CHAPTER III

RESEARCH METHODOLOGY

3.1 Research Design

The quantitative technique and causal research design are being used by the author in this study to empirically show how one variable affects another by acquiring data and facts. Causal studies look into whether one factor can affect another (Sekaran & Bougie, 2016). The objective of conducting causal research by the writer is to ascertain the impact of independent variables such as profitability and leverage on the dependent variable of dividend policy on state-owned enterprises listed in Indonesia Stock Exchange for the year 2016-2022.

3.2 Population and Sample

The term population denotes the entire collection of individuals or entities under consideration, along with the corresponding measurements obtained from every component of the population (Lind et al., 2020). In this research, the population is state-owned enterprises that listed in Indonesia Stock Exchange from the year 2016-2022 with the total of 20 companies. The main reasons why state-owned enterprises are chosen in this research are:

1. State-owned Enterprises (BUMN) are business entities where the state holds all or most of the capital through direct participation from segregated state assets.

2. State-owned enterprises play a crucial role in the execution of the national economy with the aim of achieving the well-being of the society.
3. The establishment of State-owned Enterprises aims to make a significant contribution to the overall development of the national economy, as well as to state revenues. These enterprises strive to generate profits while also providing high-quality goods and/or services that meet the needs of many individuals. Additionally, State-owned Enterprises take on a pioneering role in business activities that private companies are unable to undertake, and they engage in cooperative efforts with other organizations. Ultimately, the overarching goal of State-owned Enterprises is to promote public benefits.

Sample is a portion or part of the population of interest. The utilization of samples is a common practice in obtaining precise approximations of population parameters (Lind et al., 2020). This study employs a purposive sampling method, which is predicated on the premise that identifying optimal cases for investigation yields superior data and that research outcomes are a direct consequence of the selected cases. This approach involves the selection of "information-rich cases" that are most suitable for achieving the research objectives and addressing the research questions (Leavy, 2017).

A purposive sample consists of individuals who have been consciously or purposefully selected to possess characteristics pertinent to the research's objectives, making this sampling technique convenient and practical (Sekaran & Bougie, 2016).

Thus, the criteria of the sample are as follows:

1. State-owned Enterprises listed in Indonesia Stock Exchange in 2016-2022
2. Consistently publish annual report in 2016-2022
3. Generate profit during the period 2016-2022.
4. Distribute dividends to its shareholders during the period 2016-2022.

3.3 Data Collection Method

In this study, the author has employed a secondary data collection method by utilizing the financial statements of state-owned enterprises that are publicly available on the Indonesia Stock Exchange website or the respective company's websites between 2016 and 2022. Secondary data refers to information that has been collected by individuals or organizations for reasons other than the present research inquiry. Moreover, the documentation approach was used to investigate and collect data or information contained in the financial statements for this research's data collection procedure. The present study utilizes the library research methodology to investigate a diverse range of literary sources, scholarly journals, articles, and other written materials that pertain to the research issue and aim. (Sekaran & Bougie, 2016).

3.4 Operational Variable Definition and Variable Measurement

This research consists of two independent variables, including profitability and leverage, and one dependent variable as well, which is dividend policy. The definitions of the above-mentioned operational variables are provided below.

3.4.1. Dependent Variable

The dependent variable is the variable that is being predicted or estimated. The term commonly used to refer to the consequence or conclusion associated with a particular value of the independent variable. The variable that is being measured and analyzed in relation to the independent variable is a stochastic variable. There exist numerous alternative outcomes for the dependent variable corresponding to a particular value of the independent variable (Lind et al., 2020). Thus, in this research the dependent variable is dividend policy. Zutter & Smart (2019) elucidates that the dividend policy is a predetermined plan of action that guides the company's decision-making procedure whenever it contemplates a dividend distribution. According to Saputra & Agustin (2022), the distribution of dividends is perceived as a positive indication for investors, as it implies that the company is in an adequate economic position and has promising growth potential, potentially resulting in a rise in stock valuations. Elevated stock prices indicate an increased responsibility for the company to uphold investor confidence, prompting managers to refrain from implementing earnings management tactics that could potentially compromise the company's standing and prioritize preserving its favorable reputation.

This study will measure dividend policy using the dividend payout ratio (DPR). It has been utilized in several research conducted by Dang et al. (2018), E. Ayunku & Markjackson (2019), Soi & Buigut (2020), and Lee et al. (2022) with the formula as follows:

$$\text{Dividend Payout Ratio (DPR)} = \frac{\text{Dividend}}{\text{Net Income}}$$

3.4.2. Independent Variable

The establishment of the independent variable is fundamental in the estimation or prediction of the dependent variable (Lind et al., 2020). An independent variable is one that has an effect on the dependent variable, either positively or negatively (Sekaran & Bougie, 2016). The independent variables of this research include profitability and leverage.

3.4.2.1.Profitability

According to Wild et al. (2014), profitability is the capacity to deliver monetary benefits sufficient to generate and preserve funding. In this research profitability will be measured by using Return on Assets (ROA). Zutter & Smart (2019) proclaim that Return on Assets (ROA) metric serves as an indicator of a firm's capacity to generate profits from the assets it makes use of. The computation of the return on assets indicates the proportion of the company's profit-making capacity that is accessible to common shareholders, considering all its assets.

Return on assets will be calculated using the formula used in several previous researches by Dang et al. (2018), Soi & Buigut (2020), Aryani & Fitria (2020), and Lee et al. (2022) as follows:

$$\text{Return on Total Asset} = \frac{\text{Net Income}}{\text{Total Assets}}$$

3.4.2.2.Leverage

According to Brigham & Houston (2020), the concept of solvency relates to a collection of financial ratios that evaluate the efficacy of a company's debt management practices. In this research, leverage will be measured by Total Debt Ratio or Debt to Assets Ratio (DAR) as the debt to assets ratio, considers all debts to all creditors of all maturities. The debt ratio also shows total liabilities as a percent of total assets which also shows the proportion of assets that is financed by debt. Debt to Assets Ratio (DAR) will be calculated using the formula used in several previous research Dang et al. (2018), Soi & Buigut (2020), E. Ayunku & Markjackson (2019), and Lee et al. (2022) with the formula as follows:

$$\text{Total Debt Ratio} = \frac{\text{Total Assets} - \text{Total Equity}}{\text{Total Assets}}$$

Table 3.1 Definition and Measurement of the Variables

Variable	Definition	Indicator	Scale
Dividend Policy (Y)	A predetermined plan of action that guides the company's decision-making procedure whenever it contemplates a dividend distribution (Zutter & Smart, 2019)	$\text{DPR} = \frac{\text{Dividend}}{\text{Net Income}}$ (Dang et al., 2018; E. Ayunku & Markjackson, 2019; Lee et al., 2022; Soi & Buigut, 2020)	Ratio
Profitability (X_1)	The capacity to deliver monetary benefits sufficient to generate and preserve funding (Wild et al., 2014)	$\text{ROA} = \frac{\text{Net Income}}{\text{Total Assets}}$ (Aryani & Fitria, 2020; Dang et al., 2018; Lee et al., 2022; Soi & Buigut, 2020)	Ratio
Leverage (X_2)	A collection of financial ratios that evaluate the efficacy of a company's debt management practices (Brigham & Houston, 2020)	$\text{DAR} = \frac{\text{Total Debt}}{\text{Total Assets}}$ (Dang et al., 2018; E. Ayunku & Markjackson, 2019; Lee et al., 2022; Soi & Buigut, 2020)	Ratio

Source: Prepared by the writer (2023)

3.5 Data Analysis Method

The present research relies on a quantitative approach to data analysis, employing numerical data obtained from the population to analyze the selected research sample. Quantitative analysis involves the computation of data to assess its significance, as opposed to interpret the data to aid the writer (Dianna, 2020).

For that reason, this research will rely on quantitative analysis and the data will be processed using IBM SPSS 25.0. Furthermore, the present study applies multiple linear regression analysis as the chosen statistical method for evaluating the formulated hypothesis. Multiple regression analysis allows for the objective evaluation of the extent and characteristics of the association between the independent variables and the dependent variable. The regression coefficients provide an indication of the relative importance of each independent variable in predicting the dependent variable (Sekaran & Bougie, 2016). Moreover, in order to effectively substantiate its hypothesis, this study will employ descriptive statistical analysis, the classical assumption test, multiple linear regression, and hypothesis testing.

3.5.1 Descriptive Statistics

Descriptive statistical methods refer to the approaches employed in arranging, elaborating, and helpfully organizing data. These techniques involve the use of frequency distributions to organize and summarize data, as well as the use of charts and graphs to display frequency distributions (Lind et al., 2020). In this research, the descriptive statistics will cover maximum, minimum, mean and standard deviation.

3.5.2 Classical Assumption Test

The Best Linear Unbiased Estimator (BLUE) is derived by employing the classical assumption test to identify the most suitable estimator that fulfills the requirements of linearity and unbiasedness. Conducting standard hypothesis testing is a crucial step that should be undertaken before constructing a reliable multiple regression model. According to Ghozali (2018), the fulfillment of the normality test for the regression, in the absence of data, indicates the attainment of the Best Linear Unbiased Estimator (BLUE).

3.5.2.1 Normality Test

A normality test is performed to assess whether the residuals in a regression model exhibit a normal distribution or not. One potential approach is to utilize a regression model that assumes a normal distribution. The need for this arises from the underlying assumption of residual value normality in both t-tests and F-tests. If a particular assumption is not met, statistical tests may lose their validity (Ghozali, 2018). There are two methods available for conducting a normality test, namely:

1. The analysis of graphs

   The assessment of normality through graphical analysis can be conducted by examining a histogram. A dataset is deemed to adhere to a normal distribution when its corresponding histogram exhibits a symmetrical bell-shaped distribution without any skewness. This statement implies that the data demonstrates a uniform distribution and adheres to the normality assumption of the regression model. A histogram that exhibits either positive or negative skewness implies the presence of data that deviates from the norm.

   In addition to the histogram, the normal probability plot can serve as a tool for assessing the normality of a provided dataset. The utilization of the dispersion of data points along the diagonal axis of a graph can aid in the process of determination. The presence of diagonal dispersion in data is a defining feature of a normal distribution and satisfies the normality assumption of the regression model. Data that exhibits a random distribution and lacks adherence to a linear pattern is regarded as anomalous.

2. The Application of Statistical Analysis

To enhance the dependability of the findings, a statistical analysis will be performed. The determination of normality in residual data can be accomplished by employing a non-parametric statistical test, such as the Kolmogorov-Smirnov test. The determination of the normal distribution of a regression model in statistical analysis can be achieved by employing a significance level of 5% (0.05). If the resulting significance level (Asymp.Sig) exceeds 0.05, it can be concluded that the regression model exhibits normal distribution. If the Asymp.Sig value is less than 0.05, it can be inferred that the regression model does not exhibit a normal distribution.

3.5.2.2 Heteroscedasticity Test

The purpose of the heteroscedasticity test is to examine whether there is unequal variance among the residuals of different observations in a regression model. Homoscedasticity is a statistical concept that pertains to the consistency of variance in the residuals between different observations. Heteroscedasticity refers to the situation where the variance of a variable is not constant. For a regression model to be considered acceptable, it should exhibit homoscedasticity. Homoscedasticity refers to the presence of consistent residual variance across all observations. There are multiple methodologies available for detecting heteroscedasticity, including:

1. Scatterplot

   The current approach involves examining a scatterplot diagram that displays the relationship between the dependent variable, ZPRED, and the residual, SRESID. Heteroscedasticity is not present in the scatterplot as there is no discernible pattern and the data points are evenly dispersed above and below the zero point on the Y-axis. Heteroscedasticity refers to the occurrence of a discernible pattern in the scatterplot. The presence of surging, broadening, or narrowing tendencies in the dots indicates the occurrence of a specific pattern (Ghozali, 2018).

2. Statistical test

   The purpose of the heteroscedasticity test is to examine whether there are varying levels of variability among the residuals of different observations in a single regression model. Homoscedasticity refers to a statistical concept that describes a scenario in which the variability of the residuals remains consistent across various instances of monitoring. According to Ghozali (2018), heteroscedasticity, or the lack of homoscedasticity, occurs when there is variation in the residuals' variance. The current study utilized the Glejser test, which involves regressing the independent variable with its residual absolute value. In the opinion of Ghozali (2018), if the significance level of the correlation result is below 0.05 (5%), it can be concluded that the regression equation demonstrates heteroscedasticity. When the correlation outcome exceeds the threshold of 0.05 (5%), the equation is deemed to show neither heteroscedasticity nor non-heteroscedasticity.

3.5.2.3 Multicollinearity Test

Multicollinearity occurs when there is a correlation between independent variables. The existence of correlated independent variables presents difficulties when trying to make conclusions about individual regression coefficients and their influence on the dependent variable (Lind et al., 2020). Multicollinearity refers to a statistical occurrence where there is a strong correlation between two or more independent variables in a regression model. This correlation makes it difficult to accurately estimate the regression coefficients (Sekaran & Bougie, 2016). The current research utilizes tolerance value and Variance Inflation Factor (VIF) as methods for detecting multicollinearity. These two metrics are used to assess the degree to which independent variables influenced by other independent variables.

According to Ghozali (2018), there is a negative correlation between tolerance and VIF. Research decisions are typically made based on specific criteria. In this case, two criteria are considered: a tolerance value greater than 0.1 (>0.1) and a Variance Inflation Factor (VIF) lower than 10 (<10). These criteria are used to determine whether there is multicollinearity or correlation among the independent variables. On the other hand, a research study that demonstrates multicollinearity will produce a tolerance value below 0.1 (<0.1) and a Variance Inflation Factor (VIF) above 10 (>10), suggesting the existence of correlation among the independent variables.

3.5.2.4 Autocorrelation Test

A statistical technique called the autocorrelation test is used to determine whether there is a link between the values of similar variables in the linear regression model over successive time intervals. Autocorrelation is a phenomenon that manifests when a correlation is present between consecutive observations, a common attribute observed in time series data (Ghozali, 2018). The existence of autocorrelation within a regression model is considered unfavorable and should be addressed. The Durbin-Watson (D-W) test is utilized in this study to determine the existence of autocorrelation. The Durbin Watson test is employed to determine the presence of first-order autocorrelation in a regression model. This test necessitates the inclusion of an intercept in the model and exclusion of lag variables among the independent variables. The assessment is predicated upon the subsequent criteria:

Table 3.2 Requirement of Durbin-Watson

If	Hypothesis 0	Decision
$0 < dw < dl$	Positive autocorrelation does not exist	Rejected
$dl \leq dw \leq du$	Positive autocorrelation does not exist	No decision
$4-dl \leq dw \leq 4$	Negative autocorrelation does not exist	Rejected
$4-du \leq dw \leq 4-dl$	Negative autocorrelation does not exist	No decision
$du \leq d \leq 4-dl$	There is no autocorrelation, either positive or negative.	Accepted

Source : Ghozali (2018)

3.5.3 Multiple Linear Regression Analysis

A statistical procedure called multiple linear regression is used to assess how different variables affect just one variable. The independent variable is typically represented as the variable that has a causal effect, while the dependent variable is typically represented as the variable that is affected by the independent variable. According to George & Mallery (2019), multiple linear regression analysis is a statistical technique that involves using two or more independent variables to understand and explain the variability of a selected dependent variable. The regression coefficients provide valuable insights into the relative importance of each independent variable in predicting the dependent variable. The investigation plans to employ Multiple Linear Regression analysis to examine the effects of profitability, leverage, and liquidity. The current study employs a subsequent regression model of:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

in which:

Y = Dividend Policy (Dividend Payout Ratio)

α = Constanta

X_1 = Profitability (Return on Assets)

β_1 = Variable regression coefficients of Profitability

X_2 = Leverage (Debt to Assets Ratio)

β_2 = Variable regression coefficients of Leverage

ε = Residual (error)

3.5.4 Hypothesis Testing

The utilization of statistical methods is integral to the process of hypothesis testing, as it enables the assessment of the proposed hypothesis. The statement mentioned above refers to a methodology that is employed to assess the credibility of a hypothesis. This methodology involves determining whether the hypothesis is acceptable and should be accepted, or whether it is unacceptable and should be rejected (Lind et al., 2020). To investigate the formulated hypothesis, this study utilizes three different types of tests to assess the association between the independent and dependent variables, both individually and collectively. The tests encompassed in this analysis comprise the t-test, which is employed to evaluate the significance of individual or partial effects, the F-test, which is utilized to assess the simultaneous significance of multiple effects, and the coefficient of determination test, specifically the R^2.

3.5.4.1 Partial T-Test

The t-test is a statistical technique employed to assess the extent to which a particular independent variable account for the observed variability in the dependent variable. The t-test is a statistical method used to assess the influence of individual independent variables on the dependent variable. Based on empirical research conducted by Ghozali (2018), it has been identified that t-tests rely on two distinct decision-making frameworks:

1. The t-value derived from the test (t-count) should be observed.

In the context of statistical hypothesis testing, the null hypothesis (H0) is deemed invalid if the computed t-value (t-count) surpasses the critical value specified in the t-distribution table, or if the magnitude of the t-count is smaller than the magnitude of the critical value provided in the t-distribution table. This finding suggests that the independent variable exerts a statistically significant partial influence on the dependent variable. If the t-value is less than or equal to the critical t-value from the t-table, or if the absolute value of the t-value is greater than or equal to the absolute value of the critical t-(t-count ≤ t-table or –t-count ≥ -t-table), it can be inferred that the null hypothesis (H_0) is accepted. This suggests that the independent variable does not have a statistically significant partial impact on the dependent variable.

2. Assess the level of statistical significance.

If the statistical significance level of an independent variable, as determined by the test, is found to be less than 0.05 (Sig. < 0.05), it can be inferred that the dependent variable is significantly influenced by the independent variable. When the significance level (Sig. > 0.05) of a specific independent variable, as determined by the test, is greater than 0.05, it can be concluded that the independent variable does not have a statistically significant partial effect on the dependent variable.

Moreover, a positive beta coefficient (β) indicates a positive correlation between the independent variable and the dependent variable. In contrast, when the beta (β) value is negative, an inverse relationship can be observed between the independent variable and the dependent variable (George & Mallery, 2019).

3.5.4.2 Simultaneous F-Test

To determine if several independent factors have an overall impact on the dependent variable, simultaneous hypothesis testing is a statistical technique used. Additionally, the current study employs the F-Test Statistic to conduct simultaneous hypothesis testing, adhering to the prescribed standards:

1. If the Fcount value exceeds the Ftable value or if the significance level (Sig) is below 0.05, it is possible to accept the alternative hypothesis (H_1), which suggests a significant influence of the independent variables on the dependent variable concurrently.

2. If the value of the F value (Fcount) is lower than the critical value of the F distribution (Ftable), or if the significance level (Sig) exceeds 0.05, then the null hypothesis (H_0) will be accepted. According to Ghozali (2018), the findings suggest that the simultaneous consideration of the independent variables does not yield a statistically significant impact on the dependent variable (Ghozali, 2018).

3.5.4.3 Coefficient of Determination

According to Lind et al. (2020), the Coefficient of Determination pertains to the proportion of the total variability observed in the dependent variable that can be explained or ascribed to the variability in the independent variable. Sekaran & Bougie (2016) suggest that the coefficient of determination is a statistical measure utilized to assess the extent to which a regression line effectively represents the actual data points within a regression model. The coefficient determinant, denoted as R^2, is bounded by the interval 0 and 1, where 0 represents the lower bound and 1 represents the upper bound ($0 \leq R^2 \leq 1$). When the determinant of the coefficient matrix is equal to zero, it indicates the absence of a correlation between the independent variable(s) and the dependent variable(s). Nevertheless, when the coefficient value is equal to one, the independent variables possess the capacity to predict the variability of the dependent variable. A low coefficient of determination (R^2) suggests that independent variables have limited explanatory power related to the observed variability in the dependent variable. Conversely, a value approaching unity signifies that the independent variables provide a significant amount of the necessary data for forecasting variations in the dependent variable (Ghozali, 2018).