Jurnal Riset Mahasiswa Matematika
Volume 4 (5) (2025), Pages 218-228
p-ISSN: 2808-1552; e-ISSN: 2808-4926

Canonical Correlation Modeling on Population Growth Rate and Life Expectancy with Socio-Economic Factors

M. Rizky Ramadhan* and Gusmi Kholijah

Mathematics, Faculty of Science and Technology, University of Jambi

Abstract

People's quality of life, such as population growth rate and life expectancy, is influenced by various interrelated socio-economic factors. The objective of this study was to investigate the bidirectional relationship between population growth rate and life expectancy with economic growth rate, labor force participation rate, proportion of poor, and population density. Canonical correlation analysis, which is one of the multivariate analysis, was used to assess the simultaneous relationship between two groups of variables. The results show that there is a statistically significant canonical function with a very strong relationship between the combination of socio-economic variables with population growth rate and life expectancy. The study also provides results for bidirectional relationships between the two groups of variables. The use of the canonical correlation approach for this analysis captures bidirectional relationships between dimensions of development that are rarely considered simultaneously in a regional context. This study provides a theoretical and empirical basis for formulating policies in improving community welfare such as population growth rate, life expectancy and socioeconomic indicators in Jambi Province.

Keywords: Canonical correlation, Multivariate analysis, and Quality of life.

1 Introduction

Canonical correlation analysis is one of the multivariate analyses used to identify and measure the relationship between two groups of variables, each consisting of more than one variable [1]. Canonical correlation aims to find linear combinations of each group that have maximum correlation with each other. This method has the advantage of revealing simultaneous, complex and not always linear relationships between groups of variables [2]. Canonical correlation analysis is used to examine the relationship between two sets of variables, where human development indicators such as population growth rate and life expectancy can be used as Y variables and socio-economic factors as X variables.

Human development is a multidimensional process that involves improving the quality of life through better health, education and a decent standard of living. The Human Development Index (HDI) of the United Nations Development Program (UNDP) measures human development in three main dimensions: long and healthy life, knowledge, and a decent standard of living [3]. In this context, life expectancy and population growth rate are important indicators that reflect the quality of health and demographic dynamics of a region.

Population growth rate is a number that shows the rate of population growth in numbers over a period of time [4]. The population growth rate reflects changes in population numbers that can affect resource

*Corresponding author. E-mail: rizkibe491@gmail.com

Submitted: 10 Mei 2025 Reviewed: 25 Mei 2025 Accepted: 5 Juni 2025
DOI: https://doi.org/10.18860/jrmm.v4i5.33034

availability, development policies, and environmental management. According to data from the Central Bureau of Statistics, the population growth rate in Jambi Province gradually decreased from 2010 to 2023, and reached a maximum of 0.78% in 2021 [5]. At the same time, life expectancy provides information about general health conditions, access to health services and general quality of life [6].From 2010 to 2023, the life expectancy of the population in Jambi Province has increased every year [5]. Population growth rate and life expectancy are the main indicators for assessing the quality and dynamics of population welfare, while socio-economic factors such as economic growth rate, labor force participation rate, poverty rate, and population density play an important role in shaping these conditions. By combining these two sets of variables using canonical correlation modeling, complex multidimensional patterns of relationships can be identified, leading to a more comprehensive understanding of how socio-economic conditions affect aspects of population and public health

research on the welfare index and economic growth using canonical correlation. The results show that the human development index has the strongest contribution to economic growth, with a canonical correlation of 73% and the variation explained by the first canonical function of 88% [7]. Another study that discussed canonical correlation analysis on health behavior and socioeconomic characteristics showed that there was a significant correlation between the two groups of variables, with education and expenditure as the main contributors from the socioeconomic side [8].

The main focus of this study is to identify and analyze the multivariate relationship between a number of socioeconomic variables and two key human development indicators, namely population growth rate and life expectancy, in Jambi Province. The canonical correlation analysis method was used not only to explore the interrelationships between groups of variables, but also to form a canonical model that is able to explain the extent to which linear combinations of socioeconomic variables are significantly correlated with linear combinations of human development indicators.

Unlike previous studies such as which focused on the relationship between the aggregate Human Development Index and economic growth, and which emphasized health behavior and socio economic characteristics, this study offers novelty in two main aspects. First, it specifically examines the relationship between two key indicators of human development population growth rate and life expectancy as a unified set of dependent variables (Y), and a more segmented group of socio economic variables, namely economic growth rate, labor force participation rate, poverty rate, and population density as the set of independent variables (X). Second, this study is geographically focused on the context of Jambi Province, utilizing time series data covering the period from 2010 to 2023, which has been rarely explored in previous research. Through this approach, the study provides a new contribution to multivariate modeling that is both simultaneous and contextual, and can serve as an empirical basis for formulating more targeted regional development policies.

2 Methods

This research uses non-experimental quantitative analytics, with a multivariate approach through the canonical correlation method. This research relies on observational data from secondary sources that are already available, making it suitable for application in the socio-economic context of the region. The research data was sourced from the website of the Jambi Provincial Statistics Agency [1] and have been published. The data used are data on population growth rate, life expectancy, economic growth rate, labor force participation rate, percentage of poor people, and population density. The data taken is from 2010 to 2023 and consists of 9 regencies and 2 cities in Jambi province.

Canonical correlation modeling is carried out using the help of R software. The steps of the canonical correlation are as follows:

1. **Determine the variable**
   The variables used in this study are population growth rate (Y_1) and life expectancy (Y_2) as dependent variables, while economic growth rate (X_1), labor force participation rate (X_2), poor pop-

[1] https://jambi.bps.go.id

ulation (X_3), and population density (X_4) as independent variables. All variables used are ratio scaled.

2. **Conducting Assumption Tests**
   a. The linearity test aims to determine whether there is a linear relationship between the dependent variable and each independent variable to be tested [9]. This test is generally used as a requirement in correlation or linear regression analysis.
   b. The normality test is carried out to determine the certainty of the distribution of the data obtained [10]. The normality test method on multivariate is very difficult to do, so testing will be done by checking the normality of each variable. The method that can be done is to use the Kolmogorov Smirnov Test. The hypothesis is as follows:
   H_0 : Normally distributed data
   H_1 : Data is not normally distributed Statistical test:

$$D = maksimum|F_0(X) - S_N(X)| \tag{1}$$

   with:
   D : The statistical value of the Kolmogorov-Smirnov test
   $F_0(X)$: The predetermined cumulative frequency distribution function
   $S_N(X)$: The observed cumulative frequency distribution function
   The decision making criteria are if $D_{\text{max}} < D_{\text{tabel}}$ or value *p-value* $> \alpha$ at a significance level of 5%, so accept H_0. This means that the data is normally distributed.
   c. Multicollinearity test is conducted to test whether in a model there is a correlation between independent variables or between dependent variables. [11]. The assessment of multicollinearity testing is seen from the VIF value where the hypothesis, test statistics and decision criteria are as follows:
   H_0 : There is multicollinearity between variables
   H_1 : There is no multicollinearity between variables Staistical Test:

$$VIF_i = \frac{1}{1 - R_i^2}, \quad i = 1, 2, \ldots, k \tag{2}$$

   with:
   VIF : Variance Inflation Factor value
   R_i^2 : The coefficient of determination of the i independent variable with other variables
   The decision criteria is if the value of $VIF < 10$ so that H_0 is rejected. This means that there is no multicollinearity between the independent variables and the dependent variables.

3. **Determining Canonical Coefficients**
   Suppose there are two random variables, namely random variable X consisting of the set of variables $X_1, X_2, \ldots, X_q$, and random variable Y consisting of the set of variables $Y_1, Y_2, \ldots, Y_p$. Canonical correlation is based on linear indices, namely linear combinations as shown in the following equation:

$$a^T X \quad \text{and} \quad b^T Y \tag{3}$$

   Canonical correlation analysis is used to find the values of the vectors a and b that maximize the correlation $\rho(a, b)$ between the two linear combinations, namely:

$$\rho(a, b) = \rho(a^T X, b^T Y) \tag{4}$$

   The correlation in equation (4) between these two projections is further defined as follows. Suppose,

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \begin{pmatrix} \mu \\ \nu \end{pmatrix}, \quad \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix} \tag{5}$$

where the submatrix of this covariance structure is given as:

- $\text{Var}(X) = \Sigma_{XX}$ berukuran $q \times q$
- $\text{Var}(Y) = \Sigma_{YY}$ size $p \times p$
- $\text{Cov}(X,Y) = \mathbb{E}[(X-\mu)(Y-\nu)^T] = \Sigma_{XY} = \Sigma_{YX}^T$ size $q \times p$

The correlation between two variables X and Y can be defined from the covariance as:

$$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \tag{6}$$

The eigenvalues are obtained from the characteristic equation, so that the coefficient vectors a_k and b_k obtained in the canonical functions $a_k^T X$ and $b_k^T Y$ are the eigenvectors of the same two matrices.

4. **Conduct significance tests**

(a) **Overall significance test**

This test aims to test whether there is a significant relationship between the set of variables X and the set of variables Y as a whole. The hypotheses, test statistics, and decision criteria are as follows:

H_0: There is no significant relationship between variables X and Y
H_1: There is at least one significant relationship between variables X and Y

Statistical test:

$$F = \frac{(1-\Lambda^{1/t})}{\Lambda^{1/t}} \cdot \frac{db_2}{db_1} \tag{7}$$

with:

Λ	$= \prod_{i=1}^{k}(1-\rho_i^2)$	
db_1	$= pq$	
db_2	$= \left[n - \frac{1}{2}(p+q+3)\right]t - \frac{1}{2}pq + 1$	
t	$= \sqrt{\frac{p^2q^2-4}{p^2q^2-5}}$	
n	:	Number of observations
p	:	Number of variables in the set Y
q	:	The number of variables in the set X
k	$= \min(p,q)$	

Decision Criteria is if $F > F_{\alpha;db_1;db_2}$ or p-value $< \alpha$ with a significance level 5% ($\alpha = 0{,}05$), so reject H_0. This means that there is a significant relationship between the set of variables X and Y as a whole.

(b) **Partial significance test**

A partial significance test is carried out when at least one canonical correlation is found to be significant in the overall canonical correlation analysis [12]. This test is then applied sequentially to the second and subsequent canonical correlations up to the k^{th}. The corresponding hypotheses, test statistics, and decision criteria are outlined as follows:

H_0: There is no significant k canonical correlation
H_1: There is a significant k canonical correlation

Statistical test:

$$F = \frac{(1-\Lambda^{1/t})}{\Lambda^{1/t}} \cdot \frac{db_2}{db_1} \tag{8}$$

with:

$$\Lambda = \prod_{i=1}^{k}(1-\rho_i^2)$$

$$db_2 = \left[n - \frac{1}{2}(p+q+3)\right] t - \frac{1}{2}(p-r+1)(q-r+1)+1$$

$$t = \sqrt{\frac{(p-r+1)^2(q-r+1)^2-4}{(p-r+1)^2(q-r+1)^2-5}}$$

n : number of observations
p : many sets of variables Y
q : many sets of variables X
k : $\min(p,q)$

Decision criteria is if $F > F_{\alpha;db_1;db_2}$ or p-value $< \alpha$ with a significance level 5% ($\alpha = 0{,}05$), so reject H_0. This means that there is a significant k canonical correlation.

5. **Performing Canonical Correlation Interpretation**
   (a) Canonical weights describe the contribution of each original variable in constructing the canonical variate that has the strongest association with the canonical variate from the opposite set of variables [13]. Canonical weights indicate how much an original variable contributes to the formation of the canonical variate [14].
   (b) Canonical loadings represent the extent to which each original variable contributes to the constructed canonical variate, and are conceptually similar to factor loadings in factor analysis [1].
   (c) Cross-loadings indicate the correlation between an original variable and the canonical variate from the opposite set. These can be obtained by multiplying the canonical correlation coefficient by the corresponding canonical loading [15].
6. **Performing Redundancy Analysis**
   Redundancy analysis provides a measure of the proportion of variance in the dependent and independent variables that can be explained through their correlation with the corresponding canonical variates [16].

3 Results and Discussion

The data used in this study were obtained from the official website of the Jambi Province Statistics Office (Badan Pusat Statistik Provinsi Jambi) for the period 2010 to 2023. The dataset covers 9 regencies and 2 cities within the province. The variables analyzed include population growth rate (Y_1), life expectancy (Y_2), economic growth rate (X_1), labor force participation rate (X_2), percentage of poor population (X_3), and population density (X_4). These variables were examined using canonical correlation analysis to construct a model that comprehensively captures the relationships between the two sets of variables. The results of the canonical correlation modeling, from the initial to the final stages, are presented as follows:

3.1 Test Assumptions

Before conducting the canonical correlation analysis, the data underwent a series of assumption tests to ensure its suitability for analysis. The first test performed was the normality test using the Kolmogorov-Smirnov Test, as presented in Table 1:

Table 1: Normality Test

Variable	$p-value$
Population Growth Rate	0,832
Life Expectancy	0,999
Economic Growth Rate	0,599
Labor Force Participation Rate	0,821
Percentage of Poor Population	0,841
Population density	0,995

Table 1 shows that at a significance level of $\alpha = 0.05$, all p-values are greater than 0.05, leading to a failure to reject H_0. This indicates that the data distribution for all variables does not significantly deviate from normality.

The next assumption test is the multicollinearity test, the results of which are presented in Table 2 and Table 3 below:

Table 2: Multicollinearity Test of set Y

Variable Y	VIF
Population Growth Rate	2,866
Life Expectancy	2,866

Table 2 shows that the VIF values for the dependent variables are all less than 10, which leads to the rejection of H_0. This indicates that the two variables in the Y set do not exhibit multicollinearity, meaning they are not highly linearly correlated with each other.

Table 3: Multicollinearity Test of Variables X

Variable X	VIF
Economic Growth Rate	1,517
Tingkat Partisipasi Angkatan Kerja	1,95
Percentage of Poor Population	1,441
Population density	1,473

Table 3 shows that the VIF values for the independent variables are all below 10, which leads to the rejection of H_0. This suggests that the four variables in the X set do not exhibit multicollinearity, meaning they are not highly linearly correlated with one another.

3.2 Significance Test

After all assumption tests were satisfied, the next step was to perform both overall and partial significance tests. The results of the overall significance test are presented in Table 4 below:

Table 4: Overall Significance Test

Test	$p-value$
$Pillais$	0,050
$Hotelling$	0,00
$Wilks$	0,003

Based on Table 4, all three multivariate tests, Pillai's Trace, Hotelling's Trace, and Wilks' Lambda, indicate statistically significant results, as all significance values are less than 0.05. Therefore, H_0 is rejected, meaning that the sets of independent and dependent variables are simultaneously correlated. This provides strong evidence that, overall, there is a significant relationship between the combination of variables in set X and the combination in set Y within the constructed canonical model.

The results of the partial significance tests are presented in Table 5 below:

Table 5: Partial Significance Test

Canonical Functions	Canonical Coefficients	F-statistics	p-value
1	0,948	4,886	0,003
2	0,394	0,552	0,659

Based on the results of the partial significance tests for each canonical function, it was found that only the first canonical function shows a significant relationship between the two variable sets, namely the socioeconomic indicators (X) and the population quality of life indicators (Y). The first function has a canonical correlation value of 0.9487, which is considered very high, with an F-statistic of 4.886 and a significance value (p-value) of 0.003. Since the p-value is less than 0.05, H_0 is rejected. This indicates a strong and statistically significant linear relationship between the combination of X_1˘X_4 and Y_1˘Y_2.

The strength of the first canonical correlation indicates that approximately 90% ($0.9487^2 \approx 0.9$) of the shared variance between the two sets of variables can be explained by the first canonical function. This suggests that, in the first canonical function, the combination of socioeconomic variables such as economic growth rate (X_1), labor force participation rate (X_2), percentage of poor population (X_3), and population density (X_4) is strongly associated with quality of life, as represented by population growth rate (Y_1) and life expectancy (Y_2).

The second canonical function shows a much weaker correlation of 0.3944, with an F-statistic of 0.552 and a significance value of 0.659, which is far above the 5% significance threshold. Therefore, H_0 is accepted, indicating that there is no statistically significant partial relationship between the X and Y variable sets in the second canonical function. As a result, interpretation and discussion in this study are focused solely on the first canonical function.

3.3 Interpretation of Canonical Functions

Once it is known that the pair of canonical functions formed shows partial significance, the next step is to interpret the canonical function, namely with the following explanation:

Canonical Weight

The calculation of the canonical weights is divided into 2, namely the calculation for the variable set X and the variable set Y. The results of the canonical weight calculation are as follows.

Table 6: Canonical Weight of Dependent Variable in Canonical Function 1

Variable	**Weight Kanonik**
Y_1	0,0558
Y_2	-0,9543

Table 6 shows that the large canonical weight for Y_2, which is -0.9543, indicates that the first canonical function is heavily influenced by the life expectancy variable. This means that changes in the canonical function are more reflective of variations in life expectancy than in population growth rate.

Based on the canonical weights of the Y set, the canonical function model for U_1 can be constructed as follows:

$$U_1 = 0,0558Y_1 - 0,9543Y_2 \tag{9}$$

The negative weight on life expectancy (-0.9543) indicates the opposite direction of the relationship with the first canonical function. If the canonical function increases, then life expectancy tends to decrease, or vice versa.

Then the result of the canonical weight calculation for the set X is:

Table 7: Canonical Weights of Independent Variables in Canonical Function 1

Variable	Weight Kanonik
X_1	0,347
X_2	-0,105
X_3	0,221
X_4	-0,614

Table 7 indicates that the highest canonical weight is observed for X_4, with a value of -0.614, suggesting that the first canonical function is strongly influenced by the population density variable. This implies that changes in the canonical function are more reflective of variations in population density than in the other variables within the X set.

Based on the canonical weights of the X set, the canonical function model for V_1 can be constructed as follows:

$$V_1 = 0,347X_1 - 0,105X_2 + 0,221X_3 - 0,614X_4 \tag{10}$$

The negative weights for labor force participation rate and population density indicate an inverse relationship with the first canonical function. This means that as the canonical function increases, life expectancy tends to decrease, or vice versa. On the other hand, the positive weights for economic growth rate and percentage of poor population suggest that the canonical function increases along with higher economic growth and a greater proportion of the poor population.

Canonical loading

The calculation of the canonical load is divided into 2, namely for the variable set X and the variable set Y. The results of the canonical load calculation are as follows:

Table 8: Canonical Load of Dependent Variable on Canonical Function 1

Variable	Canonical Loading
Y_1	0,826
Y_2	-0,999

Table 8 presents the canonical loadings, which represent the correlations between the original variables and their respective canonical functions. The canonical loading for life expectancy is notably high, indicating that the first canonical function captures nearly all of the variance in life expectancy. This suggests that Y_2 is the most dominant structural contributor to the canonical variate.

Furthermore, the results of the canonical weights calculation for the X set are as follows:

Table 9: Canonical Load of Independent Variables on Canonical Function 1

Variable	Canonical Loading
X_1	0,677
X_2	-0,732
X_3	0,681
X_4	-0,872

Table 9 shows that the canonical loading for Population Density (-0.872) is the largest among the variables, indicating that the first canonical function from the X set strongly represents the variation in population density. The negative sign implies that an increase in the canonical function is associated with a decrease in both population density and labor force participation rate, whereas the positive signs indicate that the canonical function increases along with higher economic growth rate and a greater percentage of poor population.

Canonical Cross Loading

The calculation of the canonical cross-load is divided into 2, namely the cross-load of the variable set X against the set Y and the cross-load of the variable set Y against the set X. The results of the canonical day load calculation are as follows:

Table 10: Canonical Cross-Loading of Set Y on Set X

Variable	**Canonical Cross Loading**
Y_1	0,783
Y_2	-0,948

Table 10 illustrates how the Y variables—population growth rate and life expectancy—are correlated with the canonical function derived from the X set (a combination of economic growth rate, labor force participation rate, and percentage of poor population).

Life expectancy shows a very strong correlation with the canonical function of X, indicating that this quality of life indicator is the most influenced by socioeconomic conditions. This high value reinforces life expectancy as the primary variable within the Y set.

Population growth rate also demonstrates a strong correlation with the X function, suggesting that it is fairly sensitive to variations in socioeconomic variables as well.

Table 11: Canonical Cross-Load of Set X on Set Y

Variable	**Canonical Cross Loading**
X_1	0,642
X_2	-0,694
X_3	0,646
X_4	-0,827

Table 11 illustrates how the X variables—economic growth rate, labor force participation rate, and percentage of poor population—correlate with the canonical function derived from the Y set (combining population growth rate and life expectancy).

Population density shows a very strong correlation with the canonical function of Y at –0.827, indicating that this socioeconomic indicator is most influenced by the variables in the Y set. This high correlation confirms population density as the primary variable in the X set. Other variables such as economic growth rate, labor force participation rate, and percentage of poor population also exhibit strong correlations with the Y function, suggesting that they are quite sensitive to changes in the social and quality of life indicators.

3.4 Redundancy Analysis

The redundancy index shows how much of the variance of the set Y can be explained by the canonical function of the set X and vice versa. The calculation results for the redundancy analysis are as follows:

Table 12: Redundancy Analysis

Variable	Redundancy Index
Set Y	0,797
Set X	0,526

Table 12 shows that the redundancy value for the Y variable set is 0.797. This result indicates that approximately 79.7% of the variation in the Y variable set can be explained by the canonical functions derived from the X variables. This demonstrates that the information contained in the socioeconomic variables (X) is highly relevant and strongly explains the variation in the Y set.

The redundancy value for the X variable set is 0.526, indicating that about 52.6% of the combined variance of the X variables can be explained by the canonical functions formed from the Y variables. This confirms a balanced and significant bidirectional relationship. The conditions of population growth rate and life expectancy are not only influenced by socioeconomic factors but also exert feedback effects on the socioeconomic structure. This reinforces that the relationship analyzed here is not unidirectional causality but rather a reciprocal association, consistent with the canonical correlation approach.

4 Conclusion

This study examines the relationship between population growth rate and life expectancy with socioeconomic factors using a canonical correlation approach. The analysis results indicate a strong and statistically significant relationship between the two sets of variables. Furthermore, the findings reveal that the relationship between these groups is not unidirectional but reciprocal. In other words, socioeconomic factors not only influence population growth rate and life expectancy but these demographic indicators also exert feedback effects on socioeconomic conditions. This confirms that the canonical correlation approach effectively captures the complex and simultaneous bidirectional dynamics.

The application of canonical correlation as an alternative method for analyzing multidimensional relationships has been relatively underutilized in regional development studies. Nonetheless, this study has limitations, particularly regarding the limited number of observations and the scope of data representing only a single time period. Therefore, future research is recommended to employ panel data and consider more sophisticated structural approaches, such as $Structural Equation Modeling (SEM)$, to explore causal relationships among variables more comprehensively. This study is expected to provide a conceptual foundation for policymakers and researchers in formulating more integrative development strategies.

Contributor Roles Taxonomy (CRediT)

All authors have contributed significantly to the work reported in this manuscript, in accordance with the Contributor Roles Taxonomy (CRediT). **M. Rizky Ramadhan** contributed to the conceptualization of the study (*Conceptualization*), the development of the research methodology (*Methodology*), and the preparation of the original manuscript draft (*Writing–Original Draft Preparation*). **Gusmi Kholijah** was responsible for organizing and curating the data (*Data Curation*), conducting formal data analysis (*Formal Analysis*), and reviewing and editing the manuscript (*Writing–Review & Editing*).

Declaration of Generative AI and AI-assisted Technologies

No generative AI or AI-assisted technologies were used during the preparation of this manuscript.

Declaration of Competing Interest

The authors declare no competing interests.

Funding and Acknowledgments

This research received no external funding. The authors would like to thank the Department of Mathematics, University of Jambi, for providing access to computing facilities and administrative support throughout the course of this research.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request and subject to confidentiality agreements due to ethical and legal considerations.

References

[1] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis*, 7th ed. Pearson Education, 2010.

[2] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 6th ed. Pearson Prentice Hall, 2007.

[3] United Nations Development Programme (UNDP), *Human Development Report 2010: The Real Wealth of Nations – Pathways to Human Development*. New York: Palgrave Macmillan, 2010.

[4] D. E. Shodiq, "Analisis proporsi populasi penduduk terhadap rumah tangga dengan hunian layak dan terjangkau," *Jurnal Sains Edukatika Indonesia*, vol. 4, no. 01, pp. 23–28, 2022.

[5] B. P. S. P. Jambi, *Provinsi Jambi dalam Angka 2023*. Badan Pusat Statistik, 2023.

[6] S. K. Bandung, *Analisis pencapaian sasaran 2 tahun 2019: Meningkatnya derajat kesehatan masyarakat indikator*, Laporan Pemerintah Daerah Kota Bandung, 2019.

[7] J. Purwadi, B. Gumelar, T. Widiantoro, and Z. N. Ningsih, "Canonical correlation analysis of economic growth and unemployment rate," *BAREKENG: Journal of Mathematics and Its Applications*, vol. 18, no. 2, pp. 1273–1282, 2024.

[8] D. Safitri and P. Indrasari, "Analisis korelasi kanonik pada perilaku kesehatan dan karakteristik sosial ekonomi di kota pati jawa tengah," *Media Statistika*, vol. 2, no. 1, pp. 39–48, 2009.

[9] M. Djazari, D. Rahmawati, and M. A. Nugraha, "Pengaruh sikap menghindari risiko sharing dan knowledge self-efficacy terhadap informal knowledge sharing pada mahasiswa fise uny," *Nominal: Barometer Riset Akuntansi dan Manajemen*, vol. 2, no. 2, pp. 181–209, 2013.

[10] A. Quraisy, "Normalitas data menggunakan uji kolmogorov-smirnov dan shapiro-wilk," *J-HEST: Journal of Health, Education, Economics, Science and Technology*, vol. 3, no. 1, pp. 7–11, 2022.

[11] S. A. Cassany, "Pengaruh kelompok referensi dan gaya hidup terhadap keputusan menggunakan bahan bakar pertamax di kota banda aceh," *Jurnal Bisnis Net*, vol. 2, no. 3, pp. 11–23, 2019.

[12] A. C. Rencher, *Methods of Multivariate Analysis*. Canada: John Wiley & Sons, Inc., 2002.

[13] W. L. Hays, "Review of using multivariate statistics," *Contemporary Psychology*, vol. 28, no. 8, 1983.

[14] R. D. Sari, H. Yozza, and I. R. HG, "Analisis korelasi kanonik untuk menganalisis hubungan karakteristik ibu dengan tumbuh kembang balita (studi kasus di puskesmas nanggalo, padang)," *Jurnal Matematika UNAND*, vol. 4, no. 4, p. 107, 2019.

[15] F. Prastio, S. Martha, and H. Perdana, "Pengujian korelasi kanonik penumpang pesawat di bandar udara internasional supadio," *Buletin Ilmiah Math, Stat, dan Terapannya (Bimaster)*, vol. 8, no. 2, pp. 185–192, 2019.

[16] M. F. Asbah, Sudarno, and D. Safitri, "Penentuan koefisien korelasi kanonik dan interpretasi fungsi kanonik multivariat," *Jurnal Gaussian*, vol. 2, no. April, pp. 119–128, 2013.