CHAPTER I

INTRODUCTION

1.1 Background

International Data Corporation (IDC) Asia Pacific (2023) projects that the transaction value of e-commerce in Indonesia will reach 118 billion USD by 2027. This growth in transaction value is dominated by digital payment methods, as presented in Figure 1.1. This growth reflects a transformation in consumer preferences in Indonesia, which are increasingly shifting towards the convenience offered by online transactions, as opposed to traditional purchasing approaches.
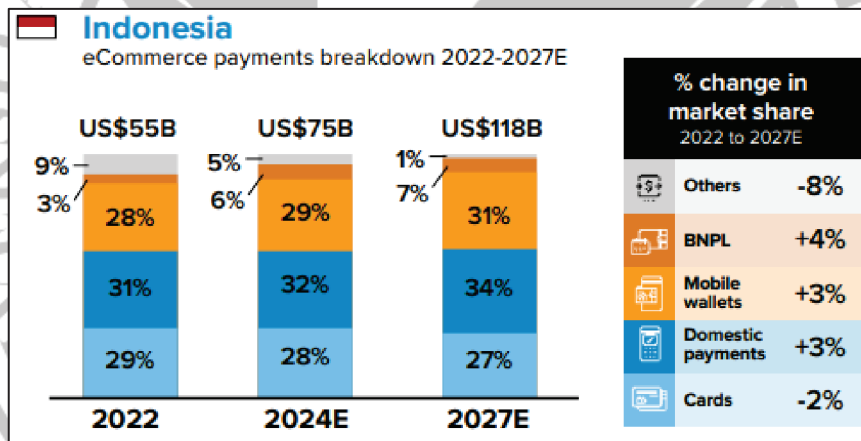

Figure 1.1 E-Commerce Breakdown 2022 - 2027E
Source: International Data Corporation (2023)

The projected growth in e-commerce transaction values aligns with Indonesia's increasing number of e-commerce users. Statista (2024) projections, the number of users in the e-commerce market in Indonesia is expected to continue rising between 2024 and 2029, reaching a total of 33.5 million users (+51.03 percent), as presented in Figure 1.2.
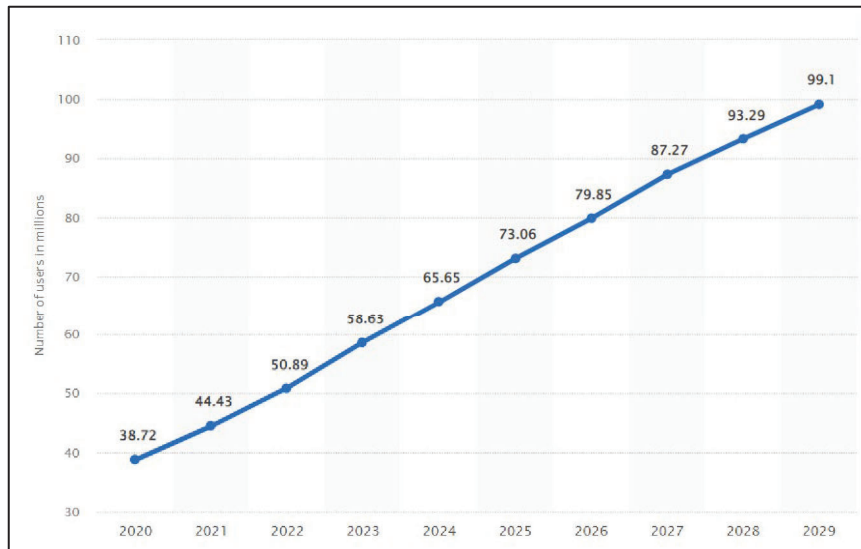
Figure 1.2 Projection of E-Commerce Users in Indonesia 2020 – 2029
Source: Statista (2024)

The increasing value of transactions and the use of e-commerce in Indonesia can lead to a significant rise in the number of e-commerce platforms. This aligns with the fact that e-commerce in Indonesia contributes 52% of the revenue (International Trade Administration, 2024). The increasing number of platforms has led to increasingly fierce competition among various e-commerce platforms to attract sellers, stores, and consumers.

Despite intense competition, Shopee remains the most visited e-commerce platform, according to data from Similar Web cited by Data Boks in 2023. Shopee has managed to maintain its position as the e-commerce platform with the highest number of visits in Indonesia, receiving approximately 2.3 billion visits, reflecting a significant increase of 41.39% since the beginning of the year, as presented in Table 1.3.

Tabel 1.1 E-Commerce Data with the Most Visitors in 2023

Num	Name Date	Value
1	Shopee	2.349.900.000
2	Tokopedia	1.254.700.000
3	Lazada	762.400.000
4	Blibli	337.400.000
5	Bukalapak	168.200.000

Source: Data Boks (2023)

The significant increase in the number of visits indicates high consumer demand, which can motivate more sellers or stores to join the Shopee e-commerce platform. The Muslim fashion industry is one industry that is interested in joining the Shopee platform. The interest in joining the Muslim fashion industry on the Shopee platform is also driven by the fact that Muslim fashion ranks third as the best-selling product, according to data from the Digital Marketing Association of Indonesia in 2020, with Muslim fashion products sold as many as 107 million units (pcs) as presented in Table 1.2.

Table 1.2 Best-Selling Products on the Shopee Platform in 2020

Num	Product	Value
1	Beauty	247.1 million
2	Home Appliances	133 million
3	Muslim Fashion	107 million
4	Women's Clothes	100 million
5	Handphone	78.2 million

Source: Asosiasi Digital Marketing Indonesia (2020)

Due to the increasing interest in purchasing Muslim fashion products, the number of sellers using the Shopee platform to market these products has significantly increased, creating more intense competition among sellers or brands

operating on the platform. The intense competition has led to a decline in sales for one of the Muslim fashion MSMEs, Dezhaf Hijab, based on data from 2022 and 2023, showing a decrease in sales of approximately 65.2%, as stated in Table 1.3.

Table 1.3 Dezhaf Hijab Sales 2022 and 2023

Month	2022	2023
Jan	120.384.318	226.542.703
Feb	164.527.985	240.574.101
Mar	304.292.304	343.900.030
Apr	874.257.742	639.196.598
May	254.941.150	113.121.414
Jun	751.073.142	302.827.612
Jul	799.438.189	224.248.129
Aug	520.215.589	225.162.577
Sep	575.327.972	167.266.761
Oct	929.611.181	154.970.760
Nov	1.059.212.244	141.634.643
Dec	1.397.895.251	142.188.618
Total	**9.555.087.127**	**3.321.734.926**

Source: Secondary Data (2024)

Therefore, Dezhaf Hijab strives to increase purchase interest to boost profits and revenue. In pursuit of this goal, the company emphasizes the importance of customer orientation by implementing strategies focused on communication, understanding, and recognition of customer needs. This ensures that customer needs are effectively met (Mintardjo, 2022). By adopting this customer-oriented approach, Dezhaf Hijab will leverage existing customers to make additional transactions or add more products to their shopping cart, ensuring customer needs are effectively fulfilled and fostering purchase intention.

To meet customer needs, Tsiptsis and Chorianopoulos (2010), emphasize in their book the importance of using data mining for business managers and retailers to extract insights from customers and enhance interactions with consumers in a more personalized manner. Data mining is the process of discovering knowledge in databases. It involves stages of data selection, data cleaning, data integration, data transformation, data mining, pattern evaluation, and visual knowledge presentation (Supoyo & Prasetyaningrum, 2022). Customer transaction data is one source that can be utilized in data mining to analyze online customers. This transaction data can then serve as the basis for evaluating and setting marketing strategies for the upcoming period (Brick, 2023).

One approach for understanding customer segmentation is RFM-D analysis. The RFM-D model, an extension of the RFM model, introduces an additional measure called Diversity; this model evaluates customer value and segmentation based on four main components: Recency (R), Frequency (F), Monetary (M), and Diversity (D) (Smaili & Hachimi, 2023). By integrating Diversity into the RFM model, marketers gain deeper insights into customer behavior, enabling more effective segmentation and marketing strategies. The researcher then conducted an analysis using the clustering method. Clustering is a technique used to group data based on the similarity of characteristics between one data point and another (Shaliha et al., 2021).

One of the clustering methods used by researcher is K-means clustering. The K-means algorithm is a clustering method that requires determining the number of clusters "k" and the objects "n" to be grouped into these clusters beforehand

(Mulyo & Heikal, 2022). Andre, Widya, and Juanda (2023) applying the K-means algorithm in data clustering is based on an iterative process to find the cluster centers, which form the fundamental basis of this algorithm. Therefore, applying the K-means algorithm in a clustering approach effectively analyzes structured data by reducing variance within clusters and enhancing differences between clusters, which is crucial for ensuring more accurate empirical conclusions.

In addition to applying the K-Means algorithm, this study utilizes the Market Basket Analysis (MBA) method for product bundling strategies. Market Basket Analysis, also known as Association Rule, is a method for discovering relationships between data by analyzing data patterns. This allows researcher to explore and identify existing patterns in the data more deeply (Amna et al., 2023). One algorithm that can be used in the Market Basket Analysis (MBA) method is the Apriori algorithm. The Apriori algorithm aims to find relationships between frequently purchased products in sales transactions; this helps companies discover related product combinations, thereby determining which products should have their stock increased or decreased to enhance sales revenue (Albab & Hidayatullah, 2022).

Furthermore, researcher using multiple linear regression (MLR) can be used as a supporting analysis in research involving RFMD analysis, k-means clustering, and market basket analysis. Multiple linear regression (MLR) is a simple and accurate machine learning model (Ehteram & Banadkooki, 2023). Multiple linear regression is a statistical modelling technique used to describe the relationship between a dependent variable and two or more independent variables (Jafar et al.,

2023). Therefore, multiple linear regression is a powerful tool for predicting and understanding relationships within data, with proper diagnostic and validation procedures crucial to ensure the reliability and stability of the results.

1.2 Research Question

Based on the background research, the research questions to be investigated in this study are as follows.

1. How can customer segmentation in online purchasing be mapped and understood, especially in the context of Dezhaf Hijab consumers on Shopee?

2. Based on the variability of purchasing patterns of Dezhaf Hijab consumers on Shopee, what types of online customer profiles can be identified?

3. How can applying market basket analysis (MBA) help optimize product bundling strategies?

4. How does price affect total (purchasing intention)?

5. How does stock code (product) affect total (purchasing intention)?

6. How does SKU ID (product variation) affect total (purchasing intention)?

7. How does quantity (demand) affect total (purchasing intention)?

8. How does province (location) affect total (purchasing intention)?

1.3 Research Objective

Based on the research problem and research question that have been outlined, the aim of this study is:

1. To understand how price, stock code (product), SKU ID (product variation), quantity (demand), and province (location) influence total (purchasing intention).

2. Identifying the online customer segmentation of Dezhaf Hijab and categorizing them based on their interactions on Shopee using the RFM method.

3. Identifying the number of customer clusters and their respective customer profiles remains integral to implementing K-Means clustering and RFM analysis.

4. Analyze product associations and perform market basket analysis on customer transaction data.

1.4   Research Scope

To reach the objectives research, the scope of this study is delineated as follows:

1. This study focuses on understanding customer segmentation and product bundling strategies.

2. The analysis is limited to e-commerce transaction data from the Shopee platform.

3. This study will focus on transaction data collected from December 9, 2023, to January 8, 2024, providing a snapshot of recent purchases.

4. This research will rely on the data mining techniques mentioned above, excluding other potentially relevant data analysis techniques.

1.5   Thesis Outline

This study consists of five chapters, with each chapter detailing specific content. The following are the five chapters:

1. Chapter I – Introduction

In this chapter, the author will delineate the research background about "Customer Profiling Using the K-Means Method and Product Recommendation Using the Market Basket Analysis for a Marketing Strategy to Increase Consumer Purchase Intention in MSMEs". Furthermore, this section will articulate the problem statement, the study's objectives, the research's significance, and the manuscript's systematic structure in relation to the conducted investigation.

2. Chapter 2 – Literature Review

In this chapter, the author will review the literature related to the basic rationale underlying the research, the hypotheses, and the theoretical framework of the study.

3. Chapter 3 – Research Methodology

In this chapter, the researcher will elucidate the conceptual framework and the meticulous steps undertaken in this investigation, encompassing the sample selection, data acquisition, and the methodology employed for data analysis, which will be expounded comprehensively.

4. Chapter 4 – Analysis and Discussion

In this chapter, the author will present the results and explain the analysis of the findings obtained through the methodology used.

5. Chapter 5 – Conclusion and Suggestion

In this chapter, the author will present the conclusions drawn from the entirety of the research, the limitations of the study, and recommendations for future research.