CHAPTER III

RESEARCH METHODOLOGY

A. Research Location

For research location, this research was conducted in Way Kambas National Park East Lampung Regency of Lampung Province. The area has been developed as a tourist attraction namely Elephant Conservation Center (PKG). This research was conducted during one month, that is February 2017.

B. Type of Data

The type of data used in this research is primary and secondary data. The primary data used in this research was obtained from the result of interviews with participants through a questionnaire. Primary data includes responses from respondents to the collection of the levy for environmental improving in Way Kambas National Park as well as the nominal value of visitors' willingness to pay for the program of development.

Secondary data is the data are obtained by gaining information from source of literature research by reading, observation, and studying the written materials such as books, magazine, articles, journal, thesis, and other writing information related to this research (Basuki, 2013). Secondary data or available because data is obtained and collected from the source that already exist.

C. Data source

According to the data source, the data used in this research is primary data which is data obtained directly from the results of interviews with the number of respondents which were sampled using the questionnaire that has been prepared.

D. Population and Sample

1. Population

According to Sugiyono (2003) population is the generalization region which consists of objects or subjects which have certain qualities and characteristics defined by the researchers to be studied and then drawn the conclusions. The population covers all the characteristics or properties owned by the subject or the object. In this research, the population was tourists who visit Way Kambas National Park.

2. Sample

The sample is a fraction of the number and characteristics owned by the population. The sample of respondents is determined by accidental sampling, which is a convenient sampling method done by selecting a free sample, according to a researcher, where respondents are readily available or accessible as a sample while maintaining the feasibility and accuracy of the selected samples.

The sample in this research is by using the formula of Slovin namely:

$$n = \frac{N}{1 + Ne^2}$$

Where:

n = Sample

N = Population

e = Standard error 0,10 (10%)

$$n = \frac{N}{1+N\ (d^2)}$$

$$= \frac{23.798}{1\ +23.798\ (0{,}1^2)} = \text{ 99,58 or 100.}$$

So, the sample which is used in this research is 100 visitors.

E. Operational Definition of Variables

1. Research variable

The variables that are used in this research were dependent and independent variable. The dependent variable is factors that are influenced by the independent variable. Meanwhile, Independent variable is factors that influence the dependent variable. The dependent variable that is used in this research is the visitors' willingness to pay (WTP), while independent variables consist of income, travel cost, education level, and age that affect WTP.

2. Operational Definition of Variables

Operational definition of variables explains the way that was used by a researcher to measure variables that would be used. A lot of variables mentioned above that were used in this research are as follows:

a) Willingness To Pay (WTP)

Willingness To Pay is an assessment method that aims to determine at what level a person is able to pay for the environmental improvement if wants to be a good environment. Willingness to pay has a relationship with income level, someone who has a high income may have a tendency to have a willingness to pay higher than people who have lower income. The estimation of the average of willingness to pay is defined as follows:

$$\text{EWTP} = \frac{\sum_{t-1}^{n} WTP_i}{n}$$

Wheres :

EWTP = the average estimation of WTP

WTP_i = the value willingness to pay

n = Total respondent

i = Respondent-i who are willing to pay (i=1,2,3,…n)

EWTP value is used to determine the respondents' willingness to pay with Dichotomous Choice. The value of dummy variable WTP is 1 if "WTP = EWTP" and 0 if "WTP ≠ EWTP".

b) Travel cost is all the cost that is expensed by tourists for travel cost to the park. This variable was measured by using a scale continuously in Rupiah including the cost of transportation, entrance fees, accommodation cost, consumption, and others.

1) The transportation cost is the cost of transportation to and from Way Kambas National Park in Rupiah.

2) The cost of accommodation is lodging expenses at the WKNP in Rupiah.
3) The cost of consumption is the cost of consumption spent for the WKNP in Rupiah.
4) The cost of admission is the admission fee or entrance fees paid in that time at WKNP in Rupiah.

c) The visitor's income per month is the total income received by individuals who have been working and earning for one month in Rupiah. For the respondent as students, their income is earned pocket money per month.

d) Visitors' Education is the level of education from each respondent measured by the length of the study of visitors.
1) SD (Elementary School) = 6
2) SMP graduate (Junior High School) = 9
3) SMA graduate (Senior High School) = 12
4) Diploma graduate = 15
5) Bachelor degree graduate = 16

e) Visitors' Age is an age of respondents which is expressed in years.

F. Data Analysis Method

1. Logit Regression Analysis

Logit regression analysis is part of regression analysis. This analysis examines the relationship between the effects of the explanatory variables (X) toward response variable (Y). However, if the response variable from

the regression analysis is categorical, the regression analysis that is used is logit regression analysis (Hosmer and lemeshow 1989).

Logistic regression uses dichotomous on the dependent variable (yes or no). The data in the data processing no longer requires a normality test and classical assumption on the independent variables. The advantage of logistic variable is more flexible than other techniques, such as in logistic regression which does not have the normality assumption on the independent variables used in the model. It means that independent variable do not have to have a linear normal distribution and also have the same variant of each group. Besides, the independent variables in logistic regression can be a mixture of continuous variables, discrete and dichotomy.

Logistic regression has also ignored the problem of heteroscedasticity, it means that the independent's variables do not require homoscedasticity on the independent variables (Gujarati dan Porter, 2012).

This research uses binary logistic regression analysis in which the respondents only have two choices whether the respondents are willing to pay or not. The equation model in this research is:

$$\text{WTP} = \beta_0 + \beta_1 Inc_i + \beta_2 TC_i + \beta_3 Edu_i + \beta_4 Age_i$$

Where :

WTP : Willingness to pay

β0 : The constant

β1 : The regression coefficient Income

Inc : Income (Rupiah)

β2 : The regression coefficient Travel Cost

TC : Travel Cost (Rupiah)

β3 : The regression coefficient Education

Edu : Education (Years)

β4 : The regression coefficient Age (Years)

2. Classification Accurate Test

Classification accurate test was done in order to predict the accuracy of the model in classifying observations expressed in percentage. Assume that the large percentage means the more accuracy of a model to classifying the observation.

3. Goodness of Fit test

a. Hosmer and Lemeshow Test

Hosmer and Lemeshow test was conducted to test whether the empirical data is in accordance with the model, and the model can be said to be fit. The hypothesis is:

H_0 = model is able to predict the value of observation

H_1 = model is unable to predict the value of observation

The test result is:

H_0 accepted if sig > 0,05, the model is able to predict the value of observation.

H_1 accepted if sig < 0,05, the model is unable to predict the value of observation.

4. Determination Coefficient Test

Nagelkerge R square is similar to R-Square in linear regression analysis that explains the percentage of fit model or value which show how the independent variables can explain the dependent variable (Basuki, 2015). Negelkerge R square in logistic regression is used as determination coefficient test.

5. Signification Test

a. Simultant Test (Overall Test)

Simultant Significant Test was conducted to determine the effect of all the independent variables altogether toward dependent variable. The hypothesis is:

H_0 = All of the independent variables altogether have no significant influence on the dependent variable.

H_1 = All of the independent variables altogether have significant influence on the dependent variable.

The test result is :

H_0 accepted if sig > 0,05, so all the independent variables altogether do not influence on dependent variable.

H_1 accepted if sig < 0,05, so all the independent variables altogether influence dependent variable.

b. Partial Test

The partial test is done to estimate the effect of independent variables which is partial toward dependent variable. The Hypothesis is :

H_0 = Independent variables have no significant influence on dependent variable.

H_1 = Independent variables have significant influence on dependent variable.

The test result is:

H_0 accepted if sig > 0,05, so the independent variables have no influence on dependent variable.

H_1 accepted if sig < 0,05, so the independent variables influence dependent variable.