CHAPTER III
RESEARCH METHOD

3.1 Research Method

The method of research used in this study is quantitative research. According to Goertzen (2017), “quantitative research” is mainly concerned with collecting and analysing data that is structured and can be represented numerically. The goals of quantitative research is to build accurate and reliable measurement to allow for statistical analysis.

Because quantitative research goal is to focus on data that can be measured, it will be very effective when given questions which answers are “what” or “how” of a given situation. Questions used are direct, quantifiable, and often contain phrases such as “what percentage?”; “what proportion?”; “to what extent?”; “how many?”; “how much?”.

3.2 Population and sample

Population is the collection of elements that consist of objects and subjects which can be observed by the researcher. Umar & Madugu (2015) stated that population is a collection of survey elements. In this research, the population will be the people who have been conducting transactions in Online Shopping services.

According to Majid et al (2018), Sampling is the process of selecting a statistically representative sample of individuals from the population of interest. The researcher determines the sample of this research by using the method of
Purposive sampling is one of the non-probability sampling techniques where the researcher determines the sampling by establishing specific characteristics that are in accordance with the research objectives so that it is expected to answer the research problems. The characteristics of the purposive sampling will be based on certain characteristics, such as:

1. Respondents who have experience of buying from online platforms.
2. Respondents are based in Surabaya.
3. Respondents have experience in making use of the discounts and promotions on a certain e-commerce platform.

Because the number of the population is indefinite, so to understand and determine the number of the population will be using the Unknown Population method. The formula used is as follows:

$$n = \frac{Z^2}{4\mu^2}$$

Details:
- \(n\) = Sample size
- \(Z\) = The sample confidence level needed in the research, with \(a = 5\% \) (percentage of confidence set is 95\%), so \(Z = 1.96\)
- \(\mu\) = Margin of error, the tolerable error level. (set on 10\%)

Based on the previous formula, the results will be:

$$n = \frac{1.96^2}{4(0.1)^2} = 96.04 \text{ (rounded to be 97)}$$

As stated by the formula above, the number of the sample will be 100 people, with the requirement of 97 accepted responses.
3.3 Data Collection Methods

3.3.1 Questionnaire

As explained by Bolarinwa (2017), the questionnaire is a tool for collecting and recording information by delivering a predetermined set of questions to the respondents. Data sources are obtained from primary data and secondary data. Primary data is obtained from the answers to the questionnaires shared with the ones concerned. The questionnaire uses a Likert scale as the level of answer given by the respondent. Likert scale is used to obtain or discover opinions, attitudes, and perceptions about the phenomenon of the research. The research indicators are as follows:

1. (STS): With the value of 1
2. (TS): With the value of 2
3. (N): With the value of 3
4. (S): With the value of 4
5. (SS): With the value of 5

3.4 Validity and Reliability

3.4.1 Validity

The definition of validity, as explained by Heale (2015), is a concept which is accurately measured in a quantitative study. The item measured will be able to be stated as valid if the existing correlation is significant with the total scores. If the correlation significance is < 0.05, then the item is stated as valid with a confidence level of 95%, as stated by Pearson (2016).
3.4.2 Reliability

According to Taherdoost (2016), the purpose of reliability is to measure a phenomenon which will provide stable and consistent results. Testing for reliability is important because it refers the consistency across the parts of a measuring instrument. The most commonly used is the Cronbach Alpha coefficient because it works well with Likert scales. The criteria will be considered reliable if the coefficient is $> 0.6$, if it’s $< 0.6$, then it is not reliable.

3.5 Operational Definition of Variables

Table 3.1 Operational Definition of Variables

<table>
<thead>
<tr>
<th>No.</th>
<th>Variables</th>
<th>Definition</th>
<th>Indicator</th>
</tr>
</thead>
</table>
| 1   | Discount       | Discount itself has many variations of meanings and used in many kinds of things. According to a research conducted by Sudrajat (2015), Discount is a price reduction given by the seller to the buyer as a token of appreciation of a certain activity from the buyer that pleases the seller. | 1. Quantity discount  
2. Seasonal Discount  
3. Cash Discount  
4. Trade Discount (Sudrajat 2015) |
| 2   | Promotion      | According to Kotler (2018), promotion is a compass for all the tools in the marketing mix whose major role is as a mean of persuasive communication. | 1. Advertising  
2. Sales promotions  
3. Direct promotions (Kotler 2018) |
| 3   | Purchase Intention | Purchase intention is a kind of decision-making that studies the reason to buy a particular brand by consumer (Shah et al., 2016). Morinez et al. (2007) defined purchase intention as a situation where the consumer tends to buy a certain product in a certain condition. Consumers’ purchase intention and decision are complex processes where consumer access and evaluate a certain product which will determine the future purchase. | 1. Attention  
2. Interest  
3. Desire  
3.6 Data Analysis Method

3.6.1 Classical Assumption Test

3.6.1.1 Normality Test

According to Deliar (2016), this test is used to determine whether or not the distribution of data is normal. A good research data is data that has a normal distribution. The researcher uses the Kolmogorov-Smirnov method where the data can be defined as a normal distribution residual if the significance level is > 0.05.

3.6.1.2 Multicollinearity Test

According to Deliar (2016), this test is used to determine the existence of significant high correlation in a multiple regression model between the independent variables. If the value of Variance Inflation Factor (VIF) is < 10, then there is no multicollinearity. However, if the value is > 10, then there is multicollinearity and the variable have to be removed.

3.6.1.3 Heteroscedasticity Test

As explained by Deliar (2015), in this test the residual variants are not the same for all observations in the regression model. Good regression should not have heteroscedasticity. The method used is the graph method by choosing the dots pattern on the regression chart.

3.6.1.4 Autocorrelation Test

According to Deliar (2015), autocorrelation is a correlation between members of the observations arranged according to time or place. In a good registration model, autocorrelation should not occur. This test method is using the Durbin Watson test (DW test) as stated below:
<table>
<thead>
<tr>
<th>DU &lt; DW &lt; 4-DU</th>
<th>There is no autocorrelation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DW &lt; DL or DW &gt; 4-DL</td>
<td>There is autocorrelation</td>
</tr>
<tr>
<td>DL &lt; DW &lt; DU or</td>
<td>There is uncertainty or valid conclusion</td>
</tr>
<tr>
<td>4-DU &lt; DW &lt; 4-DL</td>
<td></td>
</tr>
</tbody>
</table>

### 3.6.1.5 Linearity Test

Still according to Deliar (2015), Linearity test is used to determine the linearity of the data, which is whether two variables have a linear relationship or not. In this test, if the significance level is <0.05, then the relation between the dependent and independent variables are linear.

### 3.6.2 Validity

This multiple linear regression analysis is used to determine the effect of linear relationships between two or more variables with independent variables and one dependent variable. This research uses an equation as follows:

\[ Y = a + b_1x_1 + b_2x_2 + e \]

Explanation:

- **Y** = Purchase Intention
- **a** = Constant
- **b1, b2** = Regression Coefficient
- **x1** = Discount
- **x2** = Promotion
- **e** = standard error
3.6.3 Hypothesis Test

3.6.3.1 Simultaneous Test (F test)

Salehi (2018), said that a simultaneous regression coefficient test (F Testing), has the purpose to test the significance of the effect of the independent variables towards the dependent variables. This test uses a significance level of 0.05.

3.6.3.1 Partial Test (T test)

The T-test shows how deep the impact of an independent variable towards a dependent variable. If the significance of t value is < 0.05, then the independent variable does partially impact the dependent variable.

3.6.3.1 Coefficient Correlation (R) & Coefficient Determinant Test (R²)

Pearson (2015) stated that (R) is a multiple correlation, which is a correlation between 2 or more independent variables towards dependent variables. The Value of R ranges from 0 to 1. If the value is close to 1, then the relationship between the variables will get stronger. While (R²) is used to show the determinant coefficient. This number will be converted to the percentage form which means contribution percentage of the influence of independent variables toward dependent variables.