

CHAPTER III

RESEARCH METHODOLOGY

A. Research Design

In this study, the researcher uses a quantitative approach with regression analysis. Quantitative research is the research which based on positivism philosophy in which is used for observed certain populations or sampling. The technique to get sample usually has done randomly, collecting data commonly used instrument and data analysis using statistical approach to examine a hypothesis¹.

While regression analysis is a research study method that involves collecting data in order to determine the degree of relationship exists between two or more quantifiable variables. The purpose of this method is to investigate whether the raising of variable dependent can be influenced by the raising of variable independent².

A regression analysis is most appropriate for this study because the purpose of this study is to find out how much the influence of students' frequency of watching English movies to their vocabulary power is.

B. Research Setting

This study is field research in Madrasah Aliyah Nahdlatul Ulama' Banat Kudus. In this effort to find the solution to the problems in this thesis, the researcher needs some data. In conducting the study, the researcher tried to get some data from participant from students of eleventh grade of language study program of Madrasah Aliyah Nahdlatul Ulama' Banat Kudus.

The researcher chooses MA NU Banat Kudus as the research's setting because the school has the best accreditation (A) and gets many achievements such as the winner of English article and PRS at regency level, Speak contest

¹Sugiyono, *Metode Penelitian Pendidikan*, (Bandung : Alfabeta Press,2008), p. 15

²Sugiyono, *Statistik Untuk Penelitian*, (Bandung : Alfabeta Press,2010), p. 260

at SMS/MA level, English debate contest at regency level, National since Olympiad at regency level and English speech contest at province level³.

C. Population

Population is generalization area that consists of object or subject that has special quality and characteristic that determine by the researcher to be investigated and then take the conclusion⁴. In this study, the populations are students at eleventh grade of Language study program of Madrasah Aliyah Nahdlatul Ulama' Banat Kudus in the academic year of 2012/2013. The total population is 45 students. This study does not use a sample because the population is less than 100 students.

The researcher chooses eleventh grade because according to Elizabeth B. Hurlock in her book *Psychology Development*, watching movie is one of entertainment that is interesting for teenagers. The girls prefer romantic movie whereas the boys prefer adventure movie⁵. So that's why, they are suitable to do this research because this research is talking about students' interest in watch English movies.

D. Variable and Indicator

Variable is a variation object of the study. There are two types of variables; they are dependent variable and independent variable. Dependent variable is the variable of focus or the central variable on which other variables will act if there is any relationship. Independent variable is selected by researcher to determine the relationship with the dependent variable⁶.

³ MA NU Banat Kudus/Prestasi in www.banatnukudus.or.id, accessed on December 28, 2012

⁴Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif dan R & D*, (Bandung: Alfabeta, 2008), p. 80

⁵ Elizabeth B. Hurlock, *Psikologi Perkembangan*, (Jakarta: Erlangga) p. 218

⁶Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktek*, (Jakarta: PT Rineka Cipta, 2006), p. 118 -119.

Variables in this research are frequency of watching English movies (as independent variable) and students' vocabulary power (as dependent variable).

The indicators of the frequency of watching English movies are students' interest and rate of watching English movies and the indicator of students' vocabulary power is students' ability to pronounce, know the meaning and use vocabulary.

E. Data Collection Technique

The researcher used technique of data collection as follows:

1. Questionnaire

Arikunto states a questionnaire is a number of written questions, which are used to gain information from respondents about the respondents themselves, or their knowledge, believe, etc.⁷ In this case, the researcher has spread the sheets of paper as questionnaire to get the data about the result of students' frequency of watching English movies and their vocabulary power at eleventh grade in Language program of Madrasah Aliyah Nahdlatul Ulama' Banat Kudus.

The questioners are divided into two parts. The first part is to know students' frequency of watching English movies. In this part, the questions consist of five options with different point, Always=5, Often=4, sometimes=3, seldom=2, and never=1. The second part is to know how the effects of watching English movies in students' vocabulary power. The questions consist of five options with different points, strongly agree=5, Agree=4, neutral=3, Disagree=2, strongly disagree=1. And the indicators are students' frequency in watching English movies and its effect in their vocabulary power.

Before the test instrument used in this study, the researcher conducted testing instrument first. The goal is to obtain a good instrument, one that meets the criteria of valid and reliable.

⁷Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktek*, p. 151

The first is validity analysis. To test the validity of a question about, the *Pearson* correlation formula is used, that is:

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

Where:

- r_{xy} = correlation coefficient
- N = the number of students
- X = score point of question (item)
- Y = total score points of question

After Calculating, the count is calculated compared to r table with a significance level of 5%, if $r_{hitung} > r_{table}$ so, the question is valid.⁸

Then, reliability analysis. To test the reliability of the question, the alpha formula is used as follow:

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right)$$

Where,

$$\sigma^2 = \frac{\sum x^2 - \frac{(x)^2}{N}}{N}$$

Where,

- r_{11} = reliability
- n = The number of item in the test
- N = The number of the subject
- $\sum \sigma_i^2$ = total variance score each question⁹

After getting r_{11} , the price of r_{11} is compared with the price of r_{table} , if $r_{11} > r_{table}$ so the item tested is reliable.

⁸ Purwanto, *Instrument Penelitian Sosial Dan Pendidikan*, (Yogyakarta: Pustaka Pelajar, 2010), p. 127

⁹ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 2009), 9th Ed., p. 97-106

2. Test

Test is an examination or trial to find its quality, value, composition, etc.¹⁰. In this research, the researcher gives Vocabulary test about synonym, antonym and contextual task. The students have to answer thirty questions in thirty five minutes.

There are some purposes of assessing vocabulary, they are used to measure vocabulary size (useful for placement purpose or as one element of a proficiency measure), to measure what has just been learned (a short – time achievement measure) and to measure what has been learned in a course (a long – term achievement measure) and to diagnose areas of strength and weakness¹¹. In this case, the assessing vocabulary is used to diagnose areas of strength and weakness of students' vocabulary at Eleventh grade in Language study program of MA Nahdlatul Ulama' Banat Kudus in the academic year of 2012/2013.

Before the test instrument used in this study, the researcher conducted testing instrument first. The goal is to obtain a good instrument, one that meets the criteria of valid, reliable, having distinguishing the good and the moderate level of difficulty.

The first is validity analysis. To test the validity of a question about, the *Pearson* correlation formula is used, that is:

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

Where:

r_{xy}	= correlation coefficient
N	= the number of students
X	= score point of question (item)
Y	= total score points of question

¹⁰Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktek* , p.150

¹¹Scmitt and Norbert, *An Introduction to Apply Linguistic*, (New York: Edward Arnold, 2002) p. 46

After Calculating, the count is calculated compared to r table with a significance level of 5%, if $r_{hitung} > r_{table}$ so, the question is valid.

The second is reliability analysis. To test the reliability of the instrument, the formula is used as follow:

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{S - \sum pq}{S^2} \right)$$

Where,

r_{11} = The reliability coefficient items

n = The number of item the test

p = The proportion of students who give the right answer

q = The proportion of students who give the wrong answer

S^2 = The standard deviation of the test

After getting r_{11} , the price of r_{11} is compared with the price of r_{table} , if $r_{11} > r_{table}$ so the item tested is reliable.

The third is difficulty level analysis. In terms of difficulty, a good question is a question that is not too easy and not too difficult. A question that is too easy will not stimulate students to increase their effort to answer the question and a question that is too hard will cause students to be desperate and do not have the spirit to try again because beyond the reach of his ability. Difficulty level of a question is determined by the formula:

$$P = \frac{B}{JS}$$

Where:

P : Difficulty level

B : The number of students who answer correctly

JS : The number of participants

Criteria

1. $00 < P \leq 0.30$ (difficult)

0. $30 < P \leq 0.70$ (Medium)

0. $70 < P \leq 100$ (easy)¹²

The last is discrimination index analysis. It is used to distinguish between high-ability learners with low-ability learners. Figures show the amount of distinguishing features is called the index of discrimination (D), this discrimination is between 0, 00 to 1, 00. On discrimination index, negative sign is used when something about the "inverse" indicates the quality of the test¹³. The formula for determining the discrimination index is:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

$$P_A = \frac{B_A}{J_A}$$

$$P_B = \frac{B_B}{J_B}$$

Where,

D = Discrimination index

J_A = the number of participants on the top group

J_B = the number of participants on the bottom group

B_A = the number of participants on the top group who answer correctly

B_B = the number of participants on the bottom group who answer correctly

P_A = Proportion of participants on the top group who answer correctly
(P as difficulty index)

P_B = Proportion of participants on the bottom group who answer correctly

For the question $J_A = J_B = 27\% \times N$, N is the number of participants.

Discrimination index criteria (D) is as follows:

D : 0. 00 – 0. 20 = Poor

D : 0. 20 – 0. 40 = Satisfactory

¹²Daryanto, *Evaluasi Pendidikan*, (Jakarta: Rineka Cipta, 2008), p. 179 - 182

¹³Daryanto, *Evaluasi Pendidikan*, p. 183

D : 0. 40 – 0. 70 = Good

D : 0. 70 - 1. 00 = Excellent

F. Data Analysis Technique

Data analysis is the activity that is done after all of the data is collected. The activities in this analysis are grouping the data based on the variable and respondent, tabulating the data based on the variable, serving the data, doing the calculation of data to answer the research question and testing the hypothesis.¹⁴

The analysis steps are as follow:

1. Introduction Analysis

Introduction analysis is to simplify the data in order to be able to be read and interpreted easily. There are two data namely:

a. Questionnaire Data

The questionnaire data is counted up on preparation table and each alternative option has different score as follow:

1) Always / Strongly Agree (A) = 5

2) Often / Agree (B) = 4

3) Sometimes / Neutral (C) = 3

4) Seldom / Disagree (D) = 2

5) Never / Strongly Disagree (E) = 1

b. Students' English Vocabulary Test Data

The students' English vocabulary test is calculated to look for average of each student. The students have to answer thirty questions in thirty five minutes. Each correct answer gets 1 score with the formula as follows:

$$Score = \frac{\text{The correct answer}}{\text{The total of questions}} \times 100$$

¹⁴Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif dan R & D*, p. 147

2. Hypothesis Analysis

The examination uses regression one predictor analysis, while the steps are as follow:

- a. Looking for the correlation between predictor and criterion by using technique of correlation product moment technique, with formula:

$$r_{xy} = \frac{(\sum xy)}{\sqrt{(\sum x^2)(\sum y^2)}}$$

$$\sum xy = \sum XY - \frac{(\sum X)(\sum Y)}{N}$$

$$\sum x^2 = \sum X^2 - \frac{(\sum X)^2}{N} \text{ and}$$

$$\sum y^2 = \sum Y^2 - \frac{(\sum Y)^2}{N}$$
¹⁵

Where:

r_{xy} = the correlation coefficient

$\sum xy$ = the sum of X and Y score

$\sum x$ = the sum of Variable X

$\sum y$ = the sum of Variable Y

- b. Examining whether there is any significant correlation or not by consulting the result of r_{xy} with table value (r_t)

From the correlation result, the researcher will interpret that category based on the following:

0, 80 – 1,000 means very high correlation

0, 60 – 0, 799 means high correlation

0, 40 – 0, 599 means enough correlation

0, 20 – 0, 399 means low correlation

0, 00 – 0, 199 means very low correlation¹⁶

- c. Find the regression equation with the formula:

$$Y = aX + K$$

¹⁵Sutrisno Hadi, *Analisis Regresi*, (Yogyakarta: Andi Office, 2001), p. 4

¹⁶Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif dan R & D*, p. 184

Where:

Y = Criterion

X = Predictor

a = the numeral of predictor coefficient

K = the numeral of constant

To look for the score of a and K , the researcher uses deviation score method. The formula is as follows:

$$y = ax$$

Where, $y = Y - \bar{Y}$, $x = X - \bar{X}$, and $a = \frac{\sum xy}{\sum x^2}$ ¹⁷

d. Variant analysis of regression line

$$F_{reg} = \frac{RK_{reg}}{RK_{res}}$$

$$RK_{reg} = \frac{JK_{reg}}{db_{reg}}$$

$$RK_{res} = \frac{JK_{res}}{db_{res}}$$

$$JK_{reg} = \frac{(\sum xy)^2}{\sum x^2}$$

$$JK_{res} = \sum y^2 - \frac{(\sum xy)^2}{\sum x^2}$$
¹⁸

Where :

F_{reg} = The value of F numeral for line regression

RK_{reg} = Mean of regression line quadrate

RK_{res} = Mean of residue quadrate

db_{reg} = degree of residue freedom

¹⁷Sutrisno Hadi, *Analisis Regresi*, p. 6

¹⁸Sutrisno Hadi, *Analisis Regresi*, p. 13

3. Final Analysis

After getting F_{reg} , the next step is comparing the price of F_{reg} with F value on table value. The table value is 5 %:

1. It is significant if $F_{reg} > Ft 5 \%$. There is positive influence of students' frequency of watching English movies and their vocabulary power.
2. It is not significant if $F_{reg} < Ft 5 \%$. There is no influence of students' frequency of watching English movies and their vocabulary power.