American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-6, Issue-11, pp-295-300
www.ajer.org
Research Paper Open Access

Model Of Movement Awakening To Workplace By Using Binary Logistic Regression Analysis (Case Study In Sawojajar Urban Village Malang City)

Halimi Firdausi A.[1], Sobri Abusini M.T[2], Agus Widodo[3]
[1,2,3]*Department of Mathematics, PostGraduate Faculty MIPA, Brawijaya University, Malang, Indonesia*
Corresponding Author: fa.mimie@yahoo.co.id

ABSTRACT: *Transport plays an important role in the development of an urban area. Transportation if not set well will cause some problems one of which is jamming. Malang City especially Sawojajar urban village is one city that has high population level. High population levels lead to an increase in the rise of movement. Movement with the goal of working is the main movement that is done by every person every day. The number of daily movements by workers increases the volume of traffic that can lead to congestion. Therefore, it is necessary to analyze what factors influence the rise of movement to the workplace. Many methods are used to analyze the rise of movement, one of which is regression analysis. Analytical methods used are binary logistic regression method with independent variables ie job type, average income, number of motorcycles, the number of cars, the type of transportation used to go to work, the number of family members, the number of family members working, the distance home with the workplace, the cost of travel to the workplace, the number of family members aged 0-11 years, the number of family members aged 12-25 years, the number of family members aged 26-45 years, the number of family members aged 46-60 years, and number of families over 60 years old. After analyzing with binary logistic regression, got four factors that influence the rise of movement to in Sawojajar urban village. The four factors are the type of work, the number of motors, the type of vehicle used to go to the workplace and the distance home with the workplace.*

Keywords: *Transportation, Movement Generation, Binary Logistic Regression*

Date of Submission: 18-11-2017 Date of acceptance: 28-11-2017

I. INTRODUCTION

The more people who live in a city, then certainly have transportation problems, therefore it takes the concept of transportation planning. Some of the most popular transportation planning concepts are the four-stage transportation planning model. This planning model is a composite of several sub models each of which must be done separately and in successive sub models are the rise and pull of movement, movement distribution, mode selection and route selection (Ortuzar et al., 2001).

Malang City especially Sawoajajar village is one of the areas that have high population level. High population levels lead to an increase in the rise of movement. The rise of movements is classified into two, namely home-based awakening and non-home-based awakening. Home-based awakening based on movement objectives is classified into four: movement toward work, movement toward shopping, movement towards school and movement towards recreation, social, etc. (Tamin, 2000). Movement with the goal of working is the main movement that is done by every person every day. The number of daily movements by workers increases the volume of traffic that can lead to congestion. The rise of movement to the workplace is certainly influenced by several factors, such as socio-economic and land use.

Many methods are used to analyze the rise of movement, one of which is regression analysis. In the study (Mccarthy, 1969) analyzed the relationship between the rise of household-based travel with socio-economic characteristics analyzed using multiple linear regression. There are several studies that also analyze the relationship between the rise of home-based movements with socio-economics such as in the studies (Oyedepo and Makindie, 2009), (Al-Mustansiriyah et al., 2012), (Aloc and Amar, 2013) and (Dawra and Kulshreshtha, 2017).

In the study (Purvis et al., 1996), it discusses the relationship between the rise of movement to the workplace and the rise of movement that is not the goal of the workplace in the city of San Francisco using multiple linear regressions. The results show an inverse proportion between the rise of movement to the workplace and the rise of movement not to the workplace, if the rise to the workplace increases, whereas for non-work movement generation decreases.In the study (Huntsinger et al., 2013), it discusses cumulative (ordinal) logistic regression to estimate the rise of movement to the workplace and the rise of movements in addition. To estimate the rise of movement, the independent variables used are life cycle, area type and accessibility. In the study (Chang et al., 2013), comparing six models of regression model, tobit model, poisson model, logit model and multi-classification analysis applied to the rise of movement to workplace in Seoul City.

In this study discusses the model of the rise of movement to the workplace in Sawojajar Urban Malang City with binary logistic regression analysis. Binary logistic regression is a method of data analysis used to find the relationship between dependent variable (Y) which is binary or dichotomous with variable independent (X) which is polikotomus (Hosmer and Lemeshow, 2000). The trip writing data used are household-based trip generation data with dependent variable (Y) ie person who travels to workplace (1) and person not traveling to workplace (0). For independent variables are job type, average income, number of motorcycles, number of cars, number of family members, number of family members working, house distance with work rate, travel cost to work place, number of family member 0- 11 years, number of family members aged 12-25 years, Number of family members aged 26-45 years, number of family members aged 25-60 years, number of family members aged 26-60 years, and number of families aged> 60 year. The purpose of this research is to know which model and factors influencing the rise of movement to work place. Data processing in this study using SPSS software.

II. METHOD

In this study, data obtained from direct observation and distribution of questionnaires to a number of families with the number of respondents as much as 115. The data taken are as follows:

1. Job types (X_1) are: civil servant, self-employed, private, professional, in addition to all four.
2. Average family income (X_2), ie: <Rp 500,000, 2,000,000-Rp 3,000,000, Rp 500,000-Rp 2,000,000 and> Rp 3,000,000.
3. Number of motors (X_3).
4. Number of cars (X_4).
5. Vehicles used to go to work (X_5), namely: Personal vehicles, public transport, car pickup and besides the three.
6. Number of family members (X_6).
7. Number of family members working (X_7).
8. Distance of house with work place (X_8), that is: 0-3 Km, 8-12 Km, 4-7 Km and> 13 Km.
9. Traveling costs to work (X_9), with categories: Rp 0-Rp 5.000, Rp 6,000-Rp 10,000, Rp 11,000-Rp 15,000 and> Rp 15,000.
10. Number of family members aged 0-10 years (X_{10}).
11. Number of family members aged 11-25 years (X_{11}).
12. Number of family members aged 26-45 years (X_{12}).
13. Number of family members aged 46-60 years (X_{13}).
14. Number of family members> 60 years old (X_{14}).

The above data were analyzed by binary logistic regression. Stages of logistic regression analysis are multicollinearity test, parameter estimation, parameter testing simultaneously, partial parameter testing, checking suitability of model formed, applying binary logistic regression model obtained from training data to data testing, forming table of classification accuracy of training and testing model and determine the Apparent error rate.

III. RESULTS AND DISCUSSION

3.1 Test multicollinearity

Table 1. VIF Values for Each Predictor Variable

No.	Predictor Modifiers	VIF values
1.	Type of work	1,452427
2.	Amount of revenue per month	1,396953
3.	Motorbikes	1,082824
4.	Cars	1,001165
5.	The type of vehicle used to go to work	1,421814
6.	Number of families	1,048303
7.	Number of families working	1,006855

8.	Distance of the house to the workplace	1,529285
9.	The cost of travel to work	1,492383
10.	Number of families aged 0-11 years	1,027741
11.	Number of families aged 12-25 years	1,019034
12.	Number of families aged 26-45 years	1,005165
13.	Number of families aged 46-60 years	1,001306
14.	Number of families over 60 years old	1,040037

Based on Table 1. it can be seen that the VIF value for each predictor variable is less than 10, so there is no multicollinearity for each predictor variable.

3.2 Parameter Probe

The parameter estimator on binary logistic regression uses Maximum Likelihood Estimation (MLE). Here is the value of each parameter of indpendent variable.

Table 2. Binary Logistic Regression the Parameters of Estimator

Variables in the Equation						
		B	S.E.	Wald	df	Sig.
Step 1[a]	X1			8,659	4	,070
	X1(1)	15,936	10677,003	,000	1	,999
	X1(2)	-4,285	1,566	7,489	1	,006
	X1(3)	-3,432	1,300	6,970	1	,008
	X1(4)	-3,623	1,675	4,676	1	,031
	X2			2,382	3	,497
	X2(1)	-2,015	1,706	1,395	1	,237
	X2(2)	-1,112	1,698	,429	1	,513
	X2(3)	-1,676	1,787	,880	1	,348
	X3	,693	,461	2,266	1	,132
	X4	-,202	,700	,083	1	,773
	X5			7,638	3	,054
	X5(1)	1,741	1,164	2,238	1	,135
	X5(2)	3,879	1,410	7,569	1	,006
	X5(3)	3,558	4,777	,555	1	,456
	X6	1,340	,812	2,724	1	,099
	X7	,308	,589	,272	1	,602
	X8			,235	3	,972
	X8(1)	-24,080	14493,519	,000	1	,999
	X8(2)	-,827	1,976	,175	1	,676
	X8(3)	-,203	1,506	,018	1	,893
	X9			3,668	3	,300
	X9(1)	1,065	31281,774	,000	1	1,000
	X9(2)	-17,261	27721,608	,000	1	1,000
	X9(3)	-19,961	27721,608	,000	1	,999
	X10	-1,547	,914	2,865	1	,091
	X11	-1,207	,941	1,647	1	,199
	X12	-1,951	,984	3,933	1	,047
	X13	-2,273	,976	5,424	1	,020
	X14	-1,965	1,227	2,566	1	,109
	Constant	24,707	27721,608	,000	1	,999

The full model of the binary logistic regression model based on the column value B in Table 2. ie:

$$g(x) = \ln\left(\frac{\exp(24{,}707 + 5{,}936x_{1(1)} - 4{,}285x_{1(2)} - 3{,}432x_{1(3)} - 3{,}623x_{1(4)} - 2{,}015x_{2(1)}}{1 + \exp(24{,}707 + 5{,}936x_{1(1)} - 4{,}285x_{1(2)} - 3{,}432x_{1(3)} - 3{,}623x_{1(4)} - 2{,}015x_{2(1)}}\right.$$

$$\frac{-1{,}112x_{2(2)} - 1{,}676x_{2(3)} + 0{,}831x_3 - 0{,}246x_4 + 1{,}741x_{5(1)} + 3{,}879x_{5(2)} + 3{,}558x_{5(3)}}{-1{,}112x_{2(2)} - 1{,}676x_{2(3)} + 0{,}831x_3 - 0{,}246x_4 + 1{,}741x_{5(1)} + 3{,}879x_{5(2)} + 3{,}558x_{5(3)}}$$

$$\frac{+1{,}340x_6 + 0{,}308x_7 - 24{,}080x_{8(1)} - 0{,}827x_{8(2)} - 0{,}203x_{8(3)} + 1{,}065x_{9(1)} - 17{,}261x_{9(2)}}{+1{,}340x_6 + 0{,}308x_7 - 24{,}080x_{8(1)} - 0{,}827x_{8(2)} - 0{,}203x_{8(3)} + 1{,}065x_{9(1)} - 17{,}261x_{9(2)}}$$

$$\left.\frac{-19{,}961x_{9(3)} - 1{,}547x_{10} - 1{,}207x_{11} - 1{,}951x_{12} - 2{,}273x_{13} - 1{,}965x_{14})}{-19{,}961x_{9(3)} - 1{,}547x_{10} - 1{,}207x_{11} - 1{,}951x_{12} - 2{,}273x_{13} - 1{,}965x_{14})}\right).$$

3.3 Testing Parameters Simultaneosly

Testing parameters simultaneously using G (Likelihood Ratio Test) test statistic, with hypothesis:
$H_0 : \beta_0 = \beta_1 = \beta_2 = \cdots = \beta_k = 0$,the predictor variables do not affect the response variable.
H_1 : there is at least one β_k ≠ 0, at least one predictor variables affecting the response variable.

Table 3. Results of Testing Concurrently

Omnibus Tests of Model Coefficients				
		Chi-square	df	Sig.
Step 1	Step	67,526	25	,000
	Block	67,526	25	,000
	Model	67,526	25	,000

Based on the results in Table 3, the value of G can be seen in the chi-square column model in Table 3, of 61.709, so the value of G>$\lambda^2_{(20;0,05)}$(31,411). Thus, it can be concluded that there is at least one predictor variable that influences the response variable.

3.4 Parameter Test Partially

Partial parameter test using Wald statistic test, with hypothesis:
$H_0 : \beta_j = 0$,the predictor variable *j* has no effect on the response variable
H_1: there is at least one $\beta_j \neq 0$,, the *j* rammer of the predictor influences the response variable.
Table 4 is the result of the full model, to get the best model then used backward elimination method.

Table 4. Parameter Testing Partially

Variables in the Equation						
		B	S.E.	Wald	df	Sig.
Step 1[a]	X1			8,564	4	,073
	X1(1)	16,845	12336,343	,000	1	,999
	X1(2)	-3,308	1,191	7,718	1	,005
	X1(3)	-2,295	,908	6,379	1	,012
	X1(4)	-2,431	1,231	3,897	1	,048
	X3	,619	,313	3,917	1	,048
	X5			8,467	3	,037
	X5(1)	1,210	,677	3,194	1	,074
	X5(2)	2,703	,981	7,595	1	,006
	X5(3)	,383	1,766	,047	1	,828
	X8			16,873	3	,001
	X8(1)	-3,086	1,405	4,823	1	,028
	X8(2)	,715	1,366	,274	1	,600
	X8(3)	,453	1,328	,116	1	,733
	Constant	2,154	1,591	1,833	1	,176

Based on Table 4, there are four predictor variables that affect the response variable because the value of *p-value*<0.05, so reject H_0. Binary logistic regression model as follows:

$$g(x) = \ln\left(\frac{\exp(2{,}154 + 16{,}845x_{1(1)} - 3{,}308x_{1(2)} - 2{,}295x_{1(3)} - 2{,}431x_{1(4)} + 0{,}383X_3 + 1{,}210X_{5(1)} + 2{,}703X_{5(2)} + 0{,}383X_{5(3)} - 3{,}086X_{8(1)} + 0{,}715X_{8(2)} + 0{,}453X_{8(3)})}{1 + \exp(2{,}154 + 16{,}845X_{1(1)} - 3{,}308X_{1(2)} - 2{,}295X_{1(3)} - 2{,}431X_{1(4)} + 0{,}383X_3 + 1{,}210X_{5(1)} + 2{,}703X_{5(2)} + 0{,}383X_{5(3)} - 3{,}086X_{8(1)} + 0{,}715X_{8(2)} + 0{,}453X_{8(3)})}\right).$$

3.5 Model Conformity Test

The statistical test used is good of fit test. The results of the model conformity test are presented in Table 5.

Table 5. Model Conformity Test

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	8,695	8	,369

Based on Table 5, we get the sig value (p-value) of 0.369. Because p-value value> α(0,05), then accept H_0. Thus, it can be concluded that the binary logistic regression model that is formed is appropriate.

3.6 Model Interpretation

In this study the category used as a reference is not to travel. The odds ratio for each predictor variable can be seen in Table 6. in the *Exp (B)* column. Here's a brief presentation.

Tabel 6.*Odds Ratio*

Variables	Category	*P-value*	*Exp(B)*
Type of work (X_1)	Entrepreneur	0,005	0,037
	Professional	0,012	0,101
	In addition to the four (civil servants, private, self-employed, professional)	0,048	0,088
Number of motorbikes(X_3)		0,048	0,088
The type of vehicle used to go to work(X_5)	Public transportation	0,006	14,929
Distance of the house to the workplace(X_8)	3-5 Km	0,028	0,046

Here is the odds ratio interpretation:

1. Citizens whose occupations are self-employed have a tendency of 0.037 times to travel to work compared to not traveling on the job.
2. Citizens who type professional jobs have a greater tendency of 0.101 times to travel to the workplace than not traveling to the workplace.
3. Citizens whose occupations other than civil servants, private, entrepreneurs and professionals have a greater tendency of 0.088 times to travel than not travel.
4. Residents who own motorcycles tend to have bigger 0.088 times compared to not traveling.
5. People who choose public transport to go to work have a tendency 14,929 times compared to not traveling to work place.
6. Residents who are 3-5 km distance between homes and workplaces have a tendency of 0.046 times to travel to places compared to not traveling to the workplace.

3.7 Accuracy of Classification Results

The results of the precision of binary logistic regression logistic regression can be seen in Table 7.

Table 7. Accuracy of Classification Results

		Prediction		
		Y		Accuracy of Classification (%)
		Not traveling	Travelling	
Y	Not Traveling	22	12	64,7%
	Travelling	7	74	91,4%
Accuracy of the Entire Classification				83,5%

Based on Table 7 the overall classification accuracy result is 83.5%. This shows that the result of binary logistic regression analysis that can be formulated can classify the observation with exact equal to 83,5%.

IV. CONCLUSION

Based on the results of analysis and discussion that has been done, then the conclusions obtained are:

1. Of the fourteen predictor variables, four predictor variables were found that significantly affected the response variable. Four predictor variables are the type of work, the number of motorcycles, the type of vehicle used to go to the workplace and the distance home with the workplace.
2. Binary logistic regression model that formed is:

$$g(x) = \ln\left(\frac{\exp(2{,}154 + 16{,}845x_{1(1)} - 3{,}308x_{1(2)} - 2{,}295x_{1(3)} - 2{,}431x_{1(4)} + 0{,}383X_3}{1 + \exp(2{,}154 + 16{,}845X_{1(1)} - 3{,}308X_{1(2)} - 2{,}295X_{1(3)} - 2{,}431X_{1(4)} + 0{,}383X_3}\right.$$

$$\frac{+1{,}210X_{5(1)} + 2{,}703X_{5(2)} + 0{,}383X_{5(3)} - 3{,}086X_{8(1)} + 0{,}715X_{8(2)} + 0{,}453X_{8(3)})}{+1{,}210X_{5(1)} + 2{,}703X_{5(2)} + 0{,}383X_{5(3)} - 3{,}086X_{8(1)} + 0{,}715X_{8(2)} + 0{,}453X_{8(3)})}\Bigg).$$

REFERENCES

[1]. Al-Mustansiriyah, G.S., et al. 2012. Trip Generation Modeling for Selected Zone in Al-Diwaniyah City. *Journal of Enginering and Development* **16 (4)**: 1813-7822.

[2]. Dawra, M and Kulshreshtha, S.. 2017. Developing Trip Generation Model Utilizing Multiple Regresson Analisysis. *Internasional Journal of Innovative Research in Science, Engineering and Technology*.

[3]. GM Mccarthy, 1969. Multiple Regression Analysis Of Household Trip Generatiaon- A Critique.

[4]. Hosmer, D.W. and Lemeshow, S. 2000. Applied Regression. John Willey and Sons, Inc. New York.

[5]. Huntsinger, L.F., et al. 2013. Trip Generation Models Using Cumulative Logistic Regression. *Urban Planning and Development* **139:** 176-184.

[6]. L. Purvis, C., Iglesias, M. and A. Eisen, V.. 1996. Incorporating Work Trip Generaton Accessibility in Nonwork Trip Genertaion Models in San Fransisco Bay Area.

[7]. Ortuzar, J.D., Willumsen, L.G. 2001. Modelling Transport. Wileyand Sons Inc, New York.

[8]. Oyedepo, O.J andMakindie, O.O..2009. Regression Model of Household Trip Generationof Edo-Ekiti Township in Nigeria. *Europan Journal of Scientific Research* **28 (1)** : 132-142.

[9]. S. Aloc, D. and C. Amar, J.A. 2013. Trip Generation Modelling of Lipa City. *Seminar and Research Method in Civil Enginering* **199**.

[10]. S. Chang, J., et al. 2014. ComparativeAnalysis of Trip Generation Models: Results Using Home-Based Work Trips in The Seosul Metropolitan area.

[11]. Tamin, O.Z. 2000. Perencanaan dan Pemodelan Transpostasi. Teknik Sipil Institut Teknologi Bandung.

Halimi Firdausi A Model Of Movement Awakening To Workplace By Using Binary Logistic Regression Analysis (Case Study In Sawojajar Urban Village Malang City)." American Journal of Engineering Research (AJER), vol. 6, no. 11, 2017, pp. 295-300