Contents lists available at DJM

DESIMAL: JURNAL MATEMATIKA
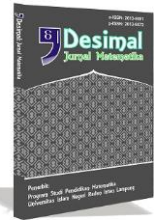
Logistic regression modelling at insurance interest: Theory of planned behavior

Lukmanul Hakim*, Agus Santoso

Sumbawa University of Technology, Indonesia

ARTICLE   INFO

ABSTRACT

One of the characteristics of people who have good financial literacy is having good risk management skills. This study aims to make logistic regression modeling of risk management or insurance interests from Sumbawa University of Technology lecturers. This research uses quantitative methods by combining mathematical theory, statistics, and management science (Theory of Planned Behavior). This research uses primary data from Sumbawa University of Technology as samples. From the data collected using the probability sampling method, it was found that of the four variables (literacy as X_1, inclusion as X_2, family as X_3, gender as D_1, and insurance interest as Y) used, only the family variable had a significant impact on increasing someone's interest in insurance (opportunity). The exponential values of the four variables are: $X_1 = 3.230$ means 1 unit in the literacy variable will have an impact on the possibility of increasing interest in insurance by 3.230 times, $X_2 = 0.548$ means 1 unit in the inclusion variable will have an impact on the possibility of increasing interest in insurance by 0.548 times, $X_3 = 3.886$ means 1 unit in the family variable will have an impact on the possibility of increasing interest in insurance by 3.886 times, and $D_1 = 0.813$ means the male gender has an impact 0.813 times greater than the female sex on increasing interest in insurance.

http://ejournal.radenintan.ac.id/index.php/desimal/index

INTRODUCTION

Insurance protection has an important role for society. Insurance is a very important form of risk transfer. Insurance has many forms, from health, life, and vehicle insurance to accident insurance. Insurance service providers are quite diverse, ranging from government to private. Based on data from the Indonesian Life Insurance Association (AAJI), 85.01 million Indonesians will be insured in 2022, and 69.62% will have life insurance (BPJS). Ownership of this insurance is very small compared to Indonesia's population, which reaches 275.77 million. Indonesia is ranked 3rd

out of other ASEAN countries. It shows that Indonesia's insurance interest still needs to grow compared to the other four ASEAN countries.
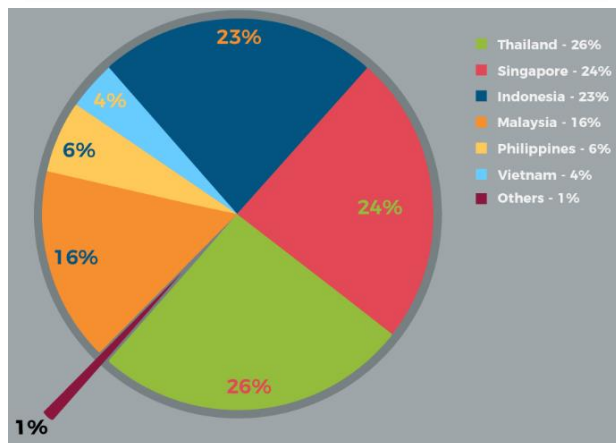
Figure 1. Insurance Ownership Rate
Source: ASEAN Insurance Pulse

The demographic bonus is a problem if it is not covered by insurance. Based on BPS 2022 data, the average health cost is 15.24% of per capita income. The large number of health costs will undoubtedly affect the community's welfare. To encourage interest in insurance, it needs encouragement from other parties, including the family. Family being the closest person can motivate individuals to behave. This can be reinforced by research on demographic factors affecting interest in insurance. The low number of insurance users shows a need for more public understanding of the benefits of insurance products. It can be seen from the OJK survey that the insurance literacy rate is 31.72%. Product knowledge is one of the determining factors for consumers to be interested. Knowledge is proof that consumers can interpret the value of a product. Product knowledge can influence interest in insurance.

The difficulty of getting access to information about insurance products is a factor in the low literacy of insurance in the community. The low insurance literacy in Indonesia cannot be separated from the influence of insurance inclusion, which is still low. Data for 2022 shows that the insurance inclusion rate in Indonesia is 16.63%. The low inclusion of insurance makes Indonesians vulnerable to health issues. This can be caused by the difficulty of accessing insurance provider institutions, both unreachable locations and limited information centers, and difficulties in managing administration. A survey conducted by prudential Indonesia in 2022 found that more quality and marketers to provide information and reach a wider community can drive an increase in the number of insurance users. Sumbawa University of Technology (UTS) is a university located in Sumbawa regency, West Nusa Tenggara. It has lecturers and staff with a total of 298 people with diverse backgrounds. Sumbawa University of Technology does not have dormitories or special residences for lecturers and staff; therefore, the distribution of lecturer residences is very diverse.

The distance of lecturers to campus is relatively far; changing natural conditions have the possibility of risks arising in carrying out routines, both in the form of health problems and work accidents. In the end, we try to figure out which factors could increase the chances of UTS lecturers being insured. This study aims to: 1) test the effect of product knowledge (X_1) on the interest of insured lecturers (Y); 2) test the influence of family (X_2) on the interest of insured lecturers (Y); 3) test the effect of gender (D_1) on the interest of insured lecturers (Y); 4) test the effect of inclusion (X_3) on the interest of insured lecturers (Y), using logistic regression model modeling as the approach used. The logistic model is used because it gives an idea of how big the chance is of the dependent variable changing if it is influenced by the independent and dummy variables. Apart from that, the logistic model is more likely to be used because it does not require normality and linearity assumptions.

METHOD

The type of research used in this study is the Quantitative Method. Creswell & Plano-Clark (2007) stated that quantitative research tries to find the right conclusions by using numbers as a tool to analyze and extract what information is desired. The quantitative analysis used in this study is a Logistic Regression analysis with the Theory of Planned Behavior as an approach model to understand the object of the research. Figure 2 is the research flow.
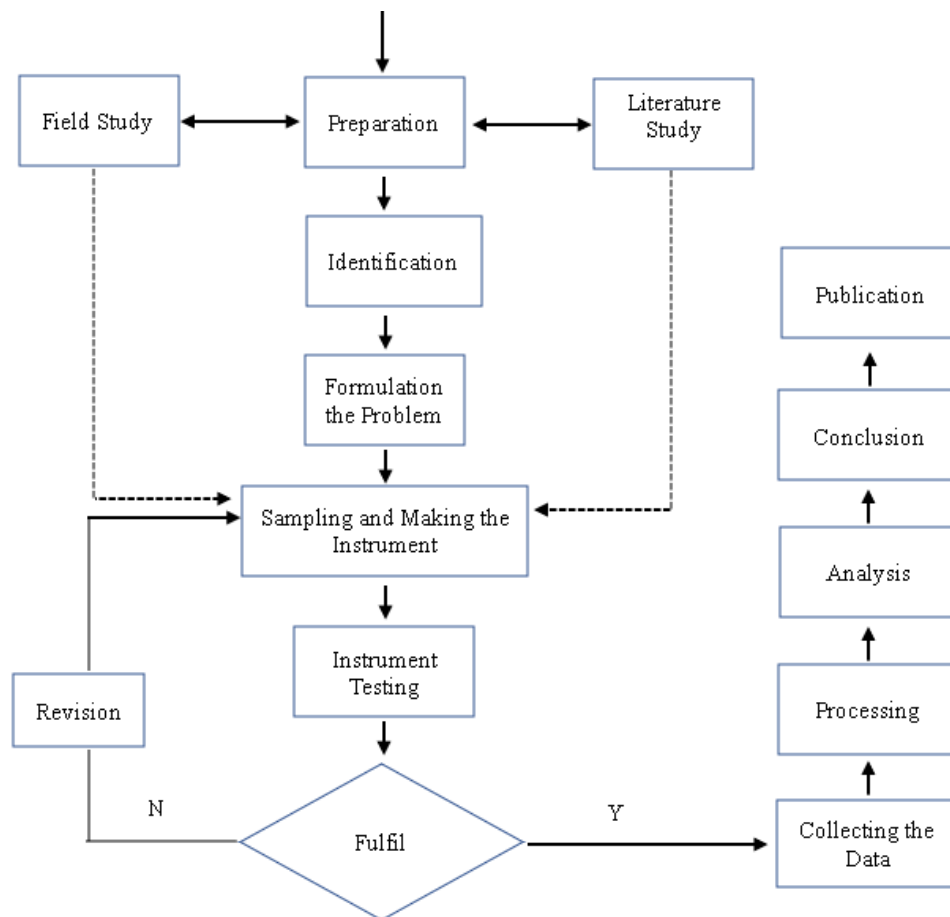

Figure 2. Research Flow Chart

This research tries to test hypotheses in mathematical and statistical modeling with the help of management science, especially the Planned Theory of Behavior. In this study, there are 5 variables used, namely: 1) product literacy and knowledge; 2) inclusion; 3) family; 4) gender; and 5) interest in insurance. The research indicators (the instrument was created by researchers with indicators) used in this study are shown in Table 1.

Tabel 1. Research Instrument Based on Theory of Planned Behavior (TPB)

Variable	Indicators	Scale
Literacy as X_1	1. Product attributes 2. Physical benefits 3. Psychological benefits 4. Product Value (Kotler & Armstrong, 2018)	Likert
Inclusion as X_2	1. Access 2. Product Services 3. Usage 4. Quality (Otoritas Jasa Keuangan, 2019)	Likert
Family as X_3	1. Information support 2. Assessment Support 3. Instrumental Support, 4. Emotional Support (Rina & Alimuddin, 2015)	Likert
Gender as D_1	-	Nominal
Interest as Y	1. Transactional Interest 2. References Interest 3. Preferential Interest 4. Exploration Interest (Ferdinand, 2006)	Nominal/Biner

In this study, the population was the lecturers and staff of Sumbawa University of Technology, which amounted to 298 people. In this study, samples were taken partially from members of the population without removing the characteristics of the population. In this study, the number of samples was determined using the Slovin formula (significance level: 10%) (Karakaya-Ozyer & Aksu-Dunya, 2018).
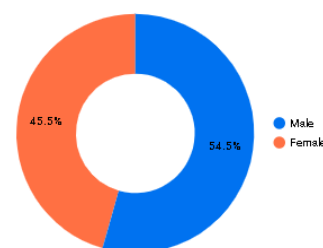
$$n = \frac{N}{1 + Ne^2}$$

So that obtained

$$n = \frac{298}{1 + (298)(0.1^2)}$$
$$n = 74.8 \approx 75$$

The sampling technique in this research is systematic random sampling. This technique is one that provides equal opportunities to every member of the population (Creswell & Plano-Clark, 2007).

Figure 3. Respondents Distribution Based on Gender

Logistic Regression

Logistic regression model analysis is used to model the level of interest of Sumbawa University of Technology lecturers and staff in insurance based on the data obtained. In the modeling process, there are several things and assumptions to pay attention to in logistic regression, including: 1) It does not require linearity and normality assumptions; 2) It does not require homoscedasticity assumptions; 3) The dependent variable is categorical. The four categories in the independent variable are exclusive. There are two types of data scales used, namely binary data scales for

family, gender, insurance interest, and Likert scales for other variables. Here is the logistic regression model used (Schmee & Anderson, 1986):

$$\ln(Y) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 D_1 + e$$

where,

$$Y = \frac{\hat{\rho}}{1 - \hat{\rho}}$$

$$\hat{\rho} = \frac{Exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 D_1 + e)}{1 + Exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 D_1 + e)}$$

Information:
X_1 : Literacy

X_2 : Inclusion

X_3 : Family

D_1 : Gender

Y : Insurance Interest

The statistical analysis of the data used in this study was a logistic regression analysis. According to Ghozali (2006), logistic regression analysis is a regression that tests whether there is a probability that the occurrence of the dependent variable can be predicted by the independent variable. Logistic regression analysis does not require classical assumptions. Logistic regression analysis has four tests (Ghozali, 2006):

Hypothesis: Hypothesis testing is used to determine the significance of the parameter. The following tests are carried out:

$$H_0: \beta_0 = \beta_1 = \cdots = \beta_j = 0$$
$$H_1: \beta_0 = \beta_1 = \cdots = \beta_j \neq 0$$

with rejection areas: reject H_0 if $G > \chi^2_{(x,\alpha)}$ or $P - value < \alpha$ where v is the degrees of freedom (many predictor variables are in the model without B0)

Overall Model Fit: The overall model fit is used to find out if all independent variables affect the dependent variable. The statistics used are based on the Likelihood theorem. Likelihood L is the probability that the hypothesized model describes the input data (Ghozali, 2006).

$$l(\beta) = \prod_{i=1}^{n} \pi(X_i)^{y_i} (1 - \pi((X_i))^{1-y_i}$$

where,

$$\pi(X_i) = \frac{\exp(\sum_{j=0}^{p} \beta_i x_{ij})}{1 + \exp(\sum_{j=0}^{p} \beta_i x_{ij})}$$

$$L(\beta) = \ln[l(\beta)]$$

Goodness of Fit Test: The feasibility test of the regression model was assessed using Hosmer and Lemeshow's as measured by the chi-square value. This model is to test the null hypothesis that the empirical data is in accordance with the model (there is no difference between the model and the data so that the model can be said to be fit) (Ghozali, 2006).

$$\hat{C} = \sum \frac{(O_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)}$$

Nagelkerke R-Square: The coefficient of determination in logistic regression is seen from the Nagelkerke R-Square because the value of the Nagelkerke R-Square can be interpreted like the value of the R-Square in multiple regression. Nagelkerke R-Square is a modification of the coefficients of Cox and Snell to ensure that the value will vary from 0-1. A Nagelkerke R-Square value close to zero indicates that the ability of the variables to explain the dependent variable is very limited, while a Nagelkerke R-Square value close to one indicates that the independent variable is able to provide all the information needed to predict the variability of the dependent variable (Ghozali, 2006).

Classification Matrix: The classification matrix is used to explain the power of

regression models to predict the likelihood of insurance interest. In the 2x2 table, the correct and incorrect estimated values are calculated. The classification table produces overall accuracy (Ghozali, 2006).

Wald Test: Ghozali (2006) said that the Wald test (t), which basically shows how far the independent variable partially influences the dependent variable, to determine the value of the Wald test (t-test), the significance level is 5%.

Omnibus Tests of Model Coefficients: Omnibus Tests of Model Coefficients is a simultaneous statistical test (F-test). This study will test whether the independent variable simultaneously affects the dependent variable (Ghozali, 2013). The significance rate is 5%.

RESULTS AND DISCUSSION

Characteristics of Respondents

The respondents used in this study were active lecturers at Sumbawa University of Technology. Here is a descriptive picture of the respondents used: 1) 62.5% of respondents are women, and the rest are men; 2) Respondents are dominated by lecturers aged 31–35 years, where this generation is still classified as the millennial generation; 3) The distribution of monthly opinions obtained by respondents is evenly distributed from the lowest income of 2 million to the highest in the range of >6 million..
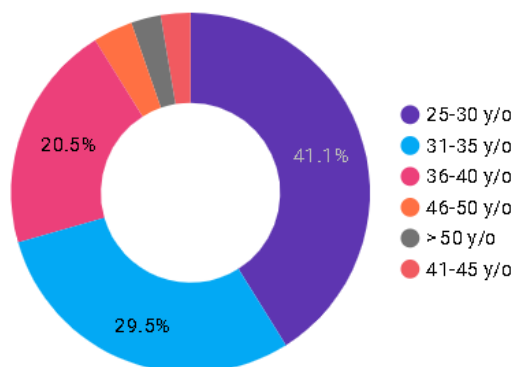
Figure 4. Respondent Distribution Based on Age
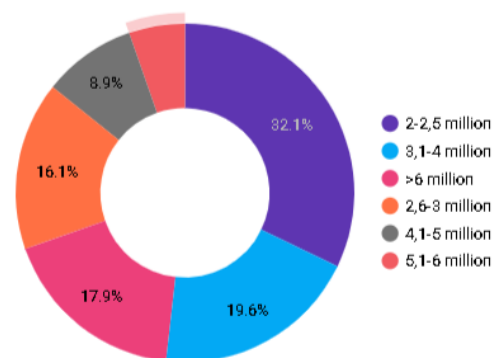
Figure 5. Respondent Distribution Based on Income

This proves that even though they have an income that is not large enough, it does not reduce respondents' interest in having insurance. When related to the type of work that respondents have, many respondents work as private employees. Some private companies sometimes do not guarantee the self-protection of their employees, so an individual who works as a private employee will seek self-protection because he is aware that the risks he has must be managed, namely through life insurance.

Instrument Testing

Before further analysis, the research team tested the research instrument by testing its validity (how well the test measures that characteristic) and reliability (how dependably or consistently a test measures a characteristic) (Ghozali, 2006). Instrument testing was done using 30 other respondents. The test results are as in Table 2.

Table 2. Reliability Test

Reliability Statistics	
Cronbach's Alpha	N of Items
.778	8
Reliability Statistics	
Cronbach's Alpha	N of Items
.788	9
Reliability Statistics	
Cronbach's Alpha	N of Items
.805	8

Based on the table mentioned above, the reliability values are 0.778, 0.788, and 0.805, respectively, where all three have values greater than 0.6, so all variables can be declared reliable or consistent (Bahri, 2018).

Logistic Regression

Analysis Furthermore, data testing is carried out to determine the hypothesis test:

$H_0: \beta_j = 0$; (X_j has no effect on Y)
$H_0: \beta_j = 0$;
$j = 1,2,3,4$

and the logistics model formed. The results of the testing data obtained, as shown in Table 3.

Table 3. Odds Ratio and Logistic Regression Test

Variables in the Equation							
		B	$S.E.$	$Wald$	df	$Sig.$	$Exp(B)$
Step 1[a]	X_1	1.172	.669	3.070	1	.080	3.230
	X_2	-.601	.606	.982	1	.322	.548
	X_3	1.357	.514	6.986	1	.008	3.886
	X_4 (1)	-.207	.699	.087	1	.768	.813
	Constant	-2.721	1.326	4.207	1	.040	.066

a. Variable(s) entered on step 1: X_1, X_2, X_3, X_4.

From Table 3, we obtained the logistic regression equation (Hair, Black, Babin, & Anderson, 2014):

$$\ln(y) = -2.71 + 1.172x_1 - 0.601x_2 + 1.357x_3 - 0.207x_4 + e$$

Based on the above equation, the following results were obtained:

- An increase of 1 unit in the literacy (x_1) variable will have an impact on the possibility of increasing interest in insurance by 3.230 times
- An increase of 1 unit in the inclusion (x_2) variable will have an impact on the possibility of increasing interest in insurance by 0.548 times
- An increase of 1 unit in the family (x_3) variable will have an impact on the possibility of increasing interest in insurance by 3,886 times
- The male gender (D_1) has an impact of 0.813 times greater than female sexual interest in insurance.

Of the four variables used, based on hypothesis testing, family variables are the only ones that have a significant influence on increasing insurance interest among Sumbawa University of Technology lecturers (significant value 0.008<0.05) (Ferdinand, 2006; Schmee & Anderson, 1986).

Literacy variability does not have a significant effect (based on hypothesis testing in Table 3) because lecturers and staff of Sumbawa University of Technology have a high sense of tolerance for health risks, with the majority of Muslims having the belief that every pain comes from Allah S.W.T. and there must be a cure. In addition, insurance products that are familiar to the public are only limited to life insurance and health insurance, even though insurance products are very broad, for example, property insurance, credit insurance, and property insurance. This happens because the availability of information on insurance products is still very limited in Indonesia. Research conducted by Yuniawati & Asiyah (2022) has the same results where the literacy variable has no effect on the variable of interest. Knowledge of insurance products by the community is limited to knowledge without any concrete implementation.

The inclusion variability does not affect the interest in insurance among lecturers and staff at Sumbawa University of Technology; this is due to the limited availability of insurance provider branch offices in Sumbawa. If you look at the OJK survey data for 2022, the insurance inclusion rate is 16.63%; this is certainly very different from financial inclusion, which has reached 74.03%.

The family aspect has a significant influence that is inseparable from the fact that the majority of lecturers and staff at Sumbawa University of Technology are family; this is reinforced by research by (Laturrakhmi, Swastikawara, & Wardasari, 2020). The portion of the decision to use insurance cannot be made entirely personally, requiring communal decisions in the form of family considerations as a reference in decision-making. Lecturers, as educators, really understand the importance of the principles of deliberation when making decisions in the family environment. With insurance for the head of the family, it becomes one form of responsibility for health and safety protection for the family. If it is related to the SDGs as the theoretical basis of this research, this condition is part of their attitude toward, which is a response to the benefits of insurance products for their families.

In addition, life insurance decisions between men and women are also phenomenally influenced by the increasing opening of job opportunities both for men and for women. As the number of working individuals increases, they will have an income and be able to set aside their income for the things they need. This phenomenon was studied by Hermawati (2013) and Pangeran (2013). They state that between men and women, there are different levels of understanding of insurance that will influence an individual's decision to have insurance. Specifically, demographics, including gender, are not a differentiating factor in a person's financial planning or do not have an influence on an individual's attitude and behavior in financial planning, including insurance planning.

Table 4. Goodness of Fit Test

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	8.042	8	.429

Based on Table 4, it can be seen that resulting in a failed decision to reject H0, the model is appropriate (there is no significant difference between the observation results and the possible prediction results of the model).

Table 5. Nagelke R-Square Test

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	58.230[a]	.235	.369

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

The value of the coefficient of determination of the four variables used is 0.369, meaning that the ability of all independent variables to explain the probability of increasing or decreasing interest in insurance at Sumbawa University of Technology lecturers is 36.9% (it means 63.1% will explain by other variables that's not included in this research, such as marriage status, demography, region, etc.).

CONCLUSIONS AND SUGGESTIONS

Based on the research results, it was concluded that UTS lecturers really consider the family's opinion in deciding

whether to buy an insurance policy or not. Family variables provide the opportunity for significant influence in decision-making because insurance usually means spending extra money on a risk that the family may not experience.

Based on the findings, researchers suggest that the next research use a mixed-methods approach in answering this research question. This can help researchers determine in advance the dominant variables that will be used during research through qualitative methods and then model them using quantitative methods.

ACKNOWLEDGEMENT

REFERENCES

Bahri, S. (2018). *Metodologi penelitian bisnis lengkap dengan teknik pengolahan data spss*. Yogyakarta: ANDI. Retrieved from https://openlibrary.telkomuniversity.ac.id/home/catalog/id/150287/slug/metodologi-penelitian-bisnis-lengkap-dengan-teknik-pengolahan-data-spss.html

Creswell, J. W., & Plano-Clark, V. I. (2007). Choosing a mmixed method design. Designing and constructing mixed methods research. *London:Sage Publication*.

Ferdinand, A. (2006). *Metode penelitian mananjemen: Pedoman penelitian untuk penulisan skripsi, tesis, dan disertasi ilmu manajemen* (2nd ed.). Semarang: Badan Penerbit Universitas Diponegoro.

Ghozali, I. (2006). Aplikasi analisis multivariate dengan program spss. In *Analisis multivariate dengan program spss*.

Ghozali, I. (2013). Structural equation modeling: Metode alternatif dengan partial least squares (PLS). In *Seminar Nasional Matematika dan Aplikasinya* (4th ed.). Semarang: Badan Penerbit Universitas Diponegoro Semarang.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.).

Hermawati, S. (2013). Pengaruh gender, tingkat pendidikan, dan usia terhadap kesadaran berasuransi pada masyarakat indonesia. *Jurnal Asuransi Dan Manajemen Risiko*, *6*(1), 64–67.

Karakaya-Ozyer, K., & Aksu-Dunya, B. (2018). A review of structural equation modeling applications in turkish educational science literature, 2010-2015. *International Journal of Research in Education and Science*, 279–291. https://doi.org/10.21890/ijres.383177

Kotler, P., & Armstrong, G. (2018). *Principles of marketing* (17th ed.). Pearson Education.

Laturrakhmi, Y. F., Swastikawara, S., & Wardasari, N. (2020). Analisis perilaku masyarakat pedesaan terhadap asuransi kesehatan nasional dalam perspektif komunikasi kesehatan. *Komuniti : Jurnal Komunikasi Dan Teknologi Informasi*, *12*(2), 87–100. https://doi.org/10.23917/komuniti.v12i2.10040

Otoritas Jasa Keuangan. (2019). *Survei nasional literasi dan inklusi keuangan 2019 (national literacy and financial inclusion survey 2019)*.

Pangeran, P. (2013). Praktik perencanaan keuangan: Studi empiris tentang sikap dan perilaku keuangan rumah tangga desa. *KINERJA*, *17*(2), 197–216. https://doi.org/10.24002/kinerja.v17i2.380

Rina, R., & Alimuddin. (2015). Education, environment family, relationships and lifestyle behaviors effect against irregularities work and finance (empirical study on employee division of finance &accounting pt. meratus line). *E-Journal Magistra Management Science*, *1*(1), 4–5.

Schmee, J., & Anderson, T. W. (1986). An introduction to multivariate statistical analysis. *Technometrics*, *28*(2), 180. https://doi.org/10.2307/1270458

Yuniawati, A. L., & Asiyah, B. N. (2022). Pengaruh literasi keuangan dan sosialisasi terhadap minat berinvestasi mahasiswa di pasar modal syariah melalui galeri investasi syariah. *Jurnal Economina*, *1*(4), 829–840. https://doi.org/10.55681/economina.v1i4.187