysis process can be continued.

Table 3. MSA Assumption Test Results

MSA results can be seen from anti-image matrices. The MSA magnitude of each attribute is indicated by a row of lift diagonals marked "a," where each attribute's MSA value cannot be less than 0.5. Based on the test results using SPSS software, it is known that there are variables that must be issued because the MSA is less than 0.5, namely 0.479 in the variable (the number of Village Maternity Huts). Because there are variables that don't meet the MSA value requirements, they must be excluded from the test and must repeat the steps to find out the KMO value.

The KMO Test and Bartlett Test repetitions were performed on data without X6 variables using SPSS software. Based on the results of these calculations, the KMO value is 0.629, which means it is greater than 0.5. That means testing can continue. It was also later discovered that the Bartlett Test value of 0.000 means less than 0.05, so the correlation between high variables and the factor analysis process can be continued.

Anti-Image Matrices							
		X_1	X_2	X_3	X_4	X_5	X_7
	X_1	.541 ^a	.063	$-.382$.033	.134	$-.094$
Anti-Image Correlation	X_2	.063	.707 ^a	$-.336$	$-.304$	$-.191$.099
	X_3	$-.382$	$-.336$.656 ^a	.020	$-.298$	$-.182$
	X_4	.033	$-.304$.020	$.615^{\circ}$	$-.161$	$-.571$
	X_5	.134	$-.191$	$-.298$	$-.161$.658 ^a	.170
	X_7	$-.094$.099	$-.182$	$-.571$.170	.562 ^a

Table 5. Second MSA Assumption Test Results

The MSA results were reviewed using Anti-Image Matrices. After this calculation, no MSA value is less than 0.5 in all variables. That means there's no need to emit any variables, and testing can continue.

Communality

The value of communality indicates the magnitude of the diversity or variance of variables that can be explained by new factors or components. Based on Table 6, it can be known that the whole variable meets the requirements of communality when its value is greater than 0.5. Table 6 also shows the total variance of variables that can be attributed to the resulting factor. Household variables using clean drinking water sources have the largest value of 0.893, or around 89.3% of the variances can be explained by the factors formed. The population density variable has the lowest value of 0.65, or about 65% of the variance of the variable can be explained by the factor formed. This also applies to other variables; the greater the extraction value, the closer it is related to the factor or component formed.

Determining the number of factors with Principal Component Analysis

The total variance explained indicates the number of factors formed using the Eigenvalue rule, where the fulfilling eigenvalue is ≥ 1 . Based on Table 7, the factors of 1,2,3 have met the requirements with an eigenvalue value of ≥ 1 . So, the number of factors formed is 3 factors, with an eigenvalue in component 1 of 2.476, an eigenvalue in component 2 of 1.157, and an eigenvalue in component 3 of 1,038. From the Table 7, it can also be seen that the total of the three factors to be formed can explain the cumulative variance of variables of 77.854%.

Figure 2. Scree Plot

The results of factor extraction can also be shown through the Scree plot with an eigenvalue of \geq 1. It can be seen from the graph that 3 components or factors also have an eigenvalue ≥ 1 .

Factor Rotation and Interpretation

The rotated component matrix aims to look at the variables that are in the factor. The method used in the rotation of this matrix is the method of orthogonal rotation of varimax.

Table 8 shows that three factors have been formed. The component value seeks the highest loading value for each variable, which will determine the dominant factor. Households using clean drinking water sources have the highest loading value at factor 3, which is 0.740, indicating that the variable is included in factor 3. This also holds for other variables. So, the factor formed and its variables, namely factor 1, consist of a population density variable with a loading value of 0.753, households with a clean and healthy living behavior variable with a loading value of 0.606, and the variable number of active integrated health posts with a loading value of 0.851, respectively. Factor 2 involves healthy houses (loading value 0.841) and a household variable with adequate sanitation access (loading value 0.896). Factor 3 involves household variables that use clean drinking water sources and has a loading value of 0.740.

According to several test steps, there are three factors with six variables, as follows:

• Factor 1, or Population Participation in Health, consists of several variables, namely:

- 1. Population Density (X_2)
- 2. The Percentage of Households with Clean and Healthy Living Behavior (X_3)
- 3. The Percentage of Integrated Health Post Active (X_5)
- Factor 2, or Sanitation, consists of several variables, namely:
- 1. The Percentage of Healthy Houses (X_4)
- 2. The Percentage of Households with Adequate Sanitation Access (X_7)
- Factor 3, or Clean Drinking Water, consists of one variable, namely:
- 1. The Percentage of Households using Clean Drinking Water Sources (X_1)

4. DISCUSSIONS

The results show that three factors are causing DHF, namely population participation in health, sanitation, and clean drinking water. This finding aligns with data showing a decreasing trend in 2018 in Central Java Province. The results indicate that high population participation in health, success in sanitation, and the use of clean drinking water are key factors in reducing DHF cases. This is consistent with the findings of previous studies, such as those by [9], which highlighted that access to clean water, sanitation, and waste management can reduce both waterborne diseases and mosquitoborne arboviral infections, which aligns with our findings. Furthermore, [10] hygiene and environmental sanitation are effective in mitigating dengue hemorrhagic fever.

The explanation of the results shows that the government's integrated approach to health promotion, sanitation, and clean water provision effectively reduces DHF cases. These findings confirm the importance of multi-faceted public health strategies in combating infectious diseases. High population participation in health and sanitation programs directly correlates with increased awareness and preventive behaviors, thereby reducing mosquito breeding sites and transmission rates. From the main factors that have been obtained, it turns out that support the success of the government through programs to manage cases of dengue fever. The low number of dengue cases in 2018 could be due to high population participation in health, success in sanitation, and the increasing number of households using clean water sources through government programs such as 3M (Menutup, Menimbun, dan Menguras) plus cleaning gutters, the Jumantik movement (larva monitoring), fogging, community-based total sanitation village development, and community-based drinking water and sanitation [11].

5. CONCLUSION

It can be concluded that of the 7 variables studied, only 6 variables were involved in further analysis, namely population density, the percentage of households with clean and healthy living behaviors, the percentage of integrated health post-active, the percentage of healthy houses, the percentage of households with access to adequate sanitation, and the percentage of households using clean drinking water sources. Based on these 6 variables, 3 factors were formed that caused the low number of dengue cases in 2018 in Central Java, namely population participation in health, sanitation, and clean drinking water. Factors of population participation in health consist of population density, the percentage of households with clean and healthy living behaviors, and the percentage of integrated health post-active. The sanitation factor consists of the percentage of healthy houses and the percentage of households with access to adequate sanitation. The clean drinking water factor consists of the percentage of households using clean drinking water sources.

6. REFERENCES

- [1] World Health Organization, *Comprehensive Guidelines for Prevention and Control of Dengue and Dengue Hemorrhagic Fever.* WHO Regional Office for South-East Asia.: SEARO Publications, 2011.
- [2] I. F. Fatati, H. Wijayanto, and A. M. Sholeh, "Analisis Regresi Spasial dan Pola Penyebaran pada Kasus Demam Berdarah Dengue (DBD) di Provinsi Jawa Tengah," *Media Statistika*, vol. 10, no. 2, p. 95, Dec. 2017, doi:
- [3] B. Sumargo, S. J. Kirana, and S. R. Rohimah, "Dengue Hemorrhagic Fever Modeling using Geographically Weighted Negative Binomial Regression," 2023, p. 050013.
- [4] Province Health Office of Central Java, "Health Profile of Central Java 2018."
- [5] S. S. Walia, M. A. Arif, and J. Liaqat, "Clinical manifestations and laboratory diagnosis," in *Dengue Virus Disease*, Elsevier, 2020, pp. 115–137.
- [6] M. M. Sintorini, N. Aliyyah, and E. R. Sinaga, "Environment Drivers Of DHF Disease In Jakarta 2017 – 2018," *International Journal of Scientific & Technology Research*, vol. 9, no. 1, Jan. 2020.
- [7] A. Lahdji and B. B. Putra, "Hubungan Curah Hujan, Suhu, Kelembaban dengan Kasus Demam Berdarah Dengue di Kota Semarang," *Syifa' Medika*, vol. 8, no. 1, Sep. 2017.
- [8] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis.*, 6th ed. Upper Saddle River, 2007.
- [9] Overgaard, H.J., Dada, N., Lenhart, A.E., Stenström, T.A., & Alexander, N.D. "Integrated disease management: arboviral infections and waterborne diarrhoea." *Bulletin of the World Health Organization, 99*, 583 - 592. 2021.
- [10] Mufti, Muhammad Rafi. "Assessing the effectiveness of CBTS in mitigating dengue hemorrhagic fever: A literature review." *World Journal of Advanced Research and Reviews*. 2023.
- [11] N. A. Pascawati, S. Sahid, S. Sukismanto, and H. Yuningrum, "Faktor yang Berhubungan dengan Pola Pengelompokkan Kasus Demam Berdarah Dengue (DBD) di Temanggung, Jawa Tengah," *Balaba: Jurnal Litbang Pengendalian Penyakit Bersumber Binatang Banjarnegara*, pp. 65–78, Jul. 2022