turnitin fmipa

ARUM CAUCHY 15

Publikasi Mahasiswa

Document Details

Submission ID
trn:oid:::3618:148074001

Submission Date
Aug 12, 2026, 8:07 PM GMT+7

Download Date
Aug 12, 2026, 8:10 PM GMT+7

File Name
ARUM CAUCHY 15.pdf

File Size
546.3 KB

15 Pages

6,595 Words

39,211 Characters

20% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Small Matches (less than 8 words)

Match Groups

102 Not Cited or Quoted 19%
Matches with neither in-text citation nor quotation marks

3 Missing Quotations 1%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

15% Internet sources
11% Publications
15% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **102** Not Cited or Quoted 19%
  Matches with neither in-text citation nor quotation marks
- **3** Missing Quotations 1%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

15% Internet sources
11% Publications
15% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	ejournal.uin-malang.ac.id <1%
2	Internet	www.mdpi.com <1%
3	Student papers	University of Huddersfield on 2021-12-14 <1%
4	Student papers	Mahidol University on 2026-01-15 <1%
5	Internet	proceedings.stis.ac.id <1%
6	Internet	ideas.repec.org <1%
7	Student papers	University of Bradford on 2023-05-29 <1%
8	Student papers	University of Hull on 2018-10-09 <1%
9	Internet	ejournal.iainpalopo.ac.id <1%
10	Internet	jurnal.syntaxliterate.co.id <1%

11	Student papers	
Kingston University on 2012-04-28		<1%
12	Student papers	
aimsr on 2026-07-17		<1%
13	Internet	
ir.uew.edu.gh:8080		<1%
14	Internet	
ojs.uph.edu		<1%
15	Internet	
proceedings.utwente.nl		<1%
16	Internet	
conferences.uin-malang.ac.id		<1%
17	Internet	
learning-gate.com		<1%
18	Internet	
www.coursehero.com		<1%
19	Student papers	
Universiti Malaysia Sarawak on 2016-09-30		<1%
20	Internet	
e-journal.unair.ac.id		<1%
21	Internet	
network.bepress.com		<1%
22	Internet	
scholarhub.unimas.my		<1%
23	Student papers	
Northcentral on 2012-05-07		<1%
24	Publication	
Pallavi R. Kamath, Yogesh P. Pai, Nandan K.P. Prabhu. "Building customer loyalty ...		<1%

25 Internet

www.nature.com <1%

26 Student papers

Binus University International on 2026-06-06 <1%

27 Publication

Retno Danu Rusmawati, Wawan Gunawan, Irwan Alnarus Kautsar, Zulidyana D... <1%

28 Student papers

Universiti Selangor on 2021-12-12 <1%

29 Student papers

University of Nottingham on 2021-12-22 <1%

30 Internet

www.atlantis-press.com <1%

31 Student papers

Universiti Teknologi MARA on 2026-07-04 <1%

32 Internet

ijsoc.goacademica.com <1%

33 Internet

jaa.kbsaa.org <1%

34 Student papers

Brunel University on 2012-06-28 <1%

35 Student papers

National Kaohsiung University of Hospitality and Tourism on 2022-06-20 <1%

36 Student papers

University of Newcastle on 2021-07-01 <1%

37 Publication

Xiu Ming Loh, Voon Hsien Lee, Lai Ying Leong. "Understanding the opposing forc... <1%

38 Internet

www.e-bcrp.org <1%

39 Internet
www.igi-global.com <1%

40 Publication
G. Aris Wulantoro, Tuty Sariwulan, Hania Aminah. "Environmental Literacy as a ... <1%

41 Student papers
LPPM on 2025-07-31 <1%

42 Publication
Mancha, Ruben, Mark T. Leung, Jan Clark, and Minghe Sun. "Finite mixture partial... <1%

43 Publication
Muhammad Fauzan, Hendra Perdana, Neva Satyahadewi. "STRUCTURAL EQUATI... <1%

44 Student papers
University of Gloucestershire on 2014-02-04 <1%

45 Internet
pmc.ncbi.nlm.nih.gov <1%

46 Student papers
College of Public Health on 2026-01-19 <1%

47 Publication
Jörg Henseler. "Guest editorial", Industrial Management & Data Systems, 2016 <1%

48 Student papers
Universiti Malaysia Perlis on 2016-07-27 <1%

49 Internet
listens.online <1%

50 Internet
smartpls.com <1%

51 Internet
1library.net <1%

52 Student papers
City University Malaysia on 2025-09-24 <1%

53	Publication	Rattanasiriwilai, Tachanun. "Adoption of Smart Technologies for Transport Mana... <1%
54	Publication	Smruti Bulsari, Kiran Pandya. "Quantitative Research Using R", Routledge, 2026 <1%
55	Student papers	UCSI Graduate Business School on 2025-08-26 <1%
56	Student papers	Universidad de Córdoba on 2025-08-01 <1%
57	Publication	Viet Thai Nguyen, Quang Vinh Nguyen, Thị Ha Truong. "When Information Qualit... <1%
58	Internet	core.ac.uk <1%
59	Internet	dokumen.pub <1%
60	Internet	journal.yrpipku.com <1%
61	Student papers	Binus University International on 2026-02-21 <1%
62	Student papers	Fakultet organizacije i informatike-Faculty of Organization and Informatics <1%
63	Publication	Hair, Joseph F., Marko Sarstedt, Torsten M. Pieper, and Christian M. Ringle. "The U... <1%
64	Publication	Paulina Pannen, Ojat Darojat, Moch Abduh. "Education Technology in the New No... <1%
65	Student papers	Taylor's Education Group on 2024-12-08 <1%
66	Student papers	Universitaet Hamburg on 2010-10-11 <1%

67 Student papers

Universiti Islam Antarabangsa Sultan Abdul Halim Mu'adzam Shah (UniSHAMS) o... <1%

68 Internet

ejournal.abcollab.id <1%

69 Internet

fair-msc-project.eu <1%

70 Internet

invotek.ppj.unp.ac.id <1%

71 Internet

jurnal2.untagsmg.ac.id <1%

72 Internet

kclpure.kcl.ac.uk <1%

73 Internet

openscholar.dut.ac.za <1%

74 Internet

publishing.globalcsrc.org <1%

75 Internet

pubs.aip.org <1%

76 Internet

pure.tue.nl <1%

77 Internet

www.aasmr.org <1%

78 Internet

www.hindawi.com <1%

79 Internet

www.ros.hw.ac.uk <1%

80 Internet

www.scielosp.org <1%

turnitin Page 9 of 23 - Integrity Submission Submission ID trn:oid:::3618:148074001

CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 11 (2) (2026), Pages 1-11
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Identifying Regional Heterogeneity and Priority Factors for Infrastructure in Indonesia Using FIMIX-PLS and Importance-Performance Map Analysis

Arum Nanda Putri Aprilia[1] and Rahmawati Erma Standsyah[2*]

[1]*Student, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya, Surabaya, Indonesia*
[2]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya, Surabaya, Indonesia*

Abstract

The development of infrastructure is closely linked to the measures taken to reduce the regional disparities among the various regions. In Indonesia, however, the situation in each province is not the same, whether in terms of economic capacity, the available human resources, or the particular features of each region. This study looks at the influence of Economy, Human Resources, and Regional Characteristics on Infrastructure, using secondary data from 38 Indonesian provinces for the years 2024 and 2025 obtained from Statistics Indonesia (BPS). Partial Least Squares-Structural Equation Modelling (PLS-SEM) is used for the analysis, followed by FIMIX-PLS and Importance-Performance Map Analysis (IPMA). The results from the PLS-SEM indicate that Regional Characteristics has a significant positive effect on Infrastructure (path coefficient = 0.951, t-statistic = 28.136, p-value = 0.000); its effect size is also the largest, with an f^2 value of 6.055, whereas those for Economy and Human Resources are 0.013 and 0.001, respectively and neither has a significant effect. The FIMIX-PLS analysis identified four segments with AIC3 = 11.052, BIC = 23.166, CAIC = 42.166, and EN = 0.865. The structural paths also vary among the four segments, particularly for Regional Characteristics, which range from 0.542 to 1.027. Moreover, the IPMA shows the Top Priority factor across the segments, and Regional Characteristics remains so. These results reveal differences between the groups of provinces and thus provide a basis for setting priorities for infrastructure improvement according to the specific characteristics of each group.

Keywords: FIMIX-PLS; Infrastructure; IPMA; PLS-SEM.

Copyright © 2026 by Authors, Published by CAUCHY Group. This is an open access article under the CC BY-SA License (https://creativecommons.org/licenses/by-sa/4.0)

[*]**Corresponding author**. E-mail: rahmawatistandsyah@unesa.ac.id

Received: m dd, yyyy Revised: mm dd, yyyy Accepted: mm dd, yyyy
DOI: https://doi.org/10.18860/cauchy.v11i1.xxxxx

turnitin Page 9 of 23 - Integrity Submission Submission ID trn:oid:::3618:148074001

1. Introduction

Sustainable development through the Sustainable Development Goals (SDGs) ranks infrastructure as one of the key foundations for stimulating economic growth, improving the well-being of the population, and expanding access to various basic services [1] [2]. The availability of adequate infrastructure is an important factor in strengthening competitiveness, promoting more inclusive development, and enhancing interregional connectivity, in addition to facilitating economic activities [3], [4]. That is why infrastructure development continues to be a strategic priority of many countries, including Indonesia, as part of its efforts to support sustainable development.

Although Indonesia has made infrastructure development a national priority, equitable development continues to face various challenges [5], [6]. Disparities in geographical conditions, regional characteristics, economic capacities, and development levels among provinces result in uneven development of the quality and availability of infrastructure [7]. This situation leads to inequalities in access to essential services, economic activities, and development opportunities, particularly in regions outside of Java [4]. Therefore, infrastructure development requires a region-specific approach to ensure that implemented policies are more effective and promote inclusive and sustainable development [4], [8].

In recent years, the use of finite mixture partial least squares (FIMIX-PLS) has expanded considerably for identifying unobserved heterogeneity across various research fields [9], [10], [11]. FIMIX-PLS allows data to be grouped into segments with different relational structures, so that the resulting model better represents the characteristics of each group. Simultaneously, FIMIX-PLS has begun to be combined with importance-performance mapping analysis (IPMA) to determine improvement priorities based on the characteristics of each segment [12], [13], [14]. This approach has been widely applied in various research areas. Various studies have shown that the combining FIMIX-PLS and IPMA not only reveals hidden heterogeneity but also provides more accurate information for strategy development compared to the use of conventional PLS-SEM.

Research on infrastructure development and underdeveloped regions in Indonesia generally still uses a single model to represent the entire study area [15], [16]. This approach provides a global view of the relationships between variables at the population level but fails to account for unobserved heterogeneity between regions. Indeed, Indonesia is a country with highly diverse geographical, social, and economic conditions, as well as varying levels of vulnerability to disasters, such that the relationships between factors influencing infrastructure development can differ from one region to another [8], [17]. Therefore, this study integrates partial least squares-structural equation modeling (PLS-SEM) [18], finite mixture partial least squares (FIMIX-PLS) [19], and importance-performance map analysis (IPMA) [20] to identify this unobserved heterogeneity and determine infrastructure improvement priorities in each segment.

According to the description, the purpose of this study is to analyze the impact of the economy, human resources, and regional characteristics on infrastructure using partial least squares - structural equation modelling (PLS-SEM). Furthermore, it identifies unobserved heterogeneity between provinces using finite mixture partial least squares (FIMIX-PLS) and determines infrastructure improvement priorities in each segment using importance-performance map analysis (IPMA). This study should result in recommendations for infrastructure development policies tailored to the characteristics of each group of provinces. Therefore, the adopted development policies should be more selective.

2. Methods

This study uses a quantitative method to analyze 38 provinces in Indonesia using Partial Least Squares-Structural Equation Modelling (SEM-PLS). After confirming and validating the validity and reliability criteria, Finite Mixture Partial Least Squares (FIMIX-PLS) analysis was used to detect the existence of unobserved heterogeneity and assign the provinces to segments with homogeneous characteristics. Then, an Importance-Performance Map Analysis (IPMA) was carried out for each segment to determine the priority indicators of infrastructure development in Indonesia.

2.1. Data Source

The research used secondary data from the official publication of the Indonesian Statistical Agency (BPS). The unit of analysis is 38 provinces of Indonesia. The data used were obtained from 2024 and 2025 from several publications of BPS, including Potensi Desa (PODES), Human Development Index (HDI), and other statistical reports, which provide indicators that are related to the latent variable used in this research. These data from different sources provided full information on the situation of the infrastructure, economy, human resources, and regional characteristics of each province.

The study consists of four constructs, namely, one endogenous variable, Infrastructure and three exogenous variables Economy, Human Resources, and Regional Characteristics. Each construct is measured using indicators are defined based on the official BPS publication and previous research. Indicators were chosen such that each construct is a valid measure of the concept it is meant to capture and is coherent with the research objectives. Table 1 details the variables and indicators used in the study.

Table 1: Latent Variables and Measurement Indicators

Latent Variable	Indicator	
Infrastructure	Asphalt / Concrete Road	I1
	Market Without Buildings	I2
	Health Facility	I3
	Basic Education Facilities	I4
Economy	Gross Regional Domestic Product (GRDP) per Capita	E1
	Percentage of the population employed in the formal sector	E2
Human Resources	Life Expectancy	H1
	Average years of schooling	H2
	Expenditure per Capita	H3
Regional Characteristics	Earthquake	R1
	Flood	R2
	Whirlwind / Putting Buyun / Typhoon	R3
	Dryness	R4

2.2. SEM-PLS

Multivariate analysis is a statistical approach that allows for the simultaneous analysis of several variables, thus revealing more complex relationships than univariate or bivariate analyses. Among the most widely used multivariate analysis methods is structural equation modelling (SEM), it examines how a latent construct is measured by several observed indicators [18].

This research uses partial least squares-structural equation modelling (PLS-SEM), to analyze the relationship between economic factors, human resources, regional characteristics, and infrastructure development in Indonesia. This method was chosen because it allows for the simultaneous estimation of measurement and structural models and is well-suited to analyzing complex relationships between latent variables. [18].

2.3.1 Evaluation of the Measurement Model (Outer Model)

The first step in PLS-SEM analysis is to evaluate the measurement model, also known as the outer model. This stage is used to check that each indicator is a good indicator of the latent construct that is being measured. Before examining correlations among constructs, the measurement instrument must fulfill reliability and validity criteria. The measurement model assessment for reflective constructs relies on internal consistency reliability and convergent validity [18].

a. Indicator Reliability

Checking the reliability of indicators is the first step when evaluating a measurement model. This process shows how well each indicator represents the underlying concept being measured. In a reflective measurement model, indicator reliability is assessed using external outer loading, which shows how

strongly the indicator is connected to the latent construct. When the external outer loading is higher, the indicator plays a bigger role in measuring the construct [18].

In general, indicators with external outer loading ≥ 0.70 are satisfactory because they explain a large part of the construct's variance. However, indicators with an outer loading between 0.40 and 0.70 can be preserved if their removal does not improve the overall reliability and validity of the construct. On the other hand, indicators with an outer loading < 0.40 are usually eliminated since they contribute very little to the measurement of the construct. [18].

b. Internal Consistency Reliability

Internal consistency reliability looks at whether the indicators of a construct are measuring the same concept consistently. In the present study, internal consistency reliability was assessed by means of Cronbach's alpha coefficient and composite reliability (CR). Cronbach's alpha coefficient examines the internal consistency of the indicators, whereas composite reliability offers a more accurate estimate by taking into account the contribution of each indicator to the construct. A construct is regarded as reliable when both the Cronbach's alpha coefficient and the composite reliability have values above 0.70 [18].

Composite reliability iscalculated using the following equation:

$$\rho_c = \frac{\left(\sum_{i=1}^{M} l_i\right)^2}{\left(\sum_{i=1}^{M} l\right)^2 + \sum_{i=1}^{M} var(e_i)}$$

In this equation, l_i represents the standardized external saturation value of the i-th indicator, M is the number of indicators measuring the latent variable, and $var(e_i)$ represents the variance of the measurement error, defined as $1 - l_i^2$.

c. Convergent Validity

Convergent validity helps determine how much a reflective construct explains the variance of the indicators measuring it. A construct is considered to exhibit good convergent validity if its indicators consistently represent the same latent construct [18].

In this study, we assessed convergent validity using the extracted mean variance (AVE). The AVE value represents the average amount of variance among the indicators that the latent construct is able to explain, and a construct is regarded as valid when its AVE is greater than 0.50 [18].

The average extracted variance (AVE) is calculated using the following equation:

$$AVE = \frac{\sum_{i=1}^{M} l_i^2}{M}$$

In this equation, l_i represents the standardized external loading value of the i-th indicator, while M M representsthe number of indicators measuring the latent construct.

A construct shows adequate convergent validity when its average variance extracted (AVE) is 0.50 or above, since in that situation the construct explains more than half of the variance found in its measurement indicators. On the other hand, if the AVE is less than 0.50, then most of the variance is due to measurement error rather than being caused by the construct [18].

2.3.2 Evaluation of the Structural Model (Inner Model)

Once you have confirmed that the measurement model is reliable and valid, you can move on to the next step, which is to focus on the structural model, also known as the inner model, to continue the analysis [18].

In this study, we evaluated the structural model using the coefficient of determination (R^2), effect size (f^2), and bootstrap resampling to test the significance of the relationships between the constructs. These three methods helped us see how well the model explains the endogenous constructs, how much

each exogenous construct contributes, and whether the relationships in the research model are significant [18].

a. Coefficient of Determination (R^2)

The coefficient of determination (R^2) shows how well a structural model explains the variance of an endogenous variable using the exogenous variables that affect it. A higher R^2 means the model explains the endogenous variable more effectively [18].

The interpretation of R^2 values, which vary from 0 to 1, depends on the research field and model complexity. In marketing and social sciences, R^2 values of 0.75, 0.50, and 0.25 are regarded as high, moderate, and low, respectively [18]. However, to make sure the model has both strong predictive power and high parsimony, R^2 should be interpreted alongside other structural model evaluation criteria [18].

b. Effect Size (f^2)

The effect size known as f^2 is a measure of the amount that each exogenous variable contributes to an endogenous variable in a structural model; whereas R^2 examines the model's overall predictive power, f^2 is concerned with the strength of the influence of an exogenous variable on a particular endogenous variable [18].

The f^2 value is obtained by comparing the R^2 value of the endogenous variable before and after removing an exogenous variable from the model. This comparison indicates the exogenous variable's contribution to explaining the variance of the endogenous variable. The f^2 value is calculated using the following equation [18].

$$f^2 = \frac{R^2_{include} - R^2_{excluded}}{1 - R^2_{included}}$$

In this equation, $R^2_{included}$ represents the coefficient of determination when the exogenous variable is included in the structural model, while $R^2_{excluded}$ represents the coefficient of determination after its exclusion from the model. The difference between these two values indicates the significance of the exogenous variable contribution to the predictive power of the endogenous variable [18].

The F^2 value allows us to assess the relative contribution of an exogenous variable to explaining the endogenous variable. F^2 values of 0.02, 0.15, and 0.35 indicate small, medium, and large effects, respectively, while values below 0.02 indicate a negligible contribution from the exogenous variable [18].

c. Bootstrapping

The importance of the relationships among the structural model constructs was assessed using a bootstrap resampling procedure. This technique allowed for the estimation of standard errors and the testing of the statistical significance of the relationships between the model constructs [18].

A bootstrap resampling procedure was carried out using 5,000 samples in accordance with the guidelines for PLS-SEM analysis. Standard error values, t-statistics, and p-values were obtained from the bootstrap resampling and then used to evaluate the significance of the relationship concerning the latent variable. A two-tailed test was performed at the 5% level of significance. The relationship between the variables was considered to be significant if the t-statistic was greater than 1.96 or the p-value was less than 0.05; if neither of these conditions was met, the relationship was regarded as not significant [18].

2.3. FIMIX-PLS

The FIMIX-PLS (Finite Mixture-Partial Least Squares) method is a segmentation technique used to detect unobserved heterogeneity in PLS-SEM models. Unlike conventional PLS-SEM, which assumes all observations are assumed to originate from a single, homogeneous population, FIMIX-PLS allows for the consideration of multiple latent segments with distinct structural relationship characteristics. Thus, this method can identify groups that are not directly observable but exhibit distinct relationship profiles between variables [19].

FIMIX-PLS combines a finite mixture modelling approach with PLS-SEM to identify latent segments based on the membership probabilities of each observation. Each observation has a probability of belonging to each segment and is then classified into the segment with the highest probability. This approach analyzes the differences in structural relationships between segments, resulting in a more representative model estimation than with a single global model [19].

The optimal number of segments is determined by comparing several information criteria, namely the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), the Corrected Akaike Information Criterion (CAIC), and entropy (EN). Lower AIC, BIC, and CAIC values indicate a better segmentation model, while EN values closer to 1 indicate better segment classification. The optimal number of segments is selected based on a combination of these criteria to obtain a model that best fits the data characteristics [18].

In this study, the FIMIX-PLS method was used to identify potential heterogeneity between Indonesian provinces. The resulting segmentation then served as the basis for implementing an importance-performance map analysis (IPMA) on each segment.

2.4. IPMA

Importance-Performance Map analysis (IPMA) is an evolution of PLS-SEM analysis that provides additional information on improvement priorities for a construct or indicator. Unlike PLS-SEM results, which only show the magnitude of influence between constructs, IPMA combines two dimensions: importance and performance. It thus helps identify factors with a strong influence but a relatively low level of performance. IPMA can therefore serve as a basis for determining improvement priorities for constructs and indicators contributing to the target construct [20].

In IPMA analysis, the importance value is calculated from the total effects generated by the PLS-SEM model. This value indicates the magnitude of a construct's contribution to the target construct. Concurrently, the performance value is derived from the latent variable score and converted to a 0 to 100 scale to facilitate interpretation. Integrating these two measures enables visualization of each construct's position based on its importance and performance levels [20].

In this study, importance-performance mapping (IPMA) analysis was applied separately to each segment generated by the FIMIX-PLS analysis. This approach allows for the evaluation of the importance and performance of each variable based on the characteristics of each segment. Therefore, the resulting improvement priorities are not based solely on overall data but also takeinto account the differences in characteristics between segments, thus enabling the formulation of more specific and targeted recommendations.

3. Results and Discussion

The research results are presented in stages, including the evaluation of the PLS-SEM, FIMIX-PLS analysis, and importance-performance map (IPMA) analysis. The discussion at each stage is conducted in a way that explains the main conclusions of this study.

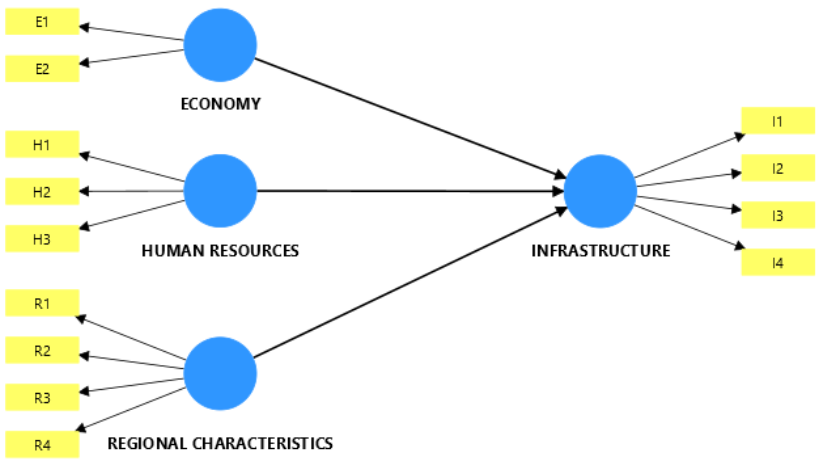


Fig. 1: Conceptual Research Model.

Fig. 1 presents the research model used in the PLS-SEM analysis. This model illustrates the relationships between the economy, human resources, and regional characteristics—acting as exogenous variables and infrastructure, acting as an endogenous variable. All constructs are measured using reflective indicators, in accordance with the conceptual model developed during the research design phase.

3.1. Evaluation of DBSCAN Parameters and Cluster Formation

The measurement model was evaluated to ensure that each construct met the validity and reliability criteria before testing the structural model. In this study, the measurement model was evaluated using external factor loading, internal consistency, convergent validity, and discriminant validity tests. The results of this evaluation are presented in Table 2.

Table 2: Measurement Model Evaluation Results

Latent Variable	Indicator	Outer Loading	Cronbach's alpha	CR	AVE
Infrastructure	I1	0,966	0,975	0,982	0,930
	I2	0,926			
	I3	0,983			
	I4	0,982			
Economy	E1	0,999	0,730	0,808	0,689
	E2	0,617			
	H1	0,993			
	H2	0,537			
	H3	0,827			
Regional Characteristics	R1	0,776	0,858	0,904	0,702
	R2	0,893			
	R3	0,840			
	R4	0,839			

According to Table 2, all explanatory variables meet the recommended validity and reliability criteria for structural equation modelling (PLS-SEM). The external factor loading of each indicator is above the minimum threshold of 0.40. Most indicators have values above 0.70, indicating that they accurately represent the latent variable. Although indicators E2 and H2 have external factor loadings of 0.617 and 0.537, respectively, they are retained because they remain above the recommended minimum threshold and do not degrade the overall quality of the model [18].

Furthermore, reliability tests showed that all constructs had Cronbach's alpha values greater than 0.70, and that composite reliability values also exceeded 0.70. These results indicate that each construct has good internal consistency, so that the indicators used allow for consistent measurement of the same construct [18].

Convergent validity was also established because all explanatory variables have an extracted mean variance (AVE) value greater than 0.50. This indicates that each explanatory variable is able to explain more than half of the variance of its constituent indicators; therefore, the indicators used have an adequate level of convergence to represent the latent variable [18].

The evaluation results indicate that the measurement model meets all the required criteria. Therefore, the model is deemed suitable to proceed to the structural model evaluation stage in order to test the relationships between the constructs in this study.

3.2. Evaluasi Model Struktural

Structural model analysis aims to assess its ability to explain relationships among latent variables and to test the significance of established relationships. In this study, these analyzes are based on the analysis of the coefficient of determination (R^2), the effect size (f^2), and on the path significance test using the bootstrap method.

3.2.1. *Coefficient of Determination (R^2)*

Table 3: Coefficient of Determination (R^2) Results

Construct	R^2
Infrastructure	0.894

According to Table 3, the coefficient of determination (R^2) for the variable Infrastructure is 0.894. These results indicate that 89.4% of the variation in this variable can be explained simultaneously by the variables Economy, Human Resources, and Regional Characteristics. The remaining 10.6% is influenced by factors external to the research model.

Hair et al. (2017) state that an R^2 of 0.894 is substantial. This means the model explains the variation in the Infrastructure variable very well.

3.2.2. *Effect Size (f^2)*

Table 4: Effect Size (f^2) of the Structural Model

Structural Path	f^2
Economy → Infrastructure	0.013
Human Resources → Infrastructure	0.001
Regional Characteristics → Infrastructure	6.055

As presented in Table 4, the results of the effect size test show that the variable Regional Characteristics has an F^2 value of 6.055, indicating a significant influence. The variables Economy and Human Resources have F^2 values of 0.013 and 0.001, respectively. Since these values are below the threshold of 0.02, the contribution of these two variables to the variations in the coefficient of determination (R^2) of infrastructure is relatively small.

In summary, these results indicate that the variable Regional Characteristics contributes most to explaining the variation in infrastructure, compared to the other variables.

3.2.3. *Bootsrapping*

Table 5: Bootstrapping Results

Construct	Original sample (O)	T statistics	P values	Description
Economy → Infrastructure	0,043	0,657	0,511	Not significant
Human Resources → Infrastructure	0,014	0,213	0,831	Not significant
Regional Characteristics → Infrastructure	0,951	28,136	0,000	Significant

As shown in Table 5, the results of the bootstrap resampling method indicate that only the relationship between regional characteristics and infrastructure has a significant effect (path coefficient of 0.951, t-statistic of 28.136, and ($p < 0.001$). In contrast, the relationships between the economy and infrastructure ($\beta = 0.043; t = 0.657; p = 0.511$) and between human resources and infrastructure ($\beta = 0.014; t = 0.213; p = 0.831$do not show a significant effect ($p > 0.05$)

Although the bootstrap resampling method allowed for testing the relationships between variables, these results reflect the conditions observed at the overall population level. Given the diversity of regional characteristics across Indonesian provinces, the existence of unobserved heterogeneity warrants further investigation. Therefore, the analysis was continued using the finite mixture-partial least squares (FIMIX-PLS) method to identify this heterogeneity and to group the provinces into segments with similar characteristics.

3.3. Finite Mixture Partial Least Squares (FIMIX-PLS)

A finite mixture partial least squares (FIMIX-PLS) analysis was performed to identify unobserved heterogeneity in the research data. The presence of heterogeneity indicates that the relationships between latent variables are not always homogeneous across all observations, such that the aggregate model estimation may mask certain group characteristics. Therefore, FIMIX-PLS is applied to detect segments with similar patterns of relationships among constructs.

Evaluation of the FIMIX-PLS model using different selection criteria: modified Akaike Information Criterion 3 (AIC3), Bayesian Information Criterion (BIC), corrected Akaike Information Criterion 3 (CAIC), and entropy (EN). For the model selection of the number of segments in FIMIX-PLS, AIC3, BIC, and CAIC values are compared. The lower the values, the better the model. Also, the quality of segmentation is evaluated by statistical entropy (EN). A higher EN value shows a better ability of the model to assign observations to the right segment.

Table 6: FIMIX-PLS Model Selection Criteria

	K = 2	K = 3	K = 4	K = 5
AIC3	24.130	29.079	11.052	22.298
BIC	29.869	38.005	23.166	37.600
CAIC	38.869	52.005	42.166	61.600
EN	0.701	0.721	0.865	0.815

Based on Table 6, the four-segment model (K = 4) was chosen as the most suitable segmentation solution. AIC3, BIC, and CAIC values of 11,052, 23,166, and 42,166, respectively, lower than those obtained from the models with different numbers of segments, support the decision. The model also satisfied these criteria and displayed an entropy (EN) of 0.865, the highest of all the segmentation options tested. Conversely, the two-, three-, and five-segment models showed higher AIC3, BIC, and CAIC values, as well as lower EN values. These results indicate that these three models did not achieve the same quality of segmentation as the four-segment model. Given these findings, the four-segment model was selected for the subsequent stage of the FIMIX-PLS analysis.

These results indicate heterogeneity that the aggregated PLS-SEM model cannot fully explain. Therefore, further analysis was conducted separately on each segment of the FIMIX-PLS results using Importance-Performance Map (IPMA) analysis to understand the priority factors in each segment better.

The results of the FIMIX-PLS process indicate that all Indonesian provinces are grouped into four segments based on the similarity of the relationships between variables. These groupings serve as the basis for conducting separate IPMA analyses for each segment.

Table 7: Provincial Classification Based on FIMIX-PLS Segmentation

Segment 1	Segment 2	Segment 3	Segment 4
Sumatera Barat	Sumatera Utara	Bengkulu	Sumatera Selatan
Jambi	Riau	Kepulauan Bangka Belitung	Jawa Tengah
Lampung	Kepulauan Riau	Nusa Tenggara Timur	Kalimantan Utara
DKI Jakarta	DI Yogyakarta	Kalimantan Barat	Gorontalo
Jawa Barat	Jawa Timur	Sulawesi Selatan	
Banten	Kalimantan Timur	Sulawesi Barat	
Bali			
Nusa Tenggara Barat			

Kalimantan Tengah
Kalimantan Selatan
Sulawesi Utara
Sulawesi Tengah
Sulawesi Tenggara
Maluku
Maluku Utara
Papua Barat
Papua Barat Daya
Papua
Papua Selatan
Papua Tengah
Papua Pegunungan

Table 7 shows the results of the FIMIX-PLS analysis. It classifies the 38 provinces of Indonesia into four segments with various characteristics. Segment 2 has relatively better average economic and human resource conditions, and Segment 3 has relatively worse average economic and human resource indicators. Segment 4 is characterized by relatively better infrastructure indicators but more pronounced regional characteristics. Most indicators in Segment 1 show moderate average values.

Table 8: Structural Path Coefficients for Each FIMIX-PLS Segment

Construct	Segment 1	Segment 2	Segment 3	Segment 4
Economy → Infrastructure	-0.043	0.148	0.647	-0.237
Human Resources → Infrastructure	0.020	0.044	-0.216	0.503
Regional Characteristics → Infrastructure	0.977	1.026	1.027	0.542

The state of infrastructure in all segments is significantly affected by regional characteristics and path coefficients reveal differences among them. Table 8 shows that Segment 3 has the highest coefficient value (1.027), followed by Segment 2 (1.026), Segment 1 (0.977), and Segment 4 (0.542). Economic factors and human resources have less influence, and the path coefficient values are different between segments. This is an indication that the relationships between variables are not the same for all provinces but are determined by the particularities of each group. Considering these differences, the Importance-Performance Map Analysis (IPMA) method is applied in this analysis to identify the priority factors for infrastructure development for each segment.

3.4. Importance–Performance Map Analysis (IPMA)

An Importance-Performance Map Analysis (IPMA) was performed to evaluate the importance and performance of each variable regarding the Infrastructure variable in the segments defined by FIMIX-PLS. This analysis allows us to identify which variables should have priority in an effort to improve performance, taking into account the specific characteristics of each segment. So, the recommendations that come out are not generic but adapted to the situation of each provincial group. The IPMA results for each segment are presented separately to give a clear picture of the differences in the priorities between the segments.

3.4.1. Analisis IPMA pada Segmen 1

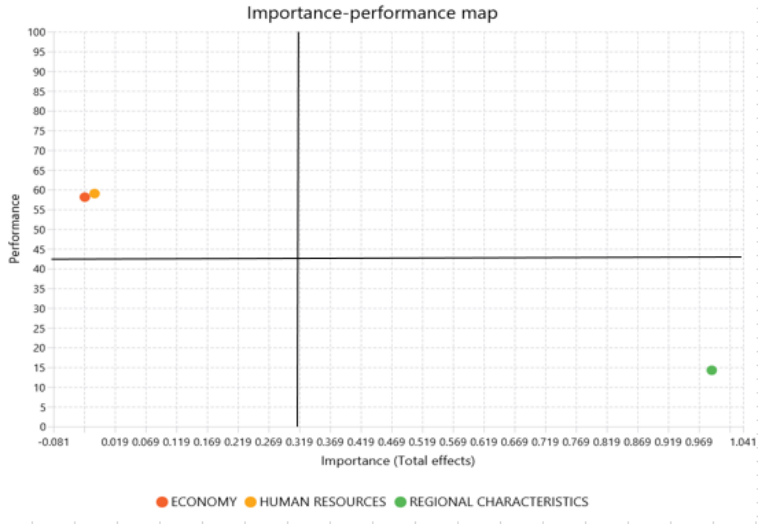


Fig. 2: Importance–Performance Map Analysis (IPMA) for Segment 1

The results of the IPMA analysis for Segment 1 are shown in Fig. 2. In this case, regional characteristics are in the top priority quadrant. This implies that the variable is very important but needs performance enhancement, making it a target for infrastructure development efforts. In contrast, the variables of the economy and human resources are in the quadrant of Possible Oversizing. This shows that even with relatively good performance, their contribution to the improvement of infrastructure in this segment is still limited. Therefore, when upgrading infrastructure, it is more appropriate to build on the regional characteristics of the provinces in Segment 1, rather than their economic factors or human resources.

3.4.2. Analisis IPMA pada Segmen 2

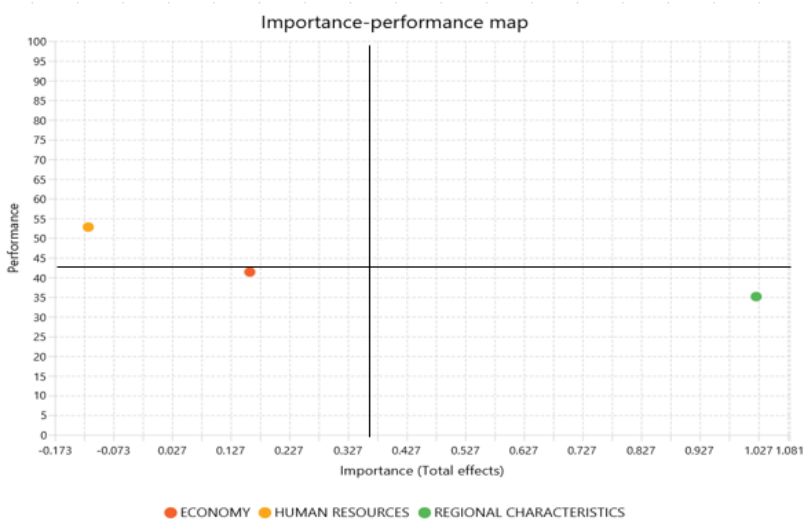


Fig. 3: Importance–Performance Map Analysis (IPMA) for Segment 2

The results of the IPMA analysis for Segment 2, shown in Fig. 3 put the regional characteristics in the high-priority quadrant, making it a key priority for infrastructure improvement efforts. In contrast, the economy factor is in the low-priority quadrant, indicating that it is relatively low in importance and performance of infrastructure. The human resources factor is in the Possible Oversizing quadrant, which means a good level of performance but a still relatively low contribution to infrastructure improvement. The results show that in the group of provinces included in the segment, the regional characteristics are still the most important factor for infrastructure development.

3.4.3. Analsis IPMA pada Segmen 3

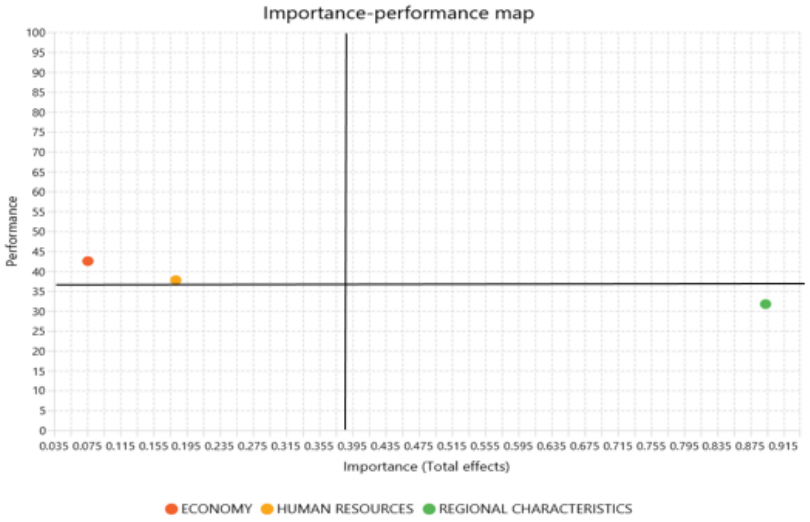


Fig. 4: Importance–Performance Map Analysis (IPMA) for Segment 3

The IPMA results presented in Fig. 4 indicate that regional characteristics fall into the Top Priority category, making them a priority factor in infrastructure improvement efforts. Conversely, the economy and human resources fall into the Possible Oversizing category, meaning that while these two factors perform relatively well, their contribution to infrastructure improvement remains lower than that of regional characteristics. These results indicate that regional conditions remain a key factor influencing infrastructure development in the provincial groups of Segment 3.

3.4.4. Analisis IPMA pada Segmen 4

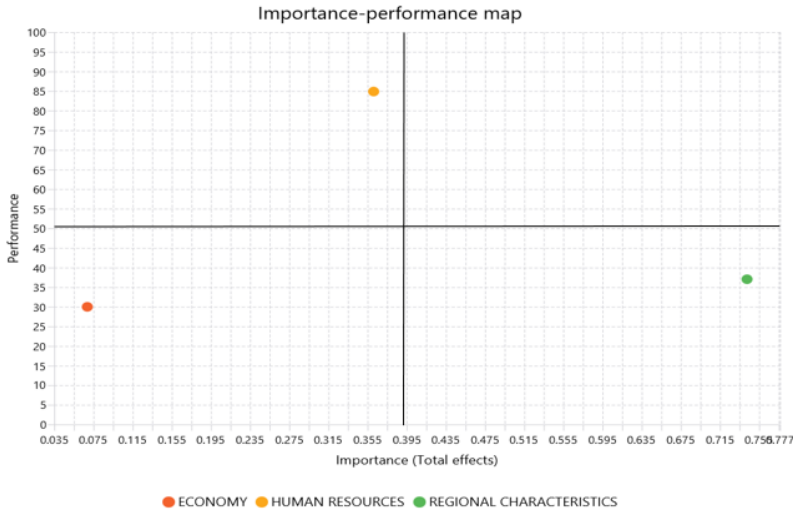


Fig. 5: Importance–Performance Map Analysis (IPMA) for Segment 4

Based on Fig. 5, the IPMA results for Segment 4 indicate that regional characteristics fall into the High Priority category, making them a priority in infrastructure improvement efforts. The economy, on the other hand, falls into the Low Priority category, meaning its influence on infrastructure is relatively weak and it has not yet become a major area for improvement. Conversely, human resources are in the Possible Oversizing category, indicating that this element already performs well, but further improvements would not have as significant an impact as those made to regional characteristics. These results show that infrastructure development in segment 4 remains more influenced by regional characteristics than by other factors.

These results indicate that regional characteristics are the predominant factor influencing infrastructure. This is also consistent with the results of effect size tests, which demonstrated that this factor contributes most to the model's ability to explain variations in infrastructure.

The predominant influence of regional characteristics indicates that these characteristics, particularly those related to disaster risk, play a significant role in the state of infrastructure. Regions highly vulnerable to disasters face increased difficulties in maintaining the quality and sustainability of

their infrastructure. These results are consistent with research indicating that various types of natural disasters can increase the vulnerability of infrastructure, particularly transport-related infrastructure [17]. Furthermore, research conducted in Indonesia also shows that natural disasters can disrupt the state of infrastructure and the continuity of public services, making the development of resilient infrastructure a crucial element of disaster risk reduction efforts [21]. Thus, the results of this study indicate that regional characteristics are an essential factor to consider in infrastructure planning and development, particularly in areas at high risk of disaster.

4. Conclusion

This study aims to analyze the impact of the economy, human resources, and regional characteristics on infrastructure in 38 Indonesian provinces using partial least squares-structural equation modelling (PLS-SEM). The results show that regional characteristics are the only factor with a significant influence on infrastructure development, while the economy and human resources have no significant influence. Finite mixture-partial least squares (FIMIX-PLS) analysis also identified four segments exhibiting previously unobserved heterogeneity between Indonesian provinces. These results indicate that the relationship between the variables is not entirely homogeneous; therefore, the approach used at the overall population level does not yet allow for a description of the characteristics of each regional group.

Analysis of the Importance-Performance Map (IPMA) reveals that each segment has different priorities for improvement. However, regional characteristics consistently emerge as a factor of high importance, associated with relatively low performance, making them a top priority for infrastructure development improvements. The findings suggest that infrastructure development policies should be formulated based on the specific characteristics of each regional group rather than a one-size-fits-all approach. The study gives a detailed view of the relations between different factors and defines development priorities taking into account the differences between provinces.

This study still has certain limitations. Since the analysis relied on secondary data from 38 observation units, it is unable to illustrate how development conditions have changed over time. Furthermore, the indicators are based only on data provided by the Central Statistics Agency (BPS), which means that some factors that could affect infrastructure development have been left out. Therefore, it is recommended that further research be conducted using panel or longitudinal data, incorporating indicators related to institutional aspects, governance, and regional fiscal capacity, and comparing the FIMIX-PLS approach with other segmentation methods to gain a more comprehensive understanding of the heterogeneity of infrastructure development in Indonesia.

CRediT Authorship Contribution Statement

Arum Nanda Putri Aprilia: Conceptualization, Methodology, Software, Formal Analysis, Data Curation, Visualization, Writing – Original Draft. **Rahmawati Erma Standsyah:** Conceptualization, Methodology, Validation, Writing – Review & Editing, Supervision.

Declaration of Generative AI and AI-assisted technologies

During the preparation of this study, the authors used ChatGPT for translation and language editing purposes. After using this tool, the authors reviewed, verified, and edited the content as needed and take full responbility for the content of the publication.

Declaration of Competing Interests

The authors declare no competing interests.

Funding and Acknowledgments

This This research was not supported by any specific grant from public, commercial, or non-profit funding agencies.

We would like to thank Statistics Indonesia (Badan Pusat Statistik/BPS) for supplying the official statistical data used in this study, including the Village Potential Statistics (PODES), Human Development Index (HDI), and other key publications that served as our main data sources.

Data and Code Availability

The data for this study come from official publications by Statistics Indonesia (Badan Pusat Statistik/BPS), such as the Village Potential Statistics (PODES), Human Development Index (HDI), and other official reports. If you would like access to the processed dataset used in this research, please contact the corresponding author.

References

[1] M. A. Berawi, M. Sari, and P. Miraj, "Fostering Infrastructure Development Aligned With Sustainable Development Goals," *CSID-JID*, vol. 4, no. 1, p. 1, May 2021, doi: 10.32783/csid-jid.v4i1.236.

[2] I. Kustova, J. Hudenko, and N. Lace, "A Systematic Review of Sustainability Criteria in Infrastructure Development," *Sustainability*, vol. 16, no. 11, p. 4564, May 2024, doi: 10.3390/su16114564.

[3] A. Jatayu *et al.*, "Measuring Levels of Infrastructure Development and its Impact on Regional Growth - Insights from Indonesia," *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 1353, no. 1, p. 012011, May 2024, doi: 10.1088/1755-1315/1353/1/012011.

[4] R. D. Ramadhan and F. Z. Fauzi, "DRIVING PROGRESS TOGETHER: EMBRACING INCLUSIVE DEVELOPMENT FOR INFRASTRUCTURE IN EAST INDONESIA," *J-3P*, pp. 93–112, Nov. 2023, doi: 10.33701/j-3p.v8i2.3691.

[5] W. Nurdina, "Infrastructure and Income Inequality in Indonesia: 2009-2017," *JISDeP*, vol. 2, no. 2, pp. 129–144, Aug. 2021, doi: 10.46456/jisdep.v2i2.132.

[6] K. Ariansyah *et al.*, "Unleashing the potential of mobile broadband: Evidence from Indonesia's underdeveloped regions on its role in reducing income inequality," *Telematics and Informatics*, vol. 82, p. 102012, Aug. 2023, doi: 10.1016/j.tele.2023.102012.

[7] R. N. T. Wahyuni, M. Ikhsan, A. Damayanti, and K. Khoirunurrofik, "Inter-District Road Infrastructure and Spatial Inequality in Rural Indonesia," *Economies*, vol. 10, no. 9, p. 229, Sep. 2022, doi: 10.3390/economies10090229.

[8] P. R. Halim, A. A. Yusuf, A. Sumner, and R. M. Purnagunawan, "Structural transformation and regional inequality differential in Indonesia," *Regional Studies, Regional Science*, vol. 12, no. 1, pp. 790–815, Dec. 2025, doi: 10.1080/21681376.2025.2565318.

[9] J. C. Londono, B. Wilson, and F. Osorio-Tinoco, "Understanding the entrepreneurial intentions of youth: a PLS multi-group and FIMIX analysis using the model of goal-directed behavior," *JEEE*, vol. 13, no. 3, pp. 301–326, Jul. 2021, doi: 10.1108/JEEE-07-2019-0090.

[10] B. W. Otok, Purhadi, R. Sriningsih, and D. S. Dila, "Segmentation of toddler nutritional status using REBUS and FIMIX partial least squares in Southeast Sulawesi," *MethodsX*, vol. 12, p. 102515, Jun. 2024, doi: 10.1016/j.mex.2023.102515.

[11] A. S. H. Wahyuningtyas *et al.*, "Consumer segmentation analysis of microalgae-based foods in Indonesia: A finite mixture partial least squares (FIMIX-PLS) approach," *Sustainable Futures*, vol. 10, p. 101555, Dec. 2025, doi: 10.1016/j.sftr.2025.101555.

[12] Y. Yuan, M. Yang, T. Feng, S. Rasouli, D. Li, and X. Ruan, "Heterogeneity in passenger satisfaction with air-rail integration services: Results of a finite mixture partial least squares model," *Transportation Research Part A: Policy and Practice*, vol. 147, pp. 133–158, May 2021, doi: 10.1016/j.tra.2021.03.003.

[13] H. Joshi and D. Chawla, "Identifying unobserved heterogeneity in mobile wallet adoption – A FIMIX-PLS approach for user segmentation," *IJBM*, vol. 41, no. 1, pp. 210–236, Jan. 2023, doi: 10.1108/IJBM-03-2022-0132.

[14] A. Jo and S. An, "Customer Segmentation in Social Media Marketing: Analyzing Hidden Customer Heterogeneity through Finite Mixture Partial Least Squares (FIMIX-PLS) and Importance-Performance Map Analysis (IPMA)," *Bus. Commun. Res. Pract.*, vol. 7, no. 1, pp. 35–47, Jun. 2024, doi: 10.22682/bcrp.2024.7.1.35.

[15] D. A. Fadhila and B. W. Otok, "Pemodelan Faktor-Faktor yang Mempengaruhi Daerah Tertinggal di Pulau Jawa Menggunakan Second-Order Confirmatory Factor Analysis (CFA)," *JSSITS*, vol. 8, no. 2, pp. D153–D160, Feb. 2020, doi: 10.12962/j23373520.v8i2.43216.

[16] B. W. Otok, A. Suharsono, Purhadi, R. E. Standsyah, and H. Al Azies, "MASEM Infrastructure in Underdeveloped Areas of Java Island," *Journal of Southwest Jiaotong University*, vol. 56, no. 1, p. 10, 2021, doi: 10.35741/issn.0258-2724.56.1.10.

[17] S. A. Argyroudis, S. A. Mitoulis, M. G. Winter, and A. M. Kaynia, "Fragility of transport assets exposed to multiple hazards: State-of-the-art review toward infrastructural resilience," *Reliability Engineering & System Safety*, vol. 191, p. 106567, Nov. 2019, doi: 10.1016/j.ress.2019.106567.

[18] J. F. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A primer on partial least squares structural equation modelling (PLS-SEM)*, Second edition. Los Angeles, London, New Delhi, Singapore, Washington, DC, Melbourne: SAGE, 2017.

[19] M. Sarstedt, J.-M. Becker, C. M. Ringle, and M. Schwaiger, "Uncovering and Treating Unobserved Heterogeneity with FIMIX-PLS: Which Model Selection Criterion Provides an Appropriate Number of Segments?," *Schmalenbach Bus Rev*, vol. 63, no. 1, pp. 34–62, Jan. 2011, doi: 10.1007/BF03396886.

[20] C. M. Ringle and M. Sarstedt, "Gain more insight from your PLS-SEM results: The importance-performance map analysis," *IMDS*, vol. 116, no. 9, pp. 1865–1886, Oct. 2016, doi: 10.1108/IMDS-10-2015-0449.

[21] D. N. Pascapurnama, A. Murakami, H. Chagan-Yasutan, T. Hattori, H. Sasaki, and S. Egawa, "Integrated health education in disaster risk reduction: Lessons learned from disease outbreak following natural disasters in Indonesia," *International Journal of Disaster Risk Reduction*, vol. 29, pp. 94–102, Aug. 2018, doi: 10.1016/j.ijdrr.2017.07.013.