

## **CHAPTER III**

### **RESEARCH METHODS**

#### **A. Object and Research Subjects**

The object of this study is on e-commerce sites while the subject of this study is millennial generation in the Special Region of Yogyakarta. The location of the study was conducted at eight universities in Yogyakarta which had been accredited A, because higher education people such as students are required to be broad-minded and this will be more easily achieved by meeting information needs quickly and efficiently using information technology devices (Utama, 2015), such as Universitas Gadjah Mada, Universitas Negeri Yogyakarta, Universitas Pembangunan Nasional Veteran, Universitas Islam Negeri Sunan Kalijaga, Universitas Muhammadiyah Yogyakarta, Universitas Islam Indonesia, Universitas Ahmad Dahlan and Universitas Atma Jaya.

#### **B. Data Types**

The data used in this study are primary data. Primary data is data obtained with field survey techniques that use all data collection techniques (Kuncoro, 2003). Primary data in this study were obtained by giving a structured statement questionnaire and calculated using a Likert scale distributed directly to millennials in the Special Region of Yogyakarta. In this study the data used are trust, price, promotion, time, risk, ease of use and quality of information on the selection of e-commerce sites by millennial generation.

### C. Sampling Technique

The sample is part of the population studied in a study and the results are considered to be representative of the entire population. Meanwhile, according to Sugiyono (2016), the population is a generalization area consisting of objects and subjects that have certain qualities and characteristics determined by researchers to be studied and then conclusions drawn. The population in this study are students who study at Yogyakarta University who have been accredited A, that is Universitas Gadjah Mada, Universitas Negeri Yogyakarta, Universitas Pembangunan Nasional Veteran, Universitas Islam Negeri Sunan Kalijaga, Universitas Muhammadiyah Yogyakarta, Universitas Islam Indonesia, Universitas Ahmad Dahlan, and Universitas Atma Jaya, therefore this study uses a nonprobability sampling technique with the type of Purposive Sampling. According to Sugiyono (2010), purposive sampling is a technique for determining research samples with certain considerations aimed at making the data obtained later more representative. In this study the criteria used as a sample are:

1. Respondents once did online shopping on an e-commerce site at least 1 time.
2. Respondents are students from one of the Universitas Gadjah Mada, Universitas Negeri Yogyakarta, Universitas Pembangunan Nasional Veteran, Universitas Islam Negeri Sunan Kalijaga, Universitas Muhammadiyah Yogyakarta, Universitas Islam Indonesia, Universitas Ahmad Dahlan or Universitas Atma Jaya.
3. Respondents aged 18-25 years, because they belong to the millennial generation and can make their own decisions.

This research uses the formula of the Slovin method (in Ayuningtyas and Gunawan 2018) as follows:

$$\eta = \frac{N}{1 + N e^2}$$

where:

n: number of samples

N: total population

e: error tolerance limit (error tolerance)

In this study the existing population of 297,916, the samples used are:

$$\eta = \frac{297.916}{1 + 297.916(0.1)^2} = 99,96 \text{ rounded to } 100.$$

So this research will use a minimum of 100 respondents. So the researchers use 200 respondents.

#### **D. Data Collection Techniques**

The method used in this study is asking questions (questionnaire). The questionnaire is a data collection tool in the form of a series of questions asked to respondents to get answers. This research uses a Likert scale as a measurement tool to measure every question given to respondents, namely strongly disagree (STS) score 1, disagree (TS) score 2, neutral (N) score 3, agree (S) score 4, and very agree (SS) score 5.

#### **E. Operational Definitions of Research Variables**

A research variable is an attribute or value or nature of people, objects or activities that have certain variations determined by researchers to be studied and then drawn in conclusions (Sugiyono 2011). This operational definition refers to the variable that you want to study, the object of research in question is the e-commerce site. This study consisted of two variables, namely the

independent variable and the dependent variable. The following is an explanation of the variables:

1. The independent variable is the variable that influences or causes the change or the occurrence of the dependent variable (Sugiyono 2011). The independent variables of this research are trust, price, promotion, time, risk, ease of use and quality of information.
  - a. Trust  
Millennial generation trust is given to other parties that can be trusted.
  - b. Price  
The amount of money that must be spent by millennials to get the goods they want.
  - c. Promotion  
That companies use to tell or offer products or services so that millennials are interested in consuming them.
  - d. Time  
How long does an e-commerce transaction take time.
  - e. Risk  
Potential millennial generation losses did in online shopping.
  - f. Ease of use  
Easily millennial generation transacts through e-commerce sites.
  - g. Quality of information  
The information presented is clear, complete, up-to-date and easily understood by the millennial generation.
2. The dependent variable is a variable that is affected because of the independent variable (Sugiyono, 2011). In this study, the dependent variable is the selection of e-commerce sites. It means, millennial generation have plan to buy a particular product that it wants/needs on e-commerce that their choose.

#### **F. Quality Test and Data Instruments**

This research uses primary data using questionnaire data, so it is necessary to conduct a validity and reliability test:

1. Validity Test

Validity Test is the level of reliability of the measuring instruments used, the extent to which researchers can measure what you want measuring, to know whether the validity or invalidity of a questionnaire to be tested. If the calculation result turns out to be  $r_{\text{count}} \geq r_{\text{table}}$  then the instrument is considered valid, conversely if the calculation results are  $r_{\text{count}} < r_{\text{table}}$  then the instrument is considered invalid (Ghozali, 2006).

## 2. Reliability Test

Reliability Test is a measure of the consistency of values achieved by people at different times but has the same results, where the results obtained can be trusted. According to Ghozali (in Ayuningtyas and Gunawan 2018), that the statement that has been declared valid in the validity test will be determined by the following criteria:

- a. If the Cronbach Alpha value  $> 0.6$  then the research questionnaire is declared reliable (very good/very convincing).
- b. If the Cronbach Alpha value  $< 0.6$  then the research questionnaire is declared unreliable (less convincing).

## G. Classic Assumption Test

The classic assumption test is performed to determine the feasibility of the regression model used in this study. This test is done so that there is no multicollinearity and heteroscedasticity so that the data generated are normal (Ghozali, 2006).

### 1. Normality Test

Normality test is carried out to look at the variable regression model, confounding or residual variables have a normal distribution. Normality test is carried out using the Kolmogorov-Smirnov approach to see whether the data is normally distributed or not. Residual variables are normally

distributed if Kolmogorov-Smirnov significance value  $> 0.05$  and vice versa (Ghozali, 2006).

#### 2. Multicollinearity Test

Multicollinearity test was conducted to look at the regression model found a correlation between independent variables in the regression model. A good regression model should be free of multicollinearity. To find out the presence or absence of multicollinearity symptoms can be seen from the value of variance inflation factor (VIF), the criterion is if the tolerance value  $> 0.1$  or VIF value  $< 10$  then multicollinearity does not occur, vice versa if the tolerance value  $< 0.1$  or VIF value  $> 10$  will occur multicollinearity (Ghozali, 2006).

#### 3. Heteroscedasticity Test

Heteroscedasticity test is performed to see whether in a regression model there is an unequal variance from the residuals of one observation to another. The way to detect the presence or absence of heteroscedasticity is to look at the significant numbers that exist in the new regression equation greater than 0.05 then it is said that heteroscedasticity does not occur (Ghozali, 2006).

### **H. Hypothesis Testing and Data Analysis**

#### 1. Descriptive Statistic

According to Sugiyono (in Ayuningtiyas and Gunawan 2018) descriptive statistics are statistics used to analyze data by describing or describing the data that has been collected as it is without intending to make conclusions that apply to the public or generalizations.

#### 2. Multiple Linear Regression Analysis

This multiple linear regression method is used to see the relationship between the dependent variables namely trust (X1), price (X2), promotion

(X3), time (X4), risk (X5), ease of use (X6), and quality of information (X7) with an independent variable namely selection of e-commerce sites, the equation model used is:

$$Y_t = a + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \beta_5 X_{5t} + \beta_6 X_{6t} + \beta_7 X_{7t} + e_t$$

Information:

Y = Selection of E-Commerce Sites

a = Coefficient

$\beta_1$  = Regression coefficient for trust variable (X1)

$\beta_2$  = Regression coefficient for price variable (X2)

$\beta_3$  = Regression coefficient for promotion variable (X3)

$\beta_4$  = Regression coefficient for time variable (X4)

$\beta_5$  = Regression coefficient for risk variable (X5)

$\beta_6$  = Regression coefficient for ease of use variable (X6)

$\beta_7$  = Regression coefficient for the variable quality of information (X7)

X1 = Trust

X2 = Price

X3 = Promotion

X4 = Time

X5 = Risk

X6 = Ease of use

X7 = Quality of information

$e_t$  = error term

a. Partial Test (T Test)

Partial test is carried out to see how far the influence of independent variables is promotion, price, income, time, risk, ease and quality of information partially in explaining the dependent variable, namely selection of e-commerce sites by millennial generation. By comparing T arithmetic and T table if T arithmetic is greater than T table then the variable has a positive relationship and vice versa if the probability value of  $t < 0.05$ , it can be seen that the independent variable has a significant effect on the dependent variable partially and vice versa if  $\text{probability} > 0.05$ , it can be seen that the independent variable does not have a significant effect on the dependent variable (Ghozali, 2011).

b. Simultaneous Influence Test (F Test)

Simultaneous influence test is done to see whether it has a significant influence of the independent variables together on the dependent variable. If the probability value  $< 0.05$ , then the independent variable simultaneously has a significant effect on the dependent variable and vice versa if the probability value  $> 0.05$ , then the independent variable simultaneously has no significant effect on the dependent variable (Ghozali, 2011).

c. Determination Coefficient Test ( $R^2$ )

Test the coefficient of determination to determine what percentage of the total variation in the dependent variable is explained in the Independent variable. The coefficient of determination lies between 0 and 1 ( $0 \leq R^2 \leq 1$ ), if  $R^2 = 1$ , it means that the independent



variables provide almost all the information needed to predict the variation of the dependent variable (Ghozali, 2011).

