

CHAPTER 1

INTRODUCTION

1.1 Overview of Research Object

William Tanuwijaya and Leontinus Alpha Edison started Tokopedia. At that point, they encountered comparable life events and obstacles; to end the issue, they had to act to close the gap between small villages and large cities. They started Tokopedia on August 17, 2009 with the goal of digital economic equality. William Tanuwijaya wished to create a firm that survived its creator. The torch has now been handed to Melissa Siska Juminto, who began her career as an accountant at Tokopedia and now runs the firm as President of E-Commerce Tokopedia. Motivating Indonesians to begin selling online has helped us to drive digital economic equality for the last 14 years. Building a platform that lets everyone start and find anything. Tokopedia logo can be seen in figure 1.1.



Figure 1. 1 Tokopedia Logo

(Source : Tokopedia)

Like other e-commerce sites, Tokopedia first entered the Indonesian market via a Consumer-to-Consumer (C2C) approach, letting people purchase and sell goods straight to one another. But as Tokopedia expanded, it moved toward a more hybrid approach combining components of Business-to-Consumer (B2C) and Business-to-Business (B2B) services. This development allowed collaborations with well-known companies and brands, hence allowing them to run official stores and provide goods to a more wide consumer base. Tokopedia envisions to democratize access to goods and services all throughout Indonesia, thereby generating equal chances for people and companies of all sizes. Their goal is to

empower local enterprises by means of a platform providing extensive product choice, smooth transactions, and integrated services including financial solutions and shipping. Tokopedia is dedicated to building a robust digital ecosystem that enables small and medium businesses (SMEs) to expand and undergo digital transformation. Tokopedia is also committed to improving consumer experiences by means of cutting-edge technologies, guaranteeing fair prices, and promoting sustainability by supporting local business and job creation. By means of these projects, Tokopedia hopes to increase its importance as a major participant in Indonesia's digital economy and a dependable partner for companies and consumers both (Tokopedia.com).

1.2 Research Background

Nowadays, it is not considered taboo for Indonesians to use electronic media or the internet. The majority of people actually can't live without social media and the internet. Therefore, many businesses use this circumstance to increase their business's efficacy and efficiency. One of the key foundations of e-marketing, which has drastically transformed company, is online trading platforms. Among the many digital marketing strategies include SEO (Search Engine Optimization), online advertising platforms like Google Ads, print media promotion, radio and television advertising, digital signage, email marketing, mobile marketing, and e-commerce.

Technological advances in this era are experiencing rapid growth, especially in the business sector. One form of technological innovation in the business field is the presence of e-commerce. APJII (Association of Indonesian Internet Service Providers) explained that the number of Indonesian internet users in 2024 reached 221,563,479 people from a total population of 278,696,200 people in Indonesia in 2023. In addition, from the results of the 2024 Indonesian internet penetration survey released by APJII, the Indonesian internet penetration rate touched 79.5%. Compared to the previous period, there was an increase of 1.4%. Especially with the existence of e-commerce which is a platform to buy and sell online.

The internet is an open global communications network and connects many computer networks of various types and types, using communication types such as

telephone, satellite and so on. (Mohammad, 2021). Internet in a broad sense and explained in a global sense, the definition is a network of computers around the world which are all interconnected using the standard internet protocol suite to be able to access information and also be able to exchange information and data with each other (Nurbaiti & Alfariysi., 2023).

According to (Rosario & Raimundo, 2021), e-commerce is the sale and purchase of goods and services over the internet in exchange for money and data transfer to complete the transaction. According to Statista, The following is data on the number of e-commerce users in Indonesia can be seen in figure 1.2

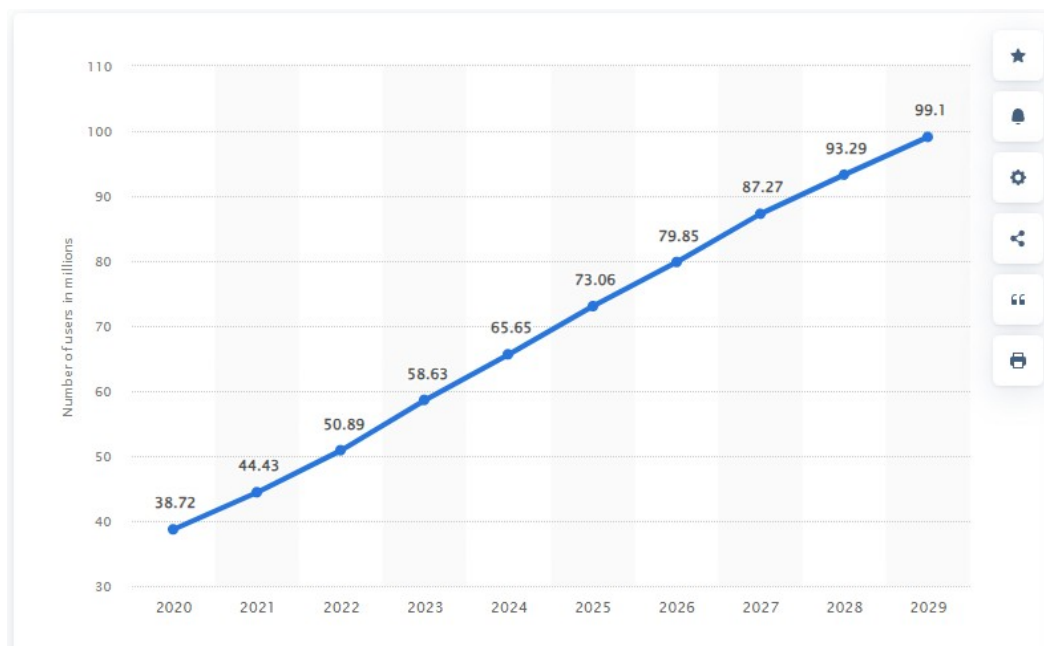


Figure 1. 2 Number of users of e-commerce in Indonesia from 2020 to 2029

(Source : <https://www.statista.com>)

According to Statista, The number of users in the e-commerce market in Indonesia was forecast to continuously increase between 2024 and 2029 by in total 33.5 million users (+51.03 percent). In Indonesia, there are currently many e-commerce service providers, such as Tokopedia, Blibli, BukaLapak, Lazada and Tokopedia. Tokopedia itself is one of the most visited e-commerce service providers

in 2023. The following is data on the 5 e-commerce with the most visitors in Indonesia (1st Quarter 2023)



Figure 1. 3 5 E-commerce with the most visitors in Indonesia (1st Quarter 2023)

(Source : <https://databoks.katadata.co.id>)

Figure 1.3 identifies the five most visited online business platforms in Indonesia in 2023, namely Shopee, Tokopedia, Lazada, Blibli and Bukalapak. The data states that Tokopedia is the e-commerce with the highest number of visitors in Indonesia in that period, with a total of 117,03 million visitors. Reporting from the (Tokopedia.com) in marketing, Tokopedia tries to maximize each marketing element including product, price, promotion, and place by prioritizing the experience that prospective buyers will get.

Marketplace dengan Kualitas Produk Terbaik (2022)

databoks

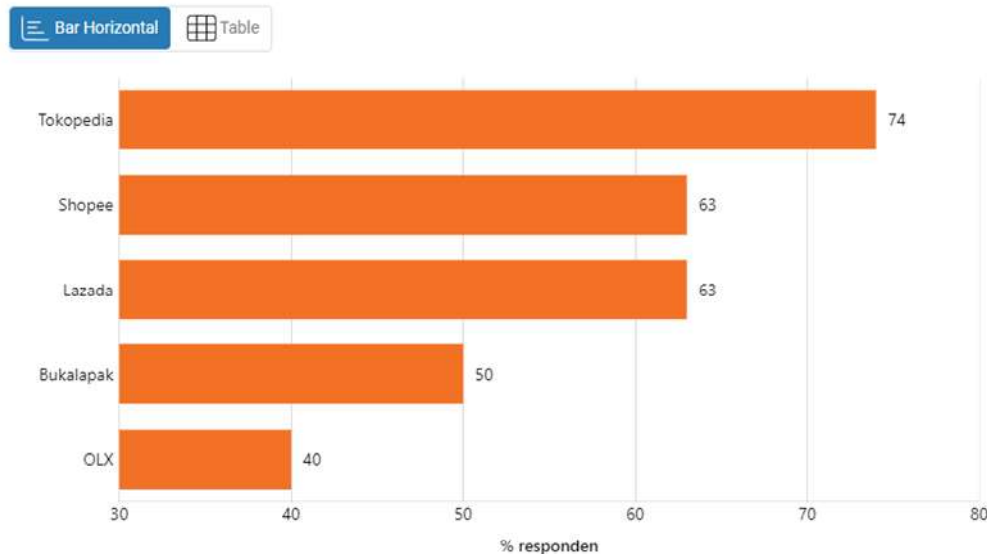


Figure 1. 4 Marketplace with the best product quality
(Source : <https://databoks.katadata.co.id>)

74% of respondents believe that Tokopedia's product quality is superior to that of other e-commerce sites, according to study done by Institut Public de Sondage d'Opinion Secteur (Ipsos) on the Databoks page (2022) of 1,051 respondents in 15 major Indonesian cities. With 63% of respondents concurring that Tokopedia's product quality is superior to that of other e-commerce sites, Tokopedia comes in first in this category. This will certainly have an impact on consumer purchase intention. According to research conducted by (Li et al., 2022) suggests that consumers prefer products that have better quality user generated content. Therefore, Tokopedia needs to improve and maintain the quality of UGC items that will be marketed to customers so that they can attract more customers and increase customer satisfaction.

E-commerce sells many products, one of which is skin care. As reported by (Kompas.id, 2024), the skincare trend has become part of the younger generation's lifestyle, signaling that skincare is not just a routine, but has become a pillar of appearance awareness and part of youth culture. The news explains that young

people start paying attention to their skincare from their teenage years. They use different skincare products to maintain healthy skin, boost their confidence and take care of their appearance. This awareness also motivates them to choose products more carefully and to make use of the product recycling facilities offered by some beauty stores.

There are already many types of local skin care products circulating in Indonesia, such as Somethinc, Scarlett, Avoskin and others. With formulas designed to meet the skin needs of Indonesians, especially given the tropical climate, these products may be more suitable compared to products from countries with colder climates (Tempo.co, 2024). Sales of local skincare brands are also quite impressive. From April to June 2022, total sales on marketplaces reached Rp 292.4 billion, with 3.8 million transactions. Supported by data from (statista.com), revenues in the 'Skincare' segment of the beauty and personal care market in Indonesia are forecast to continue to increase between 2024 and 2029, totaling US\$549.1 million (+19.88%). After nine consecutive years of growth, revenues are expected to reach US\$3.3 billion, peaking in 2029. In particular, the 'Skincare' segment of the beauty & personal care market has been growing steadily in recent years.

Revenue of the skin care market in Indonesia from 2020 to 2029 (in million U.S. dollars)

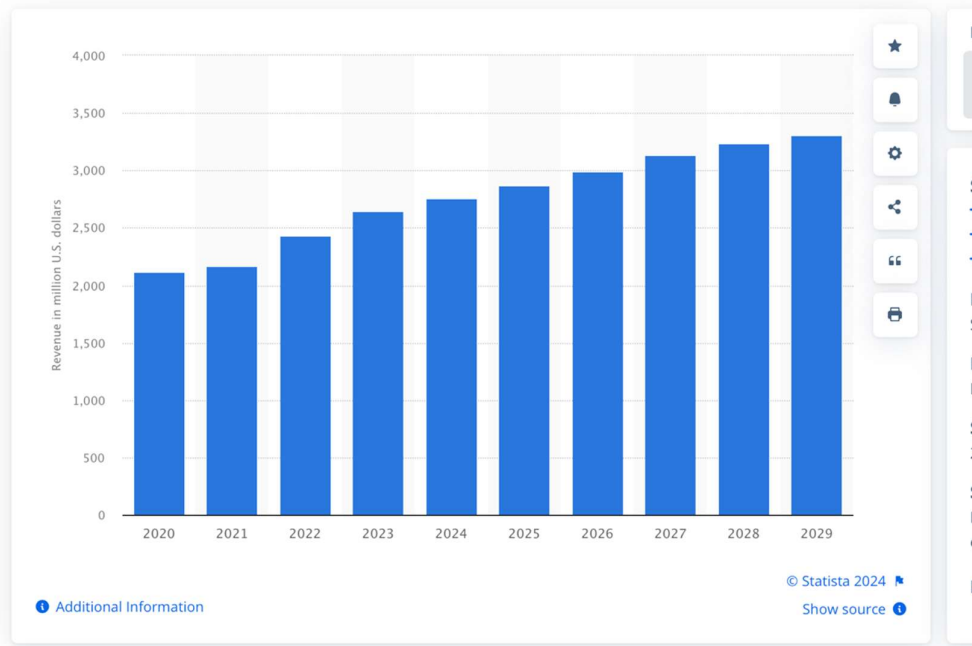


Figure 1. 5 Revenue of the skin care marker in Indonesia from 2020 to 2029

(Source : <https://www.statista.com>)

As per the report (compas.co.id, 2022) that can be seen in figure 1.5, the sales of local skin care brands have been quite impressive. From April to June 2022, the total sales in the market reached IDR 292.4 billion with a total number of transactions of 3.8 million. Supported by a statement reported by (Tempo.co, 2024) that the top 5 skincare brands that have been circulating in Indonesia are Somethinc, Scarlett, MS Glow, Avoskin, Whitelab.

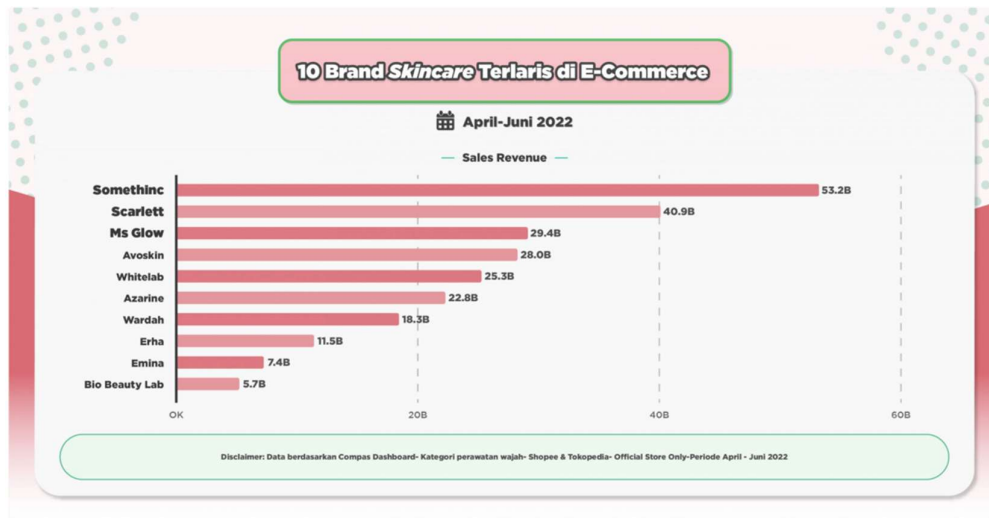


Figure 1. 6 Top 10 best selling skincare brands on e-commerce

(Source : <https://kompas.co.id/>)

User-Generated Content (UGC) is defined as content generated, shared and propagated by users through digital platforms. UGC includes various forms of unstructured data, such as product reviews, opinions, and evaluations published by consumers after consuming a product or service on an e-commerce platform (Li et al., 2022). Tokopedia has actively utilized UGC as part of its platform strategy. The platform encourages users to provide reviews, ratings, and comments after making a purchase, allowing other potential buyers to gain insights into the product's quality and performance. Additionally, Tokopedia frequently promotes UGC through campaigns that incentivize users to share their experiences, such as review rewards or rating-based discounts, thereby enriching the pool of UGC available on the platform. This strategy not only enhances consumer trust but also supports Tokopedia's recommendation system by using UGC to refine product suggestions and improve the overall shopping experience.

This is of course in line with what was stated by (Li et al., 2022) that UGC has an important role in the context of e-commerce. UGC becomes a credible source of information for consumers throughout the purchase cycle, helping to reduce the uncertainty and risk that consumers may feel when making purchase decisions. UGC data, such as customer reviews and opinions, can attract consumer attention,

encourage consumer activity, and increase interest in making purchases. according to (Antaranews.com, 2024) there was a drastic increase in sales after using Tokopedia's interactive features. The role of the Tokopedia Live feature is huge for overall revenue, where 40 percent of total revenue in the first semester of 2024 came from sales on Tokopedia Live, e-commerce platforms like Tokopedia utilize UGC to enrich its recommendation system algorithm. UGC in the form of reviews and ratings from consumers allows Tokopedia to suggest relevant products based on other users' preferences.

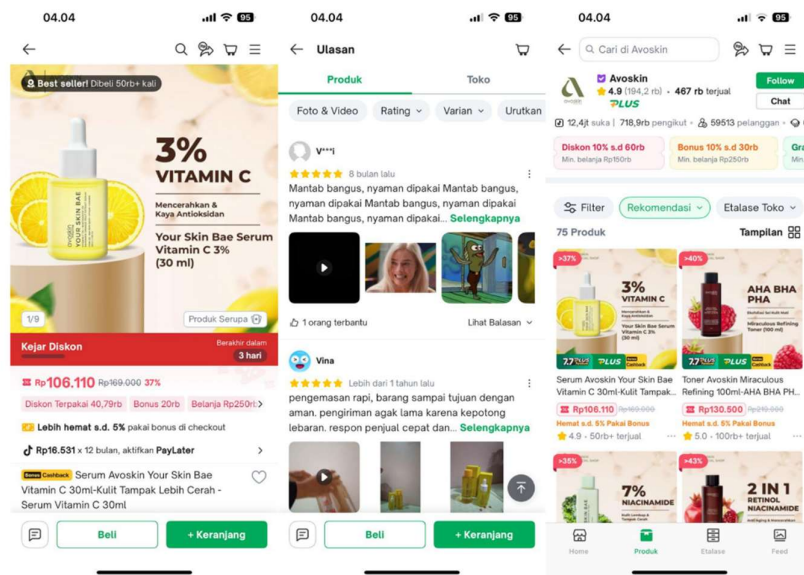


Figure 1. 7 Example of UGC in Tokopedia
(Source : Avoskin Official Store Tokopedia)

Picture 1.7 shows the increased visibility of skincare products on Tokopedia, which is influenced by the amount and quality of User-Generated Content (UGC) such as reviews and ratings from consumers. With consistently high ratings and positive reviews, skincare products appear more frequently in Tokopedia recommendations, which increases the chance of being seen by potential consumers. This not only strengthens consumer trust, but also contributes to

increased product sales on Tokopedia, making UGC a key element in a digital marketing strategy that focuses on data-driven recommendations.

Transactions have driven rapid development in digital industries, particularly in the beauty and skincare sector. As one of Indonesia's largest e-commerce platforms, Tokopedia plays a significant role in facilitating consumer access to local skincare brands. In an increasingly competitive market, companies must ensure that all aspects of customer satisfaction are addressed while identifying which service dimensions still require improvement. One effective approach to measuring customer satisfaction is by analyzing user-generated content (UGC) in the form of product reviews and feedback provided by consumers directly on the Tokopedia platform. These reviews reflect real user experiences and can serve as valuable indicators of both strengths and weaknesses in the provided services. This method can help companies better understand customer perceptions, evaluate service performance, and develop strategic improvements to meet user expectations. Ultimately, customer satisfaction can strongly influence user loyalty and purchasing behavior.

To support this research, the author has examined previous studies related to customer satisfaction as well as the implementation of text classification and topic modeling methods. (Sudirjo et al.,2023) conducted a study analyzing the effects of price, promotion, and service quality on customer satisfaction using survey methods, and found that these factors significantly influence satisfaction and loyalty. Similarly, (Fadli & Rubiyanti,2021) investigated the same variables on the Shopee application using Structural Equation Modeling (SEM) with 100 respondents, and the results also indicated a significant relationship between price, promotion, and service quality with customer satisfaction. Another study by (Mahfudiyah & Alamsyah.,2022) applied sentiment analysis and topic modeling using the IndoBERT model on Gojek app reviews. The study reported a sentiment classification accuracy of 96%, with negative perceptions dominating, particularly in the service quality dimension. In addition, (Mulianti et al.,2024) employed Support Vector Machine (SVM) and SMOTE oversampling techniques to classify

sentiment in Alfagift app reviews. The study achieved 90% accuracy on imbalanced data, which improved to 92% after balancing the dataset.

Further, (Firdaus et al.,2025) conducted aspect-based sentiment analysis on the Alfagift application using the Random Forest algorithm and Latent Dirichlet Allocation (LDA). The results revealed that negative sentiment was dominant, especially in the application aspect (833 data points), followed by delivery (455), pricing (166), and product quality (93). This indicates that sentiment analysis and topic modeling can effectively identify and evaluate e-satisfaction dimensions. (Kumar et al.,2023) explored customer satisfaction in grocery applications using LDA and identified eight key e-satisfaction dimensions: app experience, customer support, delivery, offers & discount, offline competition, payment refunds, product quality, and stock availability. Their findings showed that offers & discount had the strongest influence on satisfaction, while product quality had the least.

Building on these studies, the present research adopts the BERT algorithm for sentiment classification and the BERTopic model for topic extraction, as these methods have been proven to deliver higher accuracy than conventional algorithms (Alfarobby & Irawan, 2024). The dataset used consists of customer reviews collected from the official Tokopedia stores of five leading local skincare brands namely Somethinc, Scarlett, MS Glow, Avoskin, Whitelab, providing a reliable representation of user perceptions regarding product quality and service performance. Through this approach, the research aims to uncover the dominant dimensions that influence customer satisfaction and identify the most discussed issues in user feedback. Based on this background, the title of this study is: **"Analysis of Customer Satisfaction in Purchasing Skincare in the Tokopedia Application Using Text Classification and Topic Modeling."**

1.3 Problem Formulation

This study explores the sentiment distribution and thematic focus found in user-generated content (UGC) related to skincare purchases on the Tokopedia platform. UGC, which includes customer reviews, ratings, and comments, serves as a vital source of real-world insight into user experiences and satisfaction levels.

By analyzing this content, the study aims to determine the proportion of positive and negative sentiments expressed by users in their reviews of local skincare products. Understanding this sentiment distribution provides a measurable indication of general customer satisfaction toward Tokopedia's skincare offerings.

In addition to sentiment classification, this study also focuses on identifying which dimensions of electronic satisfaction (E-Satisfaction) are most frequently discussed in customer reviews. These dimensions such as delivery, product quality, offers & discount, and customer support serve as critical factors in shaping the overall shopping experience. By categorizing reviews into these dimensions, the research aims to map out which areas of Tokopedia's service are most recognized, valued, or criticized by users.

Further, this study seeks to determine which E-Satisfaction dimensions exhibit the highest and lowest levels of sentiment polarity. Specifically, it examines where positive sentiments are most concentrated and which areas receive the greatest volume of negative feedback. This distinction is essential for pinpointing which service aspects contribute most to customer satisfaction, and which dimensions require improvement to enhance the shopping experience.

Lastly, this study applies topic modeling techniques to uncover the dominant themes present in the most positively and negatively reviewed dimensions particularly focusing on the delivery experience, which appears frequently in both sentiment categories. Identifying key topics within these reviews allows a deeper understanding of what specific issues, experiences, or expectations users frequently mention, thereby providing actionable insights for platform improvement.

1.4 Research Question

1. What is the proportion of positive and negative sentiments found in user reviews related to skincare purchases on the Tokopedia?
2. Which E-Satisfaction dimensions are most frequently mentioned by users in reviews about skincare products on Tokopedia?

3. In which E-Satisfaction dimensions are the most positive and most negative sentiments observed regarding skincare purchases on Tokopedia?
4. What are the key topics that emerge from the most positive and most negative user reviews concerning skincare shopping experiences on the Tokopedia?

1.5 Research Objective

1. To analyze the proportion of positive and negative sentiments most frequently expressed in user reviews related to skincare purchases on the Tokopedia application.
2. To identify the E-Satisfaction dimensions most commonly mentioned by users in reviews of skincare products on Tokopedia.
3. To evaluate which E-Satisfaction dimensions receive the most positive and most negative sentiments in the context of skincare shopping on Tokopedia.
4. To uncover the key topics that emerge from user reviews with positive and negative sentiments regarding the skincare shopping experience on Tokopedia.

1.6 Research Benefits

1.6.1 Academic Benefits

This research contributes to the academic field of e-commerce and digital marketing by providing empirical insights into customer satisfaction analysis through advanced Natural Language Processing (NLP) methods. By implementing text classification using the BERT algorithm and topic modeling with BERTopic, this study offers a methodological reference for future research in sentiment analysis and user experience evaluation, particularly in the context of user-generated content (UGC) on e-commerce platforms.

1.6.2 Practical Benefits

The findings of this study can assist Tokopedia and other e-commerce platforms in identifying key satisfaction dimensions and understanding customer sentiment trends related to skincare products. By analyzing which aspects users are

most satisfied or dissatisfied with such as delivery, product quality, or promotional offers platform managers and brand owners can make data-driven decisions to improve services and enhance customer satisfaction.

1.7 Systematization of Final Project Writing

Contains a systematic and concise explanation of the research report consisting of Chapter I to Chapter V in the research report.

A. CHAPTER I INTRODUCTION

This chapter is a general, concise and concise explanation that describes exactly the content of the research. The contents of this chapter include: General Description of the Research Object, Research Background, Problem Formulation, Research Objectives, Research Benefits, and Final Project Writing Systematics.

B. CHAPTER II LITERATURE REVIEW

This chapter contains theories from general to specific, accompanied by previous research and continued with a research framework that ends with a hypothesis if needed.

C. CHAPTER III RESEARCH METHODS

This chapter confirms the approach, methods, and techniques used to collect and analyze findings that can answer the research problem. This chapter This chapter includes a description of: Type of Research, Variable Operationalization, Population and Sample (for quantitative) / Social Situation (for qualitative), Data Collection, Validity and Reliability Tests, and Data Analysis Techniques.

D. CHAPTER IV RESEARCH RESULTS AND DISCUSSION

The results of the research and discussion are described systematically in accordance with the problem formulation and research objectives and are presented in separate subtitles. This chapter contains two parts: the first part presents the research results and the second part presents the discussion or

analysis of the research results. Each aspect of the discussion should start from the results of data analysis, then be interpreted and then followed by drawing conclusions. The discussion should be compared with previous studies or relevant theoretical foundations.

E. CHAPTER V CONCLUSIONS AND SUGGESTIONS

The conclusion is the answer to the research question, then it becomes a suggestion related to the benefits of the research.