

CHAPTER III

DATA AND RESEARCH METHODOLOGY

A. Research Source and Method

The data used in this study is primary data taken directly from the location to be examined. This study uses mixed method, means that the research is mixed between quantitative and qualitative method. There are three basic mixed method design; convergent parallel mixed method, explanatory sequential mixed method, and exploratory sequential mixed method.

Convergent parallel mixed method is a method that is used in this study. Convergent parallel mixed method is a mixed method where a researcher collects a qualitative and quantitative data, analyses it separately, and compares the result to determine: are there any findings that confirm each others (Creswell, 2016).

B. Data Collection Technique

1. Population and Sample

The population is a collection of data sources, which have the same properties (Sukandarrumidi & Haryanto, 2008). The population in this study is the number of people in the Tanjung Putri village.

Sampling is the process of selecting a number of objects of research for a study that is representative of a population (Sumanto, 2014).

Sampling is to obtain information from the population which is the location of the study. Based on the formula Baiky (1982), so that research can be analyzed with a statistical minimum number of samples taken 30, although there are others suggested minimal sample size is 100 samples (Sukandarrumidi & Haryanto, 2008). This research uses Slovin's Formula to determine the sample size from the given population size.

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

Where:

n = sample size

N = pupulation size

e = marginal of error

To take sample for questionare, the researcher will use random sampling methods. Random sampling method is the sampling process that everyone in the population has an opportunity and same freedom to be selected as a sample (Sumanto, 2014). However, in taking sample for indepth-interview, the researcher will use purposive sampling. Purposive sampling method is a sample that is choosen based on the criteria or purpose of the study. In this study, the criteria for indepth-interview respondents are; one from local decision policy and one from a teacher.

2. In-Depth Interview

Data collection by interview is often conducted in the qualitative research. It happens when there is an interaction which contained

information exchange for sustainability research. According to Steward & Cash (2008),

"An interview is interactional Because there is an exchanging, or sharing of roles, responsibilities, feelings, beliefs, motives, and information. If one person does all of the talking and the other all of the listening, a speech to an audience of one, not an interview, is taking place."

In-depth interviewing is a qualitative research technique that involves conducting intensive individual interviews with a small number of respondents to explore their perspectives on a particular idea, program, or situation. For example, we might ask participants, staff, and others associated with a program about their experiences and expectations related to the program, the thoughts they have concerning program operations, processes, and outcomes, and about any changes they perceive in themselves as a result of their involvement in the program. The primary advantage of in-depth interviews is that they provide much more detailed information than what is available through other data collection methods, such as surveys. They also may provide a more relaxed atmosphere in which to collect information people may feel more comfortable having a conversation with you about their program as opposed to filling out a survey (Boyce & Neale, 2006).

In this study, interview techniques used is in-depth interviews. In-dept interview is an intensive individual interview with a small number of respondents to explore their perspective about the study.

3. Questionnaire

The questionnaire consists of a variety of questions related to the research. Data revealed by the questionnaire is the factual data or data that is reputed truth known by the author. Questions on the questionnaire in the form of direct questions directed to the information about the data. Questionnaire respondents know exactly what is being asked in the questionnaire and what information is required by related questions. The answers are in the questionnaire will not be given as a score, but given numbers as identification or classification coding answers (Azwar, 1999).

Based on the website that talking about questionnaire in research (The advantages and disadvantages of questionnaires). There are seven advantages of using questionnaires for research:

1. Practical.
2. Large amounts of information can be collected from a large number of people in a short period of time and in a relatively cost effective way.
3. Can be carried out by the researcher or by any number of people with limited affect to its validity and reliability.
4. The results of the questionnaires can usually be quickly and easily quantified by either a researcher or through the use of a software package.
5. Can be analysed more 'scientifically' and objectively than other forms of research.

6. When data has been quantified, it can be used to compare and contrast other research and may be used to measure change.
7. Positivists believe that quantitative data can be used to create new theories and / or test existing hypotheses

The disadvantages of questionnaires:

1. Is argued to be inadequate to understand some forms of information , such as: changes of emotions, behaviour, feelings etc.
2. Phenomenologists state that quantitative research is simply an artificial creation by the researcher, as it is asking only a limited amount of information without explanation.
3. Lacks validity.
4. There is no way to tell how truthful a respondent is being.
5. There is no way of telling how much thought a respondent has put in.
6. The respondent may be forgetful or not thinking within the full context of the situation.
7. People may read differently into each question and therefore reply based on their own interpretation of the question.
8. There is a level of researcher imposition, meaning that when developing the questionnaire, the researcher is making their own decisions and assumptions as to what is and is not important, therefore they may be missing something that is of importance.

The process of coding in the case of open ended questions opens a great possibility of subjectivity by the researcher

In this study, there are 50 questions in the questionnaire. There are five specific questions that talk about financial literacy; financial individual knowledge, investment, insurance, saving and loan, and financial institution knowledge.

C. Operational Definition of Research Variable

1. Dependent Variable

In this study, the dependent variable used is financial literacy, which is a person's ability to process and financial planning, not only for himself but also his family. To measure the level of financial literacy questionnaires will be distributed to the community that has defined his sample. Questions from questionnaires distributed are questions about financial literacy.

2. Independent Variable

There are three independent variables used in this study. The first one (1) is gender, this variable explains the gender, i.e. male and female. Second (2) is the educational background, in this variable is determined by the last study respondents were then classified into; SD, SMP, SMA, Diploma, S1, S2 and S3. Variable three (3) or the last one is the revenue, variable describes the amount of a person's income in each month are clarified in the form of average; < 1.000.000, 1.000.000 – 3.000.000, 3.000.000 – 5.000.000, 5.000.000 – 7.000.000, > 7.000.000.

D. Instrument Research Test

1. Validity Test

The data has been obtained from the questionnaire, before processing should be performed validity test first to test how valid questions are asked in the questionnaire with the variables studied (Algifari, 2015).

This study using SPSS version 15.0 in determining construct validity. Construct validity is to demonstrate that the measuring instrument in measuring valid questionnaires indicated by a strong correlation with existing variables in research (Abdillah & Hartono, 2015). Validity test of this research using the technique of corrected item correlation, in taking the decision is if $r_{\text{count}} > r_{\text{table}}$, it can be said valid, but on the contrary if the $r_{\text{count}} < r_{\text{table}}$ then said invalid. Once it is found valid or not, the researcher do significant testing using r table, it can be said significant at the 0.05 level or 5%.

2. Reliability Test

If the validity test is already done, the next step is doing reliability test. Reliability tests are used to test how consistent the answers given by respondents in the questionnaires given (Algifari, 2015).

Reliability test conducted with Cronbach Alpha formula:

$$r_i = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum s_i^2}{\sum s_t^2} \right)$$

Table 3.1
Criteria of Reliability Coefficient Based on Guilford

Value	Explanation
$r_{11} < 0,20$	Very Low
$0,20 < r_{11} < 0,40$	Low
$0,40 < r_{11} < 0,70$	Moderate
$0,70 < r_{11} < 0,90$	High
$0,90 < r_{11} < 0,100$	Very High

Source: (Basuki, 2015)

Data can be said reliable if the Cronbach Alpha (r_{11}) at least at 0.8. But there also are found Croanbach Alpha at 0.6 can be said reliably (Algifari, 2015).

E. Data Analysis

1. Descriptive Analysis for Interview

Descriptive research can be either quantitative or qualitative. It can involve collections of quantitative information that can be tabulated along a continuum in numerical form, such as scores on a test or the number of times a person chooses to use a-certain feature of a multimedia program, or it can describe categories of information such as gender or patterns of

interaction when using technology in a group situation. Descriptive research involves gathering data that describe events and then organizes, tabulates, depicts, and describes the data collection. It often uses visual aids such as graphs and charts to aid the reader in understanding the data distribution. Because the human mind cannot extract the full import of a large mass of raw data, descriptive statistics are very important in reducing the data to manageable form. When in-depth, narrative descriptions of small numbers of cases are involved, the research uses description as a tool to organize data into patterns that emerge during analysis. Those patterns aid the mind in comprehending a qualitative study and its implications.

Data analysis techniques in qualitative research through a procedure that has been done by many experts qualitatively. The procedure is undergoing data collection, data input, data analysis, conclusion and verification, and ending with the form of the findings in narrative form (Herdiansyah, 2010).

2. Likert scale

This study use likert skale as the data analysis in quantitative method. In response to the difficulty of measuring character and personality traits, developed a procedure for measuring attitudinal scales (Likert, 1932). The original Likert scale used a series of questions with five response alternatives: strongly approve (1), approve (2), undecided (3), disapprove (4), and strongly disapprove (5).

Respondents will choose one of five options given to answer questions. Questions contained in the questionnaire are questions that are approved and in accordance with what is being used as research material. In this study, the question on the questionnaire are questions about financial literacy, such as savings, investment, insurance, and others.

In analyzing Likert scale, researchers will create weight value table and percentage value table, then uses an index formula percent.

Table 3.1
Weight Value Table

Very Agree	5
Agree	4
Neutral	3
Disagree	2
Very Disagree	1

Source: (Likert, 1932)

Table 3.2
Percentage Value Table

Answer	Description
0% - 19.99%	Very (Disagree, Once More or Less)
20% - 39.99%	Disagree or Less Good
40% - 59.99%	Quite or Neutral
60% - 79.99%	Agree, Neither or Likes
80% - 100%	Very (Agree, Neither, Likes)

Source: (Likert, 1932)

$$\text{Index \% Formula} = \text{Total Score} / Y \times 100$$

After calculating the percent of the index is complete, it will be found the percentage of people who agree or disagree on the results of a questionnaire given to the sample.

3. Descriptive Statistics

Descriptive statistics are used to describe the data or make a summary of the data in the first phase of data analysis (Sumanto, 2014). Descriptive statistics are not concluded directly, but only provides information to use data that has been processed before, usually it is displayed in the form of tables, graphs and diagrams.

Descriptive statistics are typically distinguished from inferential statistics. With descriptive statistics the researchers are simply describing what is or what the data shows. With inferential statistics, the researchers are trying to reach conclusions that extend beyond the immediate data alone. Thus, the researchers use inferential statistics to make inferences from our data to more general conditions; we use descriptive statistics simply to describe what's going on in the data.

Descriptive Statistics are used to present quantitative descriptions in a manageable form. In a research study we may have lots of measures. Or we may measure a large number of people on any measure. Descriptive statistics help us to simplify large amounts of data in a sensible way. Each descriptive statistic reduces lots of data into a simpler summary.

4. SWOT Analysis

Usually, SWOT analysis used in the management and business research. SWOT itself is the abbreviation from Strength, Weakness, Opportunity, and Treat. Strength and weakness used for internal analysis, while opportunity and treat used for external analysis. This research used

SWOT analysis because it will be useful to get the conclusion and recommendation (FME, 2013).

5. Binary Logistic Regression Analysis

Regression analysis is one of statistical data analysis techniques that often used to examine the relationship between several variables and predict a variable (Kutner, Nachtsheim, Neter, & Li, 2004). There are two types of regression analysis; simple regression and multiple regression.

Generally, logistic regression is well suited for describing and testing hypotheses about relationships between a categorical outcome variable and one or more categorical or continuous predictor variables. Logistic regression is an approach to make a prediction model, such as in linear regression. The difference is in dependent variable; dichotomy scale. Dichotomy scale is a nominal data scale with two categories, for example: yes and no, good and bad, or high and low.

There are two main uses of logistic regression: The first is the prediction of group membership. Since logistic regression calculates the probability of success over the probability of failure, the results of the analysis are in the form of an odds ratio. Logistic regression also provides knowledge of the relationships and strengths among the variables.

This study will uses binary logistic regression. Binary logistic regression is used when there are only two possibility of respond variable (Y). The equation is:

$$\ln\left(\frac{\hat{p}}{1 - \hat{p}}\right) = B_0 + B_1X$$

Where:

\ln : Logaritma natural

$B_0 + B_1X$: OLS equation

Based on the equation above, it will be difficult to interpret the regression coefficient. So, this research will use Odds Ratio or $\text{Exp}(B)$ or OR. $\text{Exp}(B)$ is a component from regression coefficient. This interpretation means the category code in every variables:

1. Independent variable is financial literacy: 0 code is for the awareness of financial literacy is low, and code 1 is for the awareness of financial literacy is high.
2. Dependent variable is gender: 0 code is for man, and 1 code is for woman.
3. Dependent variable is indome: 0 code is for income a month $< 1.000.000$, 1 code is for income a month $1.000.000 - 3.000.00$, 2 code is for income a month $3.000.000 - 5.000.000$, 3 code is for income a month $5.000.000 - 7.000.000$, and 4 is for income a month $> 7.000.000$.
4. Dependent variable is education background: 0 code is for SD, 1 code is for SMP, 2 code is for SMA, 3 code is for S1, and 4 code is for S2.

Another difference in logistic regression is there is no “R Square” to measure the amount of simultaneous affect of independent variable to

dependent variable. In logistic regression called it Pseudo R Square, R Square similar with R Square in OLS.

If in OLS uses F Anova test to measure significant level and a goodness of equation model. In logistic regression uses Chi Square value based on Likelihood maximum measurement (Basuki, 2015).

- Assumptions of Logistic Regression:
 1. Logistic regression does not assume a linear relationship between the dependent and independent variables.
 2. The dependent variable must be a dichotomy (2 categories).
 3. The independent variables need not be interval, nor normally distributed, nor linearly related, nor of equal variance within each group.
 4. The categories (groups) must be mutually exclusive and exhaustive; a case can only be in one group and every case must be a member of one of the groups.
 5. Larger samples are needed than for linear regression because maximum likelihood coefficients are large sample estimates. A minimum of 50 cases per predictor is recommended.

F. Classic Assumption Test

1. Heterokedastisitas Test

Heterokedastisitas test is to see whether there is variance's dissimilaritis from one residual observation to another one. Heterokedastisitas detection can be done though conducting scatter plot

method by plotting the value of ZPRED (prediction value) with SRESID (residual value). Statistic test which is used in this study is Glejser test.

Another alternative is if the heterokedastisitas assumption is not appropriate, then it should be transformed into logarithm form if all the data is positive measurement (Basuki, 2015).

2. Normality Test

Normality test is conducted to see whether the value of residual is normally distributed or not. Normality test is done to residual value, not to each variable. Some statistic experts argued that the data whose number is higher than 30, then it can be said that it is normality distributed. In this study it will use Kolmogorov Smirnov as the statistic test (Basuki, 2015).

3. Multikolinearity Test

Multikolinearity test is to see whether there is high correlation between independent variables in multiple regression model. If independent variables show high correlation, that dependent variable will be disturbed. In this study it will use variance inflation factor (VIF). If VIF value < 10 then there is no multikolinearity among independent variables, and vice versa (Basuki, 2015).