Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 5, 2830-2835
2025
Publisher: Learning Gate
DOI: 10.55214/25768484.v9i5.7600

An analysis of factors influencing public transportation use by Mulawarman university students in Samarinda City using the binary logistic regression

Gultom Tiopan HM[1*], Jamal Mardewi[2], Fatah M Khairul[3]
[1,2,3]Faculty of Engineering, Mulawarman University, Indonesia; tiopanhmg@gmail.com (G.T.) wie_djamal@gmail.com (J.M.) khairulfatah01@gmail.com (F.M.K.).

Abstract: The students of Mulawarman University represent a significant potential group of public transportation users in Samarinda. Various factors influence their decision to use public transportation, such as comfort, cost, travel time, and personal preferences. To increase the use of public transportation among the students, it is crucial to understand these factors thoroughly. This research aims to analyze the factors affecting public transport usage among Mulawarman University students using the binary logistic regression method. The SPSS (Statistical Program for Social Sciences) software was employed for statistical analysis. A survey was conducted for this research applying the random sampling approach in which students were the sole respondents of the questionnaires. The analysis used binary logistic regression, including multicollinearity detection, parameter estimation via Maximum Likelihood Estimation (MLE), and significance testing. The results show that the distance from student residence to campus, public transportation cost, accessibility, private vehicle ownership, and travel time significantly affect public transport usage. However, public transportation cost, accessibility, and private vehicle ownership do not significantly influence the usage when considered partially. Residential distance has a negative and significant effect, while travel time has a positive and significant effect on public transportation usage by the students.

Keywords: *Binary logistic regression, Public transport.*

1. Introduction

Transportation can be defined as an effort to move, transport, or transfer an object from one place to another, so the object becomes more useful or useful for certain purposes [1, 2]. There are 5 main elements of transportation, namely: (1) Humans, who need transportation; (2) Goods, which are needed by humans; (3) Vehicles, as a means of transportation; (4) Roads, as transportation infrastructure; and (5) Organizations, as transportation managers. The five elements are interrelated for the implementation of transportation [3, 4]. The transportation process is created due to differences in needs between humans and one another, so that there is a movement from one place to another to get it [5, 6].

Statistics is the science of how to collect data, analyze, interpret and provide conclusions on the results of the analysis so that data is an absolute thing that must exist when statistics is used [7]. Statistics can be used as a communication tool [8, 9]. Statistics as communication is a liaison between several parties who produce statistical data or statistical analysis so that the parties will be able to make decisions by using this data or information. Then statistics as a description is the presentation and illustration of data, for example measuring production results, news coverage reports, consumer price indices, financial reports, inflation rates, population, and state income and expenditure results [10, 11].

History: Received: 13 March 2025; Revised: 6 May 2025; Accepted: 9 May 2025; Published: 28 May 2025
* Correspondence: tiopanhmg@gmail.com

Other uses of statistics include regression, which is to predict the effect of one data with other data and to anticipate future symptoms; correlation, which is to find the strength or magnitude of the data relationship in a study; and comparison, which is to compare data from two or more groups [12, 13].

Logistic regression is a statistical method applied to model a dependent variable that is categorical (nominal/ordinal scale) based on one or more independent variables that can be categorical or continuous variables (interval or ratio scale) [14, 15]. Logistic regression is part of a regression analysis that can be used if the dependent variable is a dichotomous variable, which is a variable that has different or opposite category values, usually consisting of only two values, which represent the occurrence or absence of an event which is usually given the number 0 or 1 [16, 17].

Binary logistic regression is also said to be one of the mathematical model approaches used to analyze the relationship of several factors with a variable that is dichotomous (binary) [18]. That is, in binary logistic regression, the data on the dependent variable is binary (0 and 1). The binary numbers describe two contrasting categories of data, such as 'yes or no,' 'success or failure,' and so on.

The general form of the logistic regression model:

$$g(X_{ki}) = ln\left[\frac{P(Y|X_{ki})}{P(Y|X_{ki})}\right] = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki}$$

Description:

$P(Y|X_{ki})$: $\frac{e^{(\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\cdots+\beta_k X_{ki})}}{1+e^{(\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\cdots+\beta_k X_{ki})}}$

Xki : the value of the kth independent variable of the i-th observation

βk : parameter value of the kth independent variable

k : index for independent variables (k = 0, 1, 2, ..., p)

i : index for the observation (i = 1, 2, ..., n)

Many methods can be used to estimate the value of model parameters, including the least squares method and Maximum Likelihood Estimation (MLE). However, one method that is suitable for estimating parameters in logistic regression analysis is MLE, which is a method for estimating unknown population parameters.

2. Methods

Mulawarman University is a state university in the province of East Kalimantan, established on September 27, 1962 based on the Decree of the Minister of Education and Science No. 130 of 1962 with four faculties, namely: the Faculty of Constitution and Order, the Faculty of Agriculture, the Faculty of Forestry, and the Faculty of Mining [19]. The main campus of Mulawarman University is located at Gunung Kelua, while other campuses are located on Jalan Pahlawan, Jalan Banggeris and Jalan Flores. Mulawarman University has thirteen faculties namely the Faculty of Economics and Business; Pharmacy; Law; Humanities; Social and Political Sciences; Medicine; Teacher Training and Education; Forestry; Public Health; Mathematics and Natural Sciences; Fisheries and Marine Science; Agriculture; and Engineering [20].

The number of undergraduate students by faculty until the odd semester 2023/2024 is Faculty of Economics and Business as many as 4,137 people, Faculty of Pharmacy 1,334 people, Faculty of Law 1,666 people, Faculty of Cultural Sciences 1,288 people, Faculty of Social and Political Sciences 4,365 people, Faculty of Medicine 670 people, Faculty of Teacher Training and Education 5.137 people, Faculty of Forestry 1,539 people, Faculty of Public Health 575 people, Faculty of Mathematics and Natural Sciences 1,480 people, Faculty of Fisheries and Marine Science 1,403 people, Faculty of Agriculture 1,753 people, and Faculty of Engineering 3,934 people. Using the Slovin formula with a 5% error rate, the sample size is obtained as follows:

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 5: 2830-2835, 2025
DOI: 10.55214/25768484.v9i5.7600

$$n = \frac{29.281}{1 + 29.281(0{,}05)^2} = \frac{29.281}{74{,}2025}$$
$$n = 394{,}6 \approx 395$$

The data collection technique used by researchers is interviews with respondents. Descriptive data collected by the researchers is in the form of questionnaire sheets. The independent variables in this study include the distance from student residences to the campus, the perception of the public transportation cost compared to those of other transportation modes, the accessibility of public transportation around student residences, the ownership of private vehicles, and travel time from student residences to campus.

Table 1.
Variable factors of public transportation use by Mulawarman University students.

Variables	Description	Type	Description
Y	Use of public transportation	Category	0 = Not using 1 = Using
X_1	Distance from residence to campus	Category	0 = < 1 km 1 = 1-3 km 2 = 3-5 km 3 = >10 km
X_2	Perceived cost of public transport compared to other modes of transportation	Category	0 = Very cheap 1 = Cheap 2 = Ordinary 3 = Expensive 4 = Very expensive
X_3	Public transport accessibility	Category	0 = Very easy 1 = Easy 2 = Easy enough 3 = Difficult 4 = Very difficult
X_4	Private vehicle ownership	Category	0 = Do not have 1 = Bicycle 2 = Motorcycle 3 = Car 4 = Motorcycle & car
X_5	Travel time to campus	Category	0 = <15 minutes 1 = 15-30 minutes 2 = 30-45 minutes 3 = 45-60 minutes 4 = >60 minutes

The research procedure can be broadly divided into the literature study stage, determining the number and distribution of samples, organizing data, and analyzing data. The first stage or preparation stage is a literature study to comprehend theories related to the goal of the research, then to determinethe number and distribution of samples in accordance with the research area. The second stage is data collection, and organizing required data obtained from surveys. Based on the source of data, data can be classified into primary data obtained directly from respondents or objects under study, and secondary data that have been previously collected and reported. In this research, secondary data was obtained from Mulawarman University. The third stage is data analysis to analyze all data that have been collected with descriptive analysis for each variable, followed by the binary logistic regression for the selection of the best logistic regression model. The fourth stage is a discussion of the results of both descriptive analysis and binary logistic regression. In the descriptive analysis, the collecteddata is presented in the form of a bar chart that displays the frequency of respondents' answers for each variable category. In the binary logistic regression analysis, the results of multicollinearity detection,

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 5: 2830-2835, 2025
DOI: 10.55214/25768484.v9i5.7600

simultaneous parameter significance test, partial parameter significance test, model fit test, and binary logistic regression model are presented. After the discussion, conclusions and suggestions are drawn as a closing.

3. Results and Discussion

3.1. Model Building

The formation of the model is obtained from the possible combination of k independent variables, so the number of models is 2k - 1, because in this research the number of independent variables is five, the relationship form is 25 - 1 = 31 possible models formed. The model will be tested and analyzed using SPSS software.

3.2. Model Fit Test

The model fit test is to conduct the test [21] in each model formed to determine the variables that have contributed significantly to the model. If the model I s significant, then proceed to the simultaneous test.

3.3. Concurrent Test

The simultaneous test is to test the significant effect of all predictor variables on the response variable using the G test in each appropriate model. If the model is significant, then proceed to the partial test.

3.4. Partial Test

Partial test is to test the significant effect of each independent variable on the dependent variable using the Wald test in each significant model in the simultaneous test. If in this test the model is significant, then the model is included in the best model classification.

3.5. Modeling Logistic Regression

After an analysis using SPSS *software*, the results are shown in Table 2 below.

Table 2.
Model testing results.

Variables	Description
X_1 , X_3 , X_5	Significant
X_1 , X_3 , X_4 , X_5	Significant
$X\,X_{1,2}$, X , X_{34} , X_5	Significant

Based on the G test and Wald test that have been applied to all models formed, several significant models are obtained and can be seen as the final model in Table 2 which includes all the independent variables in this study.

3.6. Determining the Best Logistic Regression Model

After obtaining several significant models presented in Table 2, the best model is selected based on the smallest G value.

Table 3 shows the best model based on the criteria of the smallest G value to determine factors that influence the use of public transportation by Mulawarman University students in Samarinda City is $(x) = \frac{\exp(0{,}087-1{,}643X_1+1{,}374X_2+1{,}146X_3-0{,}725X_4+1{,}593X_5)}{1+\exp(0{,}087-1{,}643X_1+1{,}374X_2+1{,}146X_3-0{,}725X_4+1{,}593X_5)}$ which includes variables X_1 , X_2 , X_3 , X_4 , X_5 with a G value of 471.648.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 5: 2830-2835, 2025
DOI: 10.55214/25768484.v9i5.7600

Table 3.
Best Model Classification.

Variables	Model	G
X_1 , X_3 , X_5	$\hat{\pi}(x) = \frac{\exp(0{,}605 - 1{,}311X_1 + 0{,}948X_3 + 1{,}397X_5)}{1 + \exp(0{,}605 - 1{,}311X_1 + 0{,}948X_3 + 1{,}397X_5)}$	484,743
X_1 , X_3 , X_4 , X_5	$\hat{\pi}(x) = \frac{\exp(1{,}323 - 1{,}466X_1 + 1{,}082X_3 - 0{,}582X_4 + 1{,}693X_5)}{1 + \exp(1{,}323 - 1{,}466X_1 + 1{,}082X_3 - 0{,}582X_4 + 1{,}693X_5)}$	476,236
$X\ X_{1,\ 2}$, X , X_{34} , X_5	$\hat{\pi}(x) = \frac{\exp(0{,}087 - 1{,}643X_1 + 1{,}374X_2 + 1{,}146X_3 - 0{,}725X_4 + 1{,}593X_5)}{1 + \exp(0{,}087 - 1{,}643X_1 + 1{,}374X_2 + 1{,}146X_3 - 0{,}725X_4 + 1{,}593X_5)}$	471,648

4. Conclusion

The results of the analysis using the binary logistic regression reveal variables that most influence the decision of Mulawarman University students to use public transportation. The variables are the distance from student residence to campus and the travel time. The best logistic model to describe factors that influence the use of public transportation by Mulawarman University students in Samarinda City is as follows

$$\hat{\pi}(x) = \frac{exp\,(0{,}087 - 1{,}643X_1 + 1{,}374X_2 + 1{,}146X_3 - 0{,}725X_4 + 1{,}593X_5)}{1 + exp\,(0{,}087 - 1{,}643X_1 + 1{,}374X_2 + 1{,}146X_3 - 0{,}725X_4 + 1{,}593X_5)}$$

Transparency:

The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

References

[1] E. Susanti and S. T. Nurul Hidayati, *Analysis of through traffic in Surakarta City from the West*. Indonesia: Universitas Muhammadiyah Surakarta, 2014.

[2] W. Natalia, H. Dwiatmoko, and N. Widyaningsih, "The effectiveness city of balikpapan-samarinda route selection with user characteristics analysis, and toll tariff analysis," *International Journal of Research and Review*, vol. 8, no. 8, pp. 533-545, 2021.

[3] N. Asiyah, S. Damanik, and R. Ramadhan, "Transportation mode choice to campus study in the new normal era (case study: Al-Azhar University)," *Jurnal VORTEKS*, vol. 5, no. 1, pp. 376-386, 2024.

[4] S. H. Aspan, "Knowledge, attitude, and practice towards social distancing among undergraduate students in medical fields during rapid rise of COVID-19 in Samarinda City, Indonesia: A cross sectional study," *Journal of Public Health*, pp. 1-10, 2020.

[5] C. L. B. Caling, P. V. Sandi, A. J. Fireflies, and M. Gondia, "The influence of analysis of choice of shuttle transportation modes in the province of the gunung mas travel fleet route ruteng-ende City," presented at the IOP Conference Series: Earth and Environmental Science, IOP Publishing, 2024.

[6] N. Widyaningsih, W. Mohtar, N. Yussof, and M. Putri, "Impact of large-scale social restrictions on transportation modes," *International Journal on Technical and Physical Problems of Engineering*, vol. 14, no. 51, pp. 216-221, 2022.

[7] S. Kotz, N. Balakrishnan, C. B. Read, B. Vidakovic, and N. L. Johnson, *Encyclopedia of statistical sciences*. United State: John Wiley & Sons, 2005.

[8] A. Malik and M. M. Chusni, *Introduction to educational statistics: Theory and applications*. India: Deepubish, 2018.

[9] Z. Lubis, *Applied statistics for social and economic sciences*. Indonesia: Andi Publisher, 2021.

[10] S. H. Udhana, *Application of logistic regression analysis model and correspondence analysis on self-healing method*. Indonesia: Makassar State University, 2016.

[11] R. A. Primasworo and M. Sadillah, "Identification and characteristics of urban transportation needs in Kraksaan in Probolinggo regency," *Cantilever: Jurnal Penelitian dan Kajian Bidang Teknik Sipil*, vol. 10, no. 2, pp. 101-110, 2021.

[12] S. F. Khairunnisa, S. Suharni, and D. A. Nohe, "Analysis of factors influencing human development index in West Java using binary logistic regression," in *Proceedings of National Seminar on Mathematics and Statistics*, 2022.

[13] H. Susilo, Y. Takahashi, and M. Yabe, "Evidence for mangrove restoration in the Mahakam Delta, Indonesia, based on households' willingness to pay," *Journal of Agricultural Science*, vol. 9, no. 3, pp. 30-41, 2017.
[14] A. Varamita, *Logistic regression analysis and its application to anemia disease for pregnant women at RSKD Ibu dan Anak Siti Fatimah Makassar*. Indonesia: FMIPA, 2017.
[15] R. Ananda and M. Fadhli, *Educational statistics: Theory and practice in education*. Indonesia: Deepublish, 2018.
[16] A. Nurkhin, A. Yulianto, and N. M. Daud, "Multinomial logistic regression for the analysis of career decision style in teacher education," *European Journal of Educational Research*, vol. 12, no. 1, pp. 1-10, 2023.
[17] T. A. Sari, "Application of backward method to determine the factor that affect the poverty level of Sibolga City 2011-2020," *JMEA: Journal of Mathematics Education and Application*, vol. 2, no. 1, pp. 10-19, 2023.
[18] A. A. R. Fernandes, *Metode statistika multivariat pemodelan persamaan struktural (sem) pendekatan warppls*. Indonesia: Universitas Brawijaya Press, 2017.
[19] M. Ramdhan, *Research methods*. Indonesia: Cipta Media Nusantara, 2021.
[20] J. Hogland, N. Billor, and N. Anderson, "Comparison of standard maximum likelihood classification and polytomous logistic regression used in remote sensing," *European Journal of Remote Sensing*, vol. 46, no. 1, pp. 623-640, 2013.
[21] N.-N. Sze and S. Wong, "Diagnostic analysis of the logistic model for pedestrian injury severity in traffic crashes," *Accident Analysis & Prevention*, vol. 39, no. 6, pp. 1267-1278, 2007.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 5: 2830-2835, 2025
DOI: 10.55214/25768484.v9i5.7600