

CHAPTER III

RESEARCH METHODOLOGY

3.1. Research Design

Research on the factors that influence the price of Bitcoin using quantitative descriptive methods. Descriptive method is a research method that describes the nature of a situation that is temporarily running when the study is conducted and checks the causes of these symptoms. While quantitative methods are used to develop and use mathematical models, theories and / or hypotheses related to research. This study seeks to find out the factors that influence the price of Bitcoin both through internal factor and external factors. The internal factor is the volume of bitcoin, whether the external factors are low volatility index, price of gold, and price of oil.

The data used is monthly time series data from June 2013 to July 2019. The researcher decided to start data from June 2013 because starting that date, users began using Bitcoin wallets on the blockchain. Previously, since the launch of Bitcoin in 2009 to August 2011 there was no Bitcoin wallet service yet. The analytical method used to determine the relationship of independent variables to the dependent variable is Multiple Regression using Ordinary Least Square Model.

3.2. Data Selection

The type of data in this study is quantitative data. This study uses secondary data in a time series from June 2013 to July 2019. The sources of data in this research can be seen as below:

Table 3.1
Sources of the Variables in the Study

Variable	Source
Bitcoin Price	www.investing.com
Bitcoin Volume	www.bitcoinity.org
Dow Jones Industrial Average	www.investing.com
Bitcoin wallet users	www.quandl.com
Gold Price	www.investing.com
Oil Price	www.investing.com

After that, the data is compiled in Microsoft Excel 2019 and the data is processed using the help of EViews 10.

3.3. Operational Definition of Research Variables

a. Bitcoin Price (BTC_PRICE_Y)

Bitcoin price is the amount that must be paid to get 1 BTC, the unit in USD. The price of the Bitcoin is from Bitfinex.

b. Bitcoin Volume (BTC_VOLUME_X1)

Bitcoin volume is the number of Bitcoin transactions that occur, the unit in BTC.

c. Dow Jones Industrial Average (DJIA_X2)

DJIA is an index created by Dow as a way to measure the performance of industrial components in the American stock market.

d. Bitcoin Wallet Users (BTC_WALLET_X3)

Bitcoin wallet users are the Bitcoin My Wallet number of users.

e. Gold Price (GOLD_PRICE_X4)

The price of gold is the amount that must be paid to get gold, the unit in USD.

f. Oil Price (OIL_PRICE_X5)

The price of oil is the amount that must be paid to get oil, the unit in USD.

3.4. Data Collection Technique

The data collection technique used is the documentation method. The documentation method is to look for data in the form of notes, transcripts, books, newspapers, magazines, inscriptions, minutes of meetings, agendas, etc. (Arikunto, 2010). The documentation used in this study is in the form of document data from quandl.com, investing.com, blockchain.info, regulations regarding the use of bitcoin as an investment tool and a means of payment, and news about Bitcoin

3.5. Research Application

In a research, researchers need an application which could help and simplify the calculation and analysis of the data. For this research, research used EViews as the application to process the data and use Microsoft Excel to compile the raw data.

Econometric Views (EViews) is one of the statistical and econometric data processing applications that runs on Windows OS.

EViews can be used to analyze statistical data on economic and business research such as using linear regression techniques, correlation, causality granger, ARCH and GARCH, etc. This software interface is indeed a little different than SPSS or other statistical software. One of the advantages of EViews compared to SPSS is that EViews can analyze causality or reciprocity. EViews are applications that are more suitable for processing time series data. The model of EViews that is used is EViews 10, as it is the newest version of the application.

3.6. Research Method

Data analysis technique used in this research is quantitative data analysis method. Analysis of the data used is time series data analysis with Ordinary Least Square Method.

3.6.1. Multiple Regression Model

The type of data used in this study is time series data. The equation of Ordinary Least Squares (OLS) procedure can be stated as follows:

$$\begin{aligned} \text{BTC_PRICE_}Y_t &= \beta_0 + \beta_1\text{BTC_VOLUME_}X_1 + \beta_2\text{DJIA_}X_2 + \\ &\beta_3\text{BTC_WALLET_}X_3 + \beta_4\text{GOLD_PRICE_}X_4 + \\ &\beta_5\text{OIL_PRICE_}X_5 + \mu \end{aligned}$$

Description:

BTC_PRICE_Y = variable of bitcoin price

β_0 = constant / intercept

$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ = quantitative independent variable regression coefficient

BTC_VOLUME_X1 = variable of bitcoin volume

DJIA_X2 = variable of dow jones industrial average

BTC_WALLET_X3 = variable of bitcoin wallet users

GOLD_PRICE_X4 = variable of gold price

OIL_PRICE_X5 = variable of oil price

μ = Residual value

In order to fulfill the classic assumption which can be seen in Chapter 4, the model equation with transformed data is shown as follows.

$$\begin{aligned} d(\text{BTC_PRICE_Y}) &= \beta_0 + \beta_1 d(\text{BTC_VOLUME_X1}_t) + \beta_2 d(\text{DJIA_X2}_t) + \\ &\quad \beta_3 d(\text{BTC_WALLET_X3}_t) + \beta_4 d(\text{GOLD_PRICE_X4}_t) + \\ &\quad \beta_5 d(\text{OIL_PRICE_X5}_t) + \mu \end{aligned}$$

Description:

$d(\text{BTC_PRICE_Y})$ = the first different form of the variable of bitcoin price

β_0 = constant / intercept

$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$	= quantitative independent variable regression coefficient
$d(\text{BTC_VOLUME_X1})$	= the first different form of the variable of bitcoin volume
$d(\text{DJIA_X2})$	= the first different form of the variable of dow jones industrial average
$d(\text{BTC_WALLET_X3})$	= the first different form of the variable of bitcoin wallet users
$d(\text{GOLD_PRICE_X4})$	= the first different form of the variable of gold price
$d(\text{OIL_PRICE_X5})$	= the first different form of the variable of oil price
μ	= Residual value

3.6.2. Classic Assumption Test

3.6.2.1. Normality Test

Data normality testing is testing about normality of data distribution (Santoso & Ashari, 2005). Normality testing is done with a view to seeing whether or not the data analyzed is normal. A good regression model has a normal or near normal data distribution. Residual values that are normally distributed can be seen from the shape of the curve that forms a bell-shaped curve whose sides are widening to infinity (Suliyanto, 2011).

Besides using graphics, normality can also be tested by several methods, one of which is by Jarque-Bera (JB Test). The JB test is a normality test based on the kurtosis and skewness.

H_0 : residuals are normally distributed

H_a : residuals are not normally distributed

Decision making is carried out with the following criteria:

- If prob. JB > 0.5 then H_0 accepted means that the residuals are normally distributed.
- If prob. JB < 0.5 then H_0 is rejected means that the residuals are not normally distributed.

3.6.2.2. Multicollinearity Testing

Multicollinearity test is conducted to test whether there is a correlation between independent variable in a regression model. A good regression model should not contain any correlation between independent variable (Ghozali, 2018). Multicollinearity can be detected by Variance Inflation Factor Value (VIF) and Tolerance Value as a rule of thumb. The Tolerance Value is the opposite of the opposite of the Variance Inflation Factor Value ($TOL = 1/VIF$). These two measures refer to what kind of independent variables are described by other independent variables. In other words, the independent variable becomes the dependent variable (bound) and re-created against other independent variables. Tolerance measures the variability of selected independent variables that are not explained by other independent variables.

H_0 : the model is not multicollinearity

H_a : the model is multicollinearity

Multicollinearity test is conducted by using these criteria:

- If $VIF < 10$ and tolerance value > 0.1 , then there is no multicollinearity.
- If $VIF > 10$ and tolerance value < 0.1 , then there is an indication of multicollinearity.

3.6.2.3. Homoscedasticity Testing

Homoscedasticity test to test whether the regression model occurs in the same variance of residuals. A good model is a homoscedasticity model where the variance and residual values between observations are the same so that they meet the assumptions. Testing to detect heteroscedasticity using White Heteroscedasticity (Gujarati & Porter, 2010).

H_0 : the model is homoscedasticity

H_a : the model is not homoscedasticity

Decision making is carried out with the following criteria:

- If the Prob. value of Obs*R-squared > 0.5 then H_0 is accepted, it means that the model is homoscedasticity.
- If the Prob. value of Obs*R-squared < 0.5 then H_0 is rejected means that the model is not homoscedasticity.

3.6.2.4. Non-Autocorrelation Testing

Non-autocorrelation test aims to test whether the model has a correlation between confounding errors or not. Autocorrelation occurs because observations that appear sequentially over time are related to one another. Good data is data that does not occur autocorrelation (non-autocorrelation). This test uses the LM test.

H_0 : the model does not occur autocorrelation

H_a : the autocorrelation model occurs

Decision making is carried out with the following criteria:

- If the Prob. value of Obs*R-squared > 0.5 then H_0 is accepted means that the model does not occur correlation (non-autocorrelation).
- If the Prob. value of Obs*R-squared < 0.5 then H_0 is rejected, it means that the model is autocorrelated.

3.6.3. Hypothesis Testing

a. Stimulation Test (Statistical Test F)

The F test aims to determine the effect of all independent variables simultaneously on the dependent variable. The F test is also called the model feasibility test which is used to identify a regression model that is estimated to be feasible or not. Eligible here means that the estimated model is feasible to use to explain the effect of the independent variables on the dependent variable. If the calculated F probability

value < significance level of 0.10, it can be said that the estimated regression model is feasible.

b. Partial Test (t-Test Statistics)

Regression coefficient is used to determine the effect of partially independent variables on the dependent variable. Testing of the regression results is done using the t test at a degree of confidence of 90% or $\alpha = 10\%$ with the following conditions:

H_0 : if the probability of t-Statistics < 0.10 then H_0 is rejected.

H_a : if the probability of t-Statistics > 0.10 then H_a is rejected.

c. Coefficient Determination (R^2)

The adjusted R^2 value measures the goodness of fit on how far the model's ability to explain the variation of the dependent variable. The adjusted R^2 value is a summary measure that shows how well the sample regression line matches the population data. The coefficient of determination is between zero and one.

A small adjusted R^2 value means that the ability of the independent variables to explain the dependent variable is very limited (Ghozali, 2009). Model compatibility is said to be "better" if the adjusted R^2 value gets closer to 1.