DIJEMSS:
Dinasti International Journal of Education Management and Social Science
E-ISSN: 2686-6331
P-ISSN: 2686-6358
https://dinastipub.org/DIJEMSS dinasti.info@gmail.com +62 811 7404 455
DINASTI PUBLISHER
DOI: https://doi.org/10.38035/dijemss.v6i5

Analysis of Recency Frequency Monetary (RFM) Using Python and Matplotlib in Optimizing Marketing Strategies at Dhiza Petshop Retail Store in Tanjung Selor

Eric Sapto Raharjo[1], Tinik Sugiati[2]
[1]Universitas Borneo Tarakan, Kalimantan Utara, Indonesia, ericasboel@gmail.com
[2]Universitas Borneo Tarakan, Kalimantan Utara, Indonesia, tiniksugiati@borneo.ac.id

Corresponding Author: tiniksugiati@borneo.ac.id[2]

Abstract: Petshops have been rapidly growing due to increased public awareness of pet care and food needs. The rising trend of pet adoption has fueled the expansion of petshops in various regions. The competition in the pet retail industry demands more targeted marketing strategies to enhance customer loyalty. One effective method is RFM (Recency, Frequency, Monetary) analysis, which allows for customer segmentation based on their shopping behavior. This study aims to analyze the customer behavior patterns of Dhiza Petshop Tanjung Selor using the RFM method with Python and Matplotlib, based on customer transaction data from December 21, 2024, to January 20, 2025, using 11 segments: Champions, Loyal Customers, Potential Loyalists, Recent Buyers, Promising, Need Attention, About to Sleep, Cannot Lose Them, At Risk, Hibernating, and Lost. The results show that the customer behavior patterns at Dhiza Petshop Tanjung Selor are dominated by customers with recent purchasing activity, while the number of loyal customers remains limited, indicating the need for retention and reactivation strategies to strengthen long-term relationships. The application of RFM analysis using Python and Matplotlib allows for efficient segmentation of customers, resulting in improved marketing strategies. The percentages for the 11 customer segments at Dhiza Petshop Tanjung Selor are: Champion (0.1%), Loyal (0%), Potential Loyalist (0.9%), Recent Buyers (72.6%), Promising (4.4%), Need Attention (0.1%), About to Sleep (0.2%), Cannot Lose Them (0%), At Risk (0%), Hibernating (12.1%), and Lost (9.6%).

Keyword: RFM, Customer Segmentation, Python, Matplotlib.

INTRODUCTION

In the current digital era, human life has undergone significant changes in almost all aspects, ranging from communication, work, to shopping. The rapid development of information technology has become a driving force influencing various facets of our lives (Lubis & Nasution, 2023). Technology, such as the internet, smart devices, and artificial intelligence, has created vast connectivity, making the world feel smaller and more accessible. All information can be accessed within seconds, allowing people to obtain solutions and services without time or location constraints. This transformation undoubtedly affects individuals' lives and drives major changes in global business and economic patterns. Today,

technology is involved in various fields, one of which is business (Lubis & Nasution, 2023). Consumers have become more critical and rely on technology in every step of their decision-making process, from searching for products to comparing prices and services (Ajzen, 2005). Companies must adapt to this new reality by leveraging technology to create innovation and efficiency. One of the most significant impacts of this digital era is its ability to help companies optimize marketing strategies through data-driven approaches and more personalized and effective technology (Astuti, 2018; Devenport, 2006).

Marketing strategies in the digital era have undergone significant changes compared to conventional approaches. Companies no longer rely solely on promotions through print media or television, but also utilize digital platforms such as social media, email marketing, websites, and mobile apps, given the increasingly competitive business environment (Agustini, 2003; Krisnawati, 2018). Entering the 4.0 industrial revolution era, it is crucial to understand and implement the right digital marketing strategies (Lubis & Nasution, 2023). This helps businesses reach customers more widely and directly. The digital era also offers companies the opportunity to understand customers' needs and preferences through real-time data analysis. With the ability to measure the effectiveness of campaigns directly, companies can quickly adjust their strategies to achieve more optimal results (Muliyah et al., 2020). Many businesses have lost market share due to their inability to create digital strategies for their business (Ananda et al., 2023). One increasingly relevant digital instrument in developing marketing strategies is Recency, Frequency, and Monetary (RFM) analysis.

RFM is a popular analytical method often applied in CRM (Customer Relationship Management)-based marketing strategies. This method assesses customer loyalty not only from the value of their transactions (monetary) but also considers the frequency (frequency) and the last time (recency) a customer interacted through transactions (J. Jamal & Yanto, 2019). RFM analysis has become one of the most effective methods for understanding customer behavior. As stated by Prasetyo et al. (2020) and Syah (2023), RFM performs effectively in customer segmentation using customer transaction data. This is because RFM helps companies identify and determine potential customers and those at risk of losing interest, allowing them to design more targeted marketing strategies (Smith, 1956; Syah, 2020). For example, customers with high Recency and Frequency values can be prioritized in loyalty programs, while customers with high Monetary values but low transaction frequency can be approached with special offers to increase their purchasing frequency (Putri Nugraha et al., 2021; Siahainenia, 2020). RFM analysis also provides valuable insights for companies to allocate marketing budgets more efficiently, as evidenced by Chaffey & Chadwick (2024) and Maulani & Nadhifatul Nur (2024), who argue that well-managed and analyzed customer data can maximize a company's potential and aid in making decisions to improve efficiency and purchasing decisions. In practice, RFM analysis can be done with the help of technologies like Python and Matplotlib as its visualization library. This simplifies data processing and presenting the analysis results in an informative way (Junaidi et al., 2023).

Python is a popular programming language due to its flexibility and efficiency in handling data. Python has become a primary tool in modern data analysis for many large companies (Romzi & Kurniawan, 2020). According to Kurniawan (2020), Python can handle large amounts of data, enabling more in-depth analysis. In RFM analysis, Python is used to read customer transaction data, group customers based on Recency, Frequency, and Monetary values, and create scores that describe purchasing behavior. Meanwhile, Matplotlib is used as a data visualization library to present the analysis results in the form of charts or graphs, such as histograms or heatmaps, which simplify interpretation and decision-making (Astandi, 2024; Gazali, 2017). The combination of Python and Matplotlib produces accurate reports that are easy to understand by the marketing team. In businesses like petshops, RFM analysis using Python and Matplotlib can be a strategic solution to identify potential customer segments and drive revenue growth.

Petshop businesses have become one of the rapidly growing sectors as public interest in pet ownership increases (Asgandi, 2020). Not only selling pet food and supplies, petshops now also offer additional services such as pet care, health consultations, and pet adoptions. Petshops are a potential business due to changes in people's lifestyles, particularly in urban areas, where pets are considered part of the family, leading to increased demand for pet products and services. Moreover, the support of technology in marketing processes has expanded the market reach of petshop businesses. The trend of pet ownership is rising, making this business sector highly promising (Elvida et al., 2021). One petshop that has been rapidly developing with various services and flagship products is Dhiza Petshop, located in Tanjung Selor.

Dhiza Petshop Tanjung Selor is a retail business that provides pet supplies such as food, vitamins, accessories, and pet care services. Located in the center of Tanjung Selor, Dhiza Petshop has successfully attracted local attention. The increasing number of pet owners has allowed Dhiza Petshop to see a huge potential to expand its reach through more targeted marketing strategies. In addition to focusing on product and service quality, Dhiza Petshop has also innovated by utilizing digital platforms to reach more customers for both promotion and sales. This success has made Dhiza Petshop one of the businesses that have managed to survive and grow amidst the increasingly competitive industry. However, Dhiza Petshop faces the challenge of how to increase company revenue, as the figures have been fluctuating and tend to stagnate. Sales data for Dhiza Petshop from January to November 2024 is shown in the following table:

Table 1. Sales Revenue of Dhiza Petshop from January to November 2024

2024	
Month	**Sales Revenue**
January	Rp. 553.000.000
February	Rp. 564.000.000
March	Rp. 600.000.000
April	Rp. 578.000.000
May	Rp. 583.000.000
June	Rp. 611.000.000
July	Rp. 665.000.000
August	Rp. 617.000.000
September	Rp. 591.000.000
October	Rp. 564.000.000
November	Rp. 600.000.000

Source: Dhiza Petshop Sales Data, 2024

Based on the sales revenue data from Dhiza Petshop Tanjung Selor from January to November, it appears that the sales figures are stagnant with insignificant fluctuations. Although there was an increase in July amounting to Rp. 665,000,000, the numbers dropped again in the following months and fluctuated between Rp. 564,000,000 and Rp. 617,000,000. This pattern indicates that revenue growth is not consistent and remains unstable. Therefore, it is necessary to implement an appropriate strategy to drive sustainable revenue growth.

Until now, Dhiza Petshop has not optimally utilized customer transaction data to understand their shopping behavior, such as visit frequency, spending amount, and the time of last purchase. As a result, Dhiza Petshop has struggled to identify the most potential customer segments for targeting promotions or loyalty programs. Yet, utilizing this data is crucial to enhancing customer experience and driving their loyalty to the brand. Without directed analysis, marketing strategies tend to be less precise and potentially result in wasted budgets (Zaki, 2018). This challenge is further exacerbated by the exponential growth of data. The

International Data Corporation (IDC) estimates that data storage grows by 60 percent annually, indicating that the volume of data stored will double every 20 months. This enormous amount of data burdens marketers, who struggle to measure the effectiveness of their actions with limited time and resources (Jeffery, 2010). One solution to address this problem is RFM analysis, which provides an in-depth view of customer purchasing patterns, enabling Dhiza Petshop to design more effective and efficient marketing strategies.

Previous research by Hermawan et al. (2024), which discusses marketing strategies through RFM analysis on retail transaction datasets, found that the majority of customers were in the "at risk" segment. Meanwhile, a study by Utama & Ardiansah (2023) on RFM's use in identifying customer behavior found that most customers were in the "about to sleep" segment. A research gap exists between the findings of these two studies, where the first study found that most customers were in the "at risk" segment, and the second study found that the majority were in the "about to sleep" segment. Thus, this research aims to fill the gaps in previous studies and generate better findings. As known, although RFM analysis has been widely applied in various sectors for customer segmentation, its specific application using Python and Matplotlib in the pet retail industry, particularly in Tanjung Selor, has not yet been found. The results suggest that each study yields different findings depending on the context, data, and methods used. Therefore, this research seeks to further explore the appropriate strategy based on customer segmentation to improve marketing effectiveness at Dhiza Petshop Tanjung Selor.

Based on the background, the researcher is interested in conducting a study titled "Recency Frequency Monetary (RFM) Analysis Using Python and Matplotlib in Optimizing Marketing Strategies at Dhiza Petshop Tanjung Selor."

METHOD

This research uses a quantitative design with a descriptive analysis approach to implement the RFM (Recency, Frequency, Monetary) method in customer segmentation using Python and Matplotlib. The aim of this study is to identify customer behavior patterns based on retail transaction data from Dhiza Petshop Tanjung Selor. RFM analysis is used to group customers into several segments to optimize marketing strategies.

The population in this study consists of all customer transaction data from Dhiza Petshop Tanjung Selor, which includes purchasing information such as transaction date (recency), purchase frequency (frequency), and total transaction value (monetary) made by customers within a certain period. The sample is obtained using purposive sampling technique, where the sample in this study is selected from customer transactions made between December 21, 2024, and January 20, 2025.

The instrument development process involves identifying research variables, including Recency, Frequency, and Monetary. Customer segmentation is carried out based on the results of RFM analysis, which is grouped into several categories as follows:

Table 2. Grouping Based on RFM Segments

Segment	Description
Champions	Best customers who have recently transacted, buy frequently, and generate high revenue.
Loyal Customer	Customers who frequently transact, not always recent, but contribute significantly.
Potential loyalist	New customers showing consistent purchasing patterns and potential to become loyal.
Recent Buyers	Customers who have recently bought but with relatively low frequency and monetary values.

Promising	Customers who frequently shop with large expenditures, but their last purchase was a while ago.
Need Attention	Customers with average purchase patterns; need more attention to remain engaged.
About to Sleep	Customers who are almost inactive and have relatively low purchase values.
Cannot Lose Them	Customers who have made high-value purchases but have not made a purchase for a long time.
At Risk	Customers who have high value but have been rarely transacting recently; need retention strategy to prevent churn.
Hibernating	Customers who have been inactive for a long time but still contribute slightly; need reactivation strategies.
Lost	Low RFM value.

Source: Maidla (2023)

The data used in this study includes customer transaction information, such as recency (last transaction time), frequency (purchase frequency), and monetary (amount spent) (Maskanah, 2020). Python is used to process raw data, conduct RFM analysis, and group customers based on their purchasing patterns. Python also facilitates the use of the Pandas library for data manipulation, while Matplotlib is used for data visualization, generating easily understandable graphs and plots, such as bar charts, pie charts, or histograms, to illustrate RFM analysis results. This visualization simplifies the evaluation of customer purchasing patterns and market segmentation (Bayhaqi & Aslami, 2022; Salmah, 2018). Documentation is used to obtain contextual information regarding the marketing strategies applied and to delve into how the petshop has structured its approach toward customers.

The data used in this study is primary data obtained directly from the original or primary source, collected by the researcher for the specific purpose of addressing the identified research questions or issues (Sugiyono, 2020). This data is directly collected from relevant parties or sources, namely from Dhiza Petshop Tanjung Selor's sales system, which includes transaction dates, purchase amounts, purchased products, and relevant customer data. This transaction data covers the period from December 21, 2024, to January 20, 2025. After the data is collected, the data filtering process is carried out to ensure that only valid data that meets the criteria is used in the analysis. The filtered data is then processed using Python for RFM analysis, where the Pandas library is used for data manipulation and Matplotlib is used to generate understandable data visualizations in the form of bar charts or pie charts.

This study uses analysis with 3 indicators explained in the operational definitions as follows: Recency (R) refers to the time elapsed since the customer last made a purchase at Dhiza Petshop Tanjung Selor, measured by calculating the time difference between the last transaction date made by the customer and the data collection date in days. The more recent the transaction made by the customer, the higher the recency score. Frequency (F) refers to how often a customer makes a purchase at Dhiza Petshop Tanjung Selor from December 21, 2024, to January 20, 2025. Monetary (M) is measured by calculating the amount of money spent by a customer at Dhiza Petshop Tanjung Selor during this period, which is done by summing up the total spending of the customer during the specified period.

The operational definitions can be seen in the following table:

Table 3. Operational Definitions

Variable	Indicator	Reference Source
RFM	• *Recency* • *Frequency* • *Monetary*	(Jamal Jamal & Yanto, 2019)

Python is used to process data and implement the calculations and RFM analysis in this study. Python is used to process customer data, calculate Recency, Frequency, and Monetary values, and create customer segmentation models or analyses for Dhiza Petshop Tanjung Selor. The Python libraries used include Pandas for data manipulation. Matplotlib is a data visualization library in Python used to create graphs and diagrams that help in illustrating the results of the analysis used to create visualizations of RFM analysis results.

The data preprocessing process is carried out by obtaining data from the company's database from December 21, 2024, to January 20, 2025. Data preprocessing is a process that must be carried out before analyzing the data, involving steps such as cleaning, removing, and formatting the data so that it meets the analysis requirements. Data categorization is performed to select the necessary columns for the analysis process. Four columns have string or categorical data types: customer id, invoice id, amount$, and invoice date. In the data categorization process, these columns are used as parameters to measure transaction amounts, the time interval of the first and last transactions, and the frequency of transactions made by each customer.

The calculation of R, F, and M values is performed using Python with quintiles, where the RFM score boundaries are obtained by calculating quintiles, which divides the values into 5 equal parts (20%) based on the minimum and maximum values for each attribute according to the following formula and conditions (Sutresno et al., 2018):

$$Q_i = \frac{i(n+1)}{5}$$

Explanation:
Q_i = The i-th quintile
i = The quintile to be calculated
n = Number of data

The calculation of Recency is divided into 5 parts based on the number of days since the last transaction up to the data analysis date. The 5 parts are as follows:

1. If a customer's Recency data falls into quintile 1, it means the customer has not transacted for a long time.
2. If a customer's Recency data falls into quintile 2, the customer has not transacted recently, but not too long ago.
3. If a customer's Recency data falls into quintile 3, it means the customer has a medium transaction time.
4. If a customer's Recency data falls into quintile 4, it means the customer has transacted fairly recently.
5. If a customer's Recency data falls into quintile 5, it means the customer has transacted very recently.

The calculation of Frequency is divided into 5 parts based on the intensity of the customer's transactions. The 5 parts are as follows:

1. If a customer's Frequency data falls into quintile 1, it means the customer rarely transacts.

2. If a customer's Frequency data falls into quintile 2, it means the customer still transacts infrequently.
3. If a customer's Frequency data falls into quintile 3, it means the customer has an average frequency of transactions.
4. If a customer's Frequency data falls into quintile 4, it means the customer often transacts.
5. If a customer's Frequency data falls into quintile 5, it means the customer transacts very frequently.

The calculation of Monetary is divided into 5 parts based on the total GMV spent from each transaction. The 5 parts are as follows:

1. If the total GMV of a customer falls into quintile 1, it means the customer has very low purchasing value.
2. If the total GMV of a customer falls into quintile 2, it means the customer's purchase value is relatively low.
3. If the total GMV of a customer falls into quintile 3, it means the customer has an average monetary value.
4. If the total GMV of a customer falls into quintile 4, it means the customer's purchase value is quite high.
5. If the total GMV of a customer falls into quintile 5, it means the customer has a very high purchasing value.

The customer segmentation clustering is carried out based on the R, F, and M values obtained earlier. The categorization results are grouped into 10 categories with each R, F, and M value as shown in the table below:

Table 4. RFM Segmentation Clustering

Segment	RFM Score	Description
Champions	555, 554, 544, 545, 454, 455, 445	Best customers who have recently transacted, buy frequently, and generate high revenue.
Loyal Customer	543, 444, 435, 355, 354, 345, 344, 335	Customers who haven't transacted recently but contribute significantly.
Potential loyalist	553, 551, 552, 541, 542, 533, 532, 531, 452, 451, 442, 441, 431, 453, 433, 432, 423, 353, 352, 351, 342, 341, 333, 323	Customers who frequently transact but not always recently and contribute significantly.
Recent Buyers	512, 511, 422, 421 412, 411, 311	Customers who have high values but haven't transacted in a while; require retention strategies to prevent churn.
Promising	525, 524, 523, 522, 521, 515, 514, 513, 425,424, 413,414,415, 315, 314, 313	Customers with average purchase patterns; need more attention to stay engaged.
Need Attention	535, 534, 443, 434, 343, 334, 325, 324	New customers who have potential to become loyal if facilitated well.
About to Sleep	331, 321, 312, 221, 213, 231, 241, 251	Customers who are almost inactive and have relatively low purchase values.

Cannot Lose Them	155, 154, 144, 214,215,115, 114, 113	Customers who haven't transacted in a while, are very rare buyers, and have low value.
At Risk	255, 254, 245, 244, 253, 252, 243, 242, 235, 234, 225, 224, 153, 152, 145, 143, 142, 135, 134, 133, 125, 124	Customers showing consistent purchase patterns with potential to become loyal.
Hibernating	332, 322, 233, 232, 223, 222, 132, 123, 122, 212, 211	Customers who have been inactive for a long time but still contribute slightly; need reactivation strategies (Sitorus & Nugraha, 2025).
Lost	111, 112, 121, 131,141,151	Last purchase was made a long time ago.

Source: Maidla (2023)

Customer segmentation through clustering techniques can reflect their transaction patterns (Kotler & Keller, 2016). With this method, companies can more easily predict customer behavior because each segment represents a group with similar characteristics. This also supports companies in formulating more targeted business strategies for the future (Utama & Ardiansah, 2023).

RFM analysis is performed using Python to cluster each customer based on RFM values and Matplotlib to present research data visualization. Any customers whose RFM values do not fall within the Table 3.3 categories will be grouped into the "Others" segmentation.

RESULTS AND DISCUSSION

The data used in this study was obtained from Dhiza Petshop Tanjung Selor, with a total of 5,729 transactions during the shopping period from December 21, 2024, to January 20, 2025. The data includes customer information such as Customer ID, Invoice ID, Amount (transaction value), and Invoice Date (transaction date). Based on this data, recency is calculated based on the number of days since the last transaction, frequency is calculated by the number of transactions per customer, and monetary is determined by the total amount spent by each customer. The results are displayed in a table with columns for R, F, and M.

The characteristics of the transaction data are analyzed to identify customer segments based on the purchasing behavior of Dhiza Petshop Tanjung Selor's customers. The characteristics of the customer transaction data during the study period can be seen in the following table:

Table 5. Characteristics of Transaction Data

Total Transactions	Total Transaction Value	Average Transaction Value	Min	Max
5.729	Rp 467.143.000	Rp 81,554.29	Rp. 6.000	Rp. 958.000

Source: Research Data, 2025

Based on the transaction data characteristics provided, it is evident that the total number of recorded transactions is 5,729, with a total transaction value of Rp 467,143,000. The average transaction value per customer is Rp 81,554.29. The highest transaction value is Rp 958,000, while the lowest transaction value is Rp 6,000. This indicates a high volume of transactions with relatively significant average transaction values per customer, suggesting that customers make frequent transactions with notable monetary values.

The characteristics of the products sold are analyzed to understand sales trends and identify the most popular products among customers at Dhiza Petshop Tanjung Selor. This analysis aims to provide an overview of product demand patterns, which can be used as a basis for stock planning and marketing strategy. The characteristics of the products sold during the study period can be seen in the following table:

Table 6. Characteristics Based on Products Sold

No.	Product Name	Qty Sold	Nominal (Rp)
1	Bolt Tuna 800gr	3272	71,984,000.00
2	Cat Choize Adult Tuna	1154	25,388,000.00
3	Cat Choize Kitten Tuna	667	19,710,000.00
4	Whiskas Pouch J Tuna	553	3,871,000.00
5	Whiskas Pouch J Mackrl	471	3,297,000.00
6	Whiskas Pouch 1+ Tuna	449	3,143,000.00
7	Whiskas Pouch Grilleds	369	2,583,000.00
8	Whiskas Pouch Mac Salm	341	2,387,000.00
9	Cat Choize Adult Salmo	272	5,984,000.00
10	Whiskas Pouch Macar1+	262	1,834,000.00
…	…	…	…
…	…	…	…
412	Rc Digestive Care 400	1	87,000.00
413	Sankoi Spirulina 1kg	1	82,000.00
414	Under 60x60(10)	1	60,000.00
415	Imo Kitten, Baby & Mom	1	181,000.00
416	Groming J/K 1KG	1	60,000.00

Source: Processed Data, 2025

Based on the sales data at Dhiza Petshop Tanjung Selor during one month, the majority of the most sold products are cat food. The best-selling product is Bolt Tuna 800gr, with 3,272 units sold, generating Rp 71,984,000 in revenue. Other best-selling products include Cat Choize brands, such as Cat Choize Adult Tuna (1,154 units, Rp 25,388,000) and Cat Choize Kitten Tuna (667 units, Rp 19,710,000). The wet food products in pouch packaging from the Whiskas brand are also quite popular, though their revenue is smaller due to the smaller packaging sizes.

Conversely, some products were less popular, with only one unit sold during the research period, such as Rc Digestive Care 400 (Rp 87,000), Sankoi Spirulina 1kg (Rp 82,000), Under 60x60(10) (Rp 60,000), Imo Kitten, Baby & Mom (Rp 181,000), and Groming J/K 1KG (Rp 60,000). The low sales of these products are attributed to their relatively high prices and varying customer preferences, leading to a lack of interest.

From this analysis, it is evident that cat food products dominate the sales, indicating that the majority of Dhiza Petshop's customers are cat owners. Therefore, marketing strategies can focus on optimizing stock for the best-selling cat food, promoting high-demand products, and evaluating less popular products to increase sales in the upcoming months (Hamdat, 2020).

Further data analysis in this research underwent several processes, including library installation and import, dataset import, reference date determination, RFM value calculation, customer segmentation, and data visualization using Matplotlib.

1. *Instal* and *Import Library*

Library installation and import aim to prepare the tools required for data analysis. Using the !pip install command, we ensure that the pandas library (for data manipulation) and matplotlib (for visualization) are integrated into the Colab environment. The installation and import of libraries were carried out with the dataset seen in Appendix 1. Once the dataset and libraries are successfully imported, the code can be used for analysis, calculations, and graph creation to support the research.

2. *Import Dataset*

The process of reading the dataset is the first step in preparing the customer data for analysis. This process is carried out using the pd.read_csv() command to load the data stored in CSV format into the Python environment. The dataset contains information related to customer behavior, such as the date of the last purchase (Recency), purchase frequency (Frequency), and total customer spending (Monetary). The purpose of reading the dataset is to obtain the necessary data for calculating the RFM scores. Additionally, this step ensures that the available data is in the correct format before further analysis. In this process, the pd.read_csv() command is used to load the data from the Research Data file, which is then stored in the variable df. After loading the data, the df.head() command is used to display the first five rows of the dataset, allowing for an initial glance at the structure of the data and ensuring that the data has been correctly loaded into the Python environment. This step is crucial for performing an initial check before proceeding to the next stage of analysis. Once the dataset import process is successfully executed, the system will display the data from the uploaded summary file. The result of reading the data in Google Colab is shown below.

	Customer ID	Invoice ID	Amount	Invoice Date
0	D0000406	FJ001-211224-0001	127000	21/12/2024
1	PLG0870	FJ001-211224-0002	17000	21/12/2024
2	PLG0871	FJ001-211224-0003	44000	21/12/2024
3	PLG0872	FJ001-211224-0004	44000	21/12/2024
4	PLG0873	FJ001-211224-0005	40000	21/12/2024

Figure 1. Data Read Result in Import Dataset

In the figure, it shows that the data has been successfully loaded into the coding system and is now available for further customer segmentation analysis.

3. *Determining the Reference Date*

Determining the reference date in RFM analysis is an important step in calculating the Recency (R) score, which measures how recent a customer's purchase was. The reference date is selected as the most recent date in the dataset, which is January 20, 2025. This date represents the latest recorded transaction in the data. By using this reference date, the time difference between each customer's transaction and the reference date can be calculated, resulting in the Recency score. This score indicates how active the customer has been in recent purchases, which is a key indicator in customer segmentation for marketing strategies. The reference date is determined with the following code, which can be found in Appendix 1:

The command df['Invoice Date'].max() is used to retrieve the latest transaction date (maximum) recorded in the 'Invoice Date' column of the dataset, which is then considered as the reference date. This date becomes the basis for measuring the time difference between each customer's transaction and the last recorded transaction date.

Once the code to determine the reference date has been successfully executed, the system will display the reference date in the research data. The result of reading the reference date in Google Colab is shown below.

```
Tanggal referensi: 2025-01-20
```

Figure 2. Reference Date

The appearance of the reference date in the figure indicates that the data has been successfully loaded into the coding system and is now available for further customer segmentation analysis. The reference date is set as January 20, 2025.

4. *Calculating RFM Values*

The calculation of RFM (Recency, Frequency, and Monetary) values is at the core of RFM analysis for customer segmentation. Recency is measured to determine how recent a customer's purchase was, calculated by the time difference between the reference date and the last purchase date. Frequency indicates how often a customer makes a purchase over the research period, based on the number of transactions performed by the customer. Monetary is used to measure the total expenditure of a customer, calculated by the total transaction value during the analysis period.

a. RFM Data Calculation for Customers

Each RFM variable is calculated for each customer, and the results are used to assign scores. The RFM scores obtained will serve as the basis for segmenting customers into specific groups. RFM value calculation in Python is done with the code provided in Appendix 1. The result is a DataFrame containing RFM values for each customer, which generates the output of the calculations, as shown in the following image:

	Customer ID	Frequency	Monetary	Recency
0	D0000002	6	300000	1
1	D0000003	10	987000	3
2	D0000004	29	2141000	1
3	D0000005	19	857000	3
4	D0000006	10	875000	3
...				
	Customer ID	Frequency	Monetary	Recency
2008	PLG2198	1	720000	1
2009	PLG2199	1	28000	1
2010	PLG2200	1	15000	1
2011	PLG2201	1	27000	1
2012	PLG2202	1	66000	0

Figure 3. RFM Calculation Results for Customers

The RFM analysis results provide information on customer behavior based on three main variables: Recency, Frequency, and Monetary. For example, customer with Customer ID D0000002 has a Frequency of 6 transactions, a Monetary value of 300,000, and a Recency of 1 day. This data indicates that the customer recently made a purchase with relatively low transaction frequency. Meanwhile, Customer ID D0000003 has a higher purchase frequency of 10 transactions, a significantly higher Monetary value of 987,000, and a Recency of 3 days.

This shows that this customer has the highest monetary value, even though there are other customers with a higher shopping frequency, indicating that this customer tends to spend more per purchase. This interpretation is performed for each customer to analyze their purchasing behavior at Dhiza Petshop Tanjung Selor.

b. RFM Scoring for Customers

RFM scoring is performed to categorize customers based on their purchasing behavior. The scoring process begins by assigning values to each Recency (R), Frequency (F), and Monetary (M) dimension. Recency values are assigned based on how recent a customer made a purchase, where the smaller the Recency value, the higher the score given. Frequency indicates how often a customer transacts, so the more transactions performed, the higher the score. Meanwhile, Monetary reflects how much a customer spends, so the larger the amount spent, the higher the score. Each dimension is then scored on a scale of 1 to 5. The final customer score is the combination of scores across the three dimensions. RFM scoring results are shown in the image provided in Appendix 1.

The purpose of this code is to assign scores for each RFM dimension to categorize customers based on their transaction behavior. The code uses the pd.cut function to divide the values of each dimension into 5 categories, each labeled from 1 to 5, where 5 indicates the best performance and 1 indicates the worst performance.

Using this code, Python will display the top five records as a sample to show that the scoring has been successfully applied. The result can be seen in the following image:

	Customer ID	Frequency	Monetary	Recency	R_Score	F_Score	M_Score	RFM
0	D0000002	6	300000	1	5	1	1	
1	D0000003	10	987000	3	5	2	2	
2	D0000004	29	2141000	1	5	4	3	
3	D0000005	19	857000	3	5	3	2	
4	D0000006	10	875000	3	5	2	2	
...								
	Customer ID	Frequency	Monetary	Recency	R_Score	F_Score	M_Score	
2008	PLG2198	1	720000	1	5	1	1	
2009	PLG2199	1	28000	1	5	1	1	
2010	PLG2200	1	15000	1	5	1	1	
2011	PLG2201	1	27000	1	5	1	1	
2012	PLG2202	1	66000	0	5	1	1	

Figure 4. RFM Scoring Output

The RFM analysis results show scoring values that represent customer behavior based on the three dimensions: Recency, Frequency, and Monetary. For example, Customer ID D0000002, with a Recency score of 5 (indicating a recent purchase), Frequency of 6 (purchasing a few times), and Monetary of 300,000, indicates that this customer recently made a purchase with relatively low transaction value and purchase frequency. This scoring result helps to identify the segments for each customer based on their different shopping behaviors.

5. *Segmenting Customers Based on Segments*

In this section, customers are grouped based on the segments obtained from the RFM analysis. Each customer is classified into a segment that reflects their purchasing behavior and loyalty potential, with the goal of making it easier for the company to design more targeted and effective marketing strategies. The segments used include:

a. Recent Buyers
b. Hibernating
c. Lost
d. Promising
e. Potential Loyalists
f. About to Sleep
g. Champions

h. Need Attention
i. At Risk
j. Loyal Customers
k. Cannot Lost Them

The Python code used for the customer segmentation process at Dhiza Petshop Tanjung Selor can be found in Appendix 1. The goal of this code is to classify each customer into 11 segments based on the combination of previously calculated RFM scores. The assign_segment function is used to map each customer's RFM score combination to the relevant segments. Each combination of RFM scores is matched with the corresponding segment criteria to gain a deeper understanding of the customer's status and potential.

After running this function, a new column containing the customer segments is automatically added to the dataset, and the segmentation results and segment distribution can be seen in the following image:

```
[2012 rows x 9 columns]
Segment
Recent Buyers          1460
Hibernating             244
Lost                    194
Promising                88
Potential Loyalists      17
About To Sleep            5
Champions                 2
Need Attention            2
Name: count, dtype: int64
```

Figure 5. Customer Segmentation Output

The results of the RFM analysis show that the majority of customers, specifically 1,460 individuals, fall into the "Recent Buyers" category, indicating that they have made recent purchases but with lower transaction frequency and value. The "Hibernating" and "Lost" segments, consisting of 244 and 194 customers respectively, indicate customers who were once active but have not made any recent purchases. Meanwhile, the "Promising" and "Potential Loyalists" segments, with 88 and 17 individuals respectively, represent customers who have the potential to become loyal customers in the future, though their numbers are relatively small.

There are also other segments with fewer customers, including "About to Sleep," "Champions," and "Need Attention," each consisting of customers with different levels of loyalty and transaction frequency. However, no customers were found in the "At Risk," "Loyal Customers," or "Cannot Lose Them" segments, as no shopping behavior at Dhiza Petshop Tanjung Selor matched the criteria for these segments. This finding reveals the existence of customer groups with varying behaviors.

6. *Data Visualization with Matplotlib*

After obtaining the results from customer segmentation, the next step is to visualize the data using Matplotlib to support the RFM analysis by displaying the distribution of customers based on the defined segments. Visualizations in the form of bar charts and pie charts are used to show the percentage contribution of each segment to the total customer population.

Both types of visualizations help to identify customer distribution patterns, making it easier to analyze them intuitively and providing a clearer picture of how customers are spread across each segment, supporting more effective strategic decision-making in managing customer behavior at Dhiza Petshop Tanjung Selor. The data visualization for customer segmentation is shown in the following images:

a. Data Visualization with Bar Chart

A bar chart is used to depict the number of customers in each segment, making it easier to identify which segments have the largest or smallest customer base. The results can be seen in the following image:

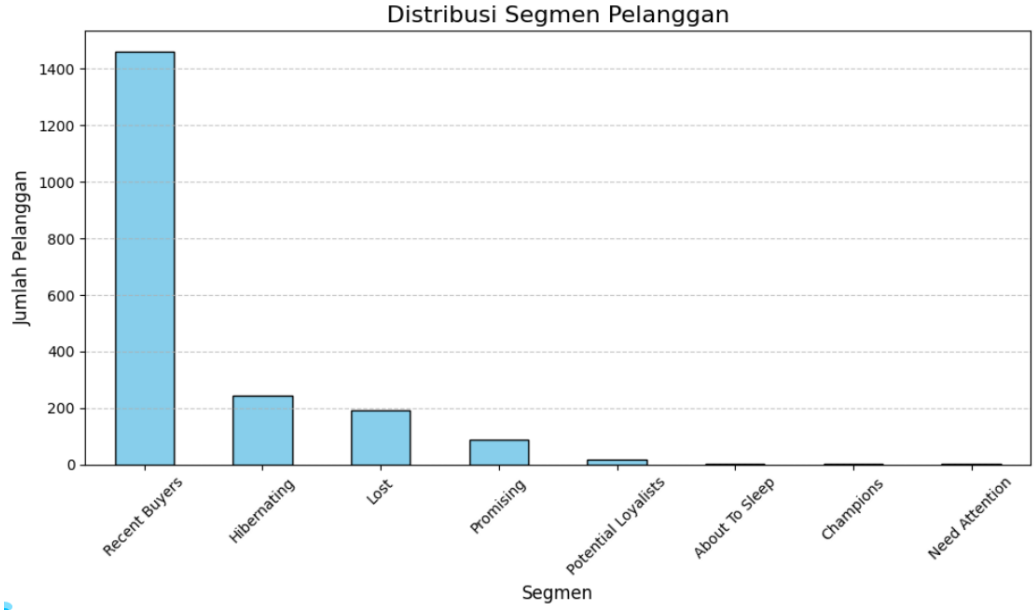


Figure 6. Data Visualization with Bar Chart

In this chart, it can be observed that the majority of customers fall under the "Recent Buyers" segment, while the "Loyal Customers" segment has the fewest customers.

b. Data Visualization with Pie Chart.

A pie chart is used to show the percentage contribution of each segment to the total customer population. The results are displayed in the following image:

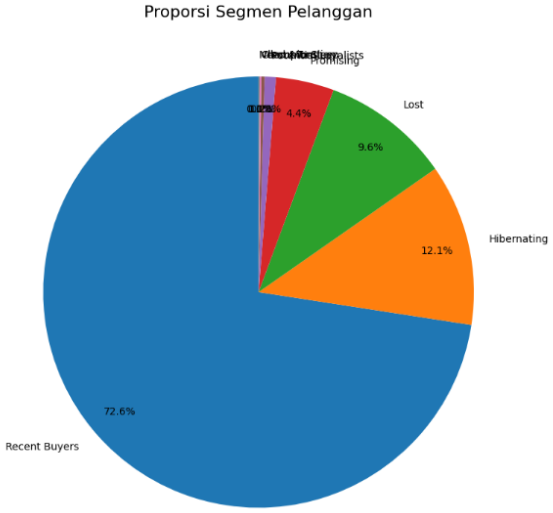


Figure 7. Data Visualization with Pie Chart

From the bar chart and pie chart values above, the number of customers in each segment and their respective percentages are calculated. The results are presented in the following table:

Table 7. Customer Segmentation

No.	Segment	Number of Customers	Percentage (%)	Strategy
1	*Champions*	2	0,1	1. Exclusive program for VIP customers, higher reward points on premium food, cat litter, or pet care products. 2. Loyalty rewards with tiered discounts. For example, every purchase of 5 kg of premium cat food, get 1 kg free. 3. Special invitations to brand events, such as Cat Lovers Day, for loyal customers who frequently purchase cat products. Events can include discussions, dog fun days, or pet care training.
2	*Loyal*	0	0	1. Long-term loyalty program. For example, each purchase of a certain amount of cat or dog food earns points that can be redeemed for discounts, free products, or services such as free pet health consultations with a veterinarian. 2. Special discounts for repeat purchases. For instance, a 10% discount on every purchase of a certain quantity of cat food within a month.
3	*Potential Loyalist*	17	0,9	1. Offer attractive promotions such as buy 1 get 1 free for cat food or discounts on bundled cat litter and food in a single transaction. 2. Discount for next purchase. For example, after purchasing large quantities of cat food or cat litter, customers receive a 10% discount coupon for their next purchase within a certain period. 3. Membership programs to enhance their loyalty. For instance, customers who frequently purchase premium cat food can collect points that can be redeemed for free products, while those who buy pet care supplies frequently can receive special discounts on grooming services.

4	*Recent Buyers*	1460	72,6	1. Warm welcome: When a customer makes their first purchase, offer friendly and informative service. 2. Discount offers for the next purchase. 3. Referral programs to encourage them to return. For example, if one customer invites a friend to shop at Dhiza Petshop, both will receive a discount or free products such as food samples or cat and dog snacks.
5	*Promising*	88	4,4	1. Offer more personalized promotions. For example, if customers frequently buy cat food, send them special offers like discounts for their favorite cat food brands. 2. Special offers such as a lottery coupon. For example, for every purchase of cat or dog food worth at least IDR 100,000, customers receive a lottery ticket with a prize of a big discount, free food packages, or pet accessories. 3. Encourage repeat purchases by reminding customers to maximize the benefits of the products. If they've just bought cat litter, send a reminder that their stock may be running low, along with a discount offer.
6	*Need Attention*	2	0,1	1. Send a reminder via WhatsApp. For example: "We noticed that you haven't bought your pet's favorite food for a while. Is your stock running low? We have new product recommendations that you can try." 2. Offer special discounts, such as a "Buy 2, Get 1 Free" or a 10% discount on food purchases. 3. Time-limited promotions: Use urgency strategies by offering limited-time promotions, for example: "15% off for the next 3 days! Don't miss the opportunity to buy your favorite cat/dog food at a discounted price."

7	*About to Sleep*	5	0,2	1. Create a retention campaign with a “last chance” offer. This should be clear and appealing, encouraging customers to feel they would regret not taking advantage of it. 2. Aggressive promotions before they become completely inactive. For customers who usually buy dog products such as shampoo or treats, offer additional discounts on bundled food and vitamins.
8	*Cannot Lose Them*	0	0	1. Special offer: “We Miss You.” Send a personal message, such as: "We miss seeing you shop. As a token of appreciation, we’re offering you an exclusive 25% discount on your favorite premium food or cat litter. Only for special customers like you!” 2. Direct consultation or personal approach: If customers frequently purchase special food for cats with specific dietary needs, invite them to discuss the best options available. 3. New and exciting loyalty program: Points that can be exchanged for free products or special services like free grooming after several purchases.
9	*At Risk*	0	0	1. Re-engage with appealing offers like big discounts. 2. Special loyalty programs for returning customers. For example, for every purchase of a certain amount of cat food, customers receive free cat litter as a bonus.
10	*Hibernating*	244	12,1	1. Create a special campaign such as “We Miss You.” Send a personal message to customers with an emotional touch, for example: "We miss you and your beloved cat or dog! It’s been a while since you shopped with us. We’re giving you a special discount. Come back and pamper your pets!" 2. Offer discounted product packages to entice them back. For

				example, a "Comeback Special" package for cat owners containing premium food, cat litter, and snacks at a 20% discount.
11	*Lost*	194	9,6	1. Special program for customers returning after a long period of inactivity. If customers haven't made a purchase in the last 6 months, they can receive a 25% discount voucher for their first purchase after returning to Dhiza Petshop. 2. Run a reactivation campaign through WhatsApp or social media with appealing offers and urgency to encourage customers to return to shopping at Dhiza Petshop.

Source: Hermawan et al. (2024)

Based on the data visualization in the form of a pie chart, it can be seen that the "Recent Buyers" segment dominates with 72.6% of the total customers. This indicates that the majority of customers have recently made a purchase, providing a significant opportunity to retain them through targeted marketing strategies. One shopping activity from a customer in the "Recent Buyers" segment can be seen in Customer ID D000025, who made a purchase on December 25, 2024. This transaction shows that the customer purchased Lifecat cat food and Chiro cat litter, with a total transaction of IDR 84,000.

The "Hibernating" segment ranks second with 12.1%, indicating a group of customers who have not been active in making transactions for a long time (Sitorus & Nugraha, 2025). One shopping activity from a customer in the "Hibernating" segment can be seen in Customer ID D000304, who made a purchase on December 31, 2024. This transaction shows that the customer purchased Felibite and Lezato cat food, with a total transaction of IDR 90,000.

The "Lost" segment ranks third with 9.6%, indicating customers who have stopped transacting and require a special approach to be re-engaged. One shopping activity from a customer in the "Lost" segment can be seen in Customer ID PLG0916, who made a purchase on December 21, 2024. This transaction shows that the customer bought MEO Kitten cat food for IDR 70,000.

Meanwhile, other segments such as "Promising" (4.4%), "Potential Loyalist" (0.9%), and "About to Sleep" (0.2%) have a relatively small number of customers but show potential for increased loyalty. One shopping activity from a customer in the "Promising" segment can be seen in Customer ID PLG1336, who made a purchase on January 13, 2025. This transaction shows that the customer purchased Maxi, KF, and Taro cat food for a total transaction of IDR 208,000.

The segments with very small percentages, namely "Champions" and "Need Attention", each contributing 0.1% or less, indicate the need for extra attention to maintain or enhance performance in these segments. One shopping activity from a customer in the "Need Attention" segment can be seen in Customer ID D0000179, who made a purchase on January 7, 2025. This transaction shows that the customer bought Kitekat, Beauty Gold Premium, and Cat Choize cat

food for a total of IDR 132,000. Meanwhile, one shopping activity from a customer in the "Champions" segment can be seen in Customer ID D0000014, who made a purchase on January 18, 2025. This transaction shows that the customer bought RC skin and fur care vitamins and Lifecat cat food for a total of IDR 105,000. This customer tends to purchase premium products when shopping at Dhiza Petshop Tanjung Selor.

The analysis results show that the majority of customers belong to a category that reflects recent purchasing activity, with only a few customers categorized as loyal or highly engaged. This suggests that Dhiza Petshop has significant potential to retain a substantial customer base, but also faces challenges in establishing strong long-term relationships. This situation reflects Dhiza Petshop's success in attracting many new customers, but the behavior pattern dominated by customers who have only recently made purchases indicates that customer retention efforts require more attention. In marketing terms, this highlights both significant opportunities and critical challenges in maintaining long-term customer relationships, particularly with the increasing competition in the market.

From a customer management perspective, these results emphasize the importance of strategies focused on increasing customer loyalty and building emotional connections between customers and Dhiza Petshop. Marketing strategies such as loyalty programs, personalized services, and consistent communication are key steps in ensuring that customers continue choosing Dhiza Petshop for their needs. Additionally, the analysis also indicates the presence of customer segments that are starting to become inactive or have stopped transacting altogether. This phenomenon signals that customer reactivation programs should be a priority, particularly for customers who have the potential to become active again if provided with the right promotions. Research by Blattberg et al. (2009) and Christian et al. (2022) emphasizes that data-driven approaches, such as RFM analysis, can enhance marketing strategy efficiency and have a significant impact on overall customer value.

The segmentation results show that several customer categories exhibit different behavioral patterns. For example, customers in the "Champion" category, such as Customer ID D0000493, with a frequency of 29, monetary value of IDR 4,028,000, and recency of 0, often purchase premium products like Whiskas Pouch Junior. On the other hand, customers in the "Potential Loyalist" category, such as Customer ID D0000526, with a frequency of 21, monetary value of IDR 674,000, and recency of 0, frequently purchase Bolt Tuna 800g products. The "Need Attention" segment includes customers like Customer ID D0000179, who shows a frequency of 19 and a monetary value of IDR 3,353,000 but with a recency of 1, frequently buying Kitchen Flavour (KF). Meanwhile, the "Lost" segment includes customers like Customer ID D0000377, who has a frequency of 2, monetary value of IDR 149,000, and recency of 25, purchasing products such as Lifecat Tuna 400g. This segmentation provides clear guidance for strengthening marketing strategies and customer relationship management at Dhiza Petshop Tanjung Selor, indicating that management must be more adaptable to customer needs and create relevant value aligned with their expectations.

The application of RFM analysis using Python and Matplotlib at Dhiza Petshop Tanjung Selor has been effective in segmenting customers into 11 segments, including champions, loyal customers, potential loyalists, recent buyers, promising, need attention, about to sleep, cannot lose them, at risk, hibernating, and lost. Python is used for its flexibility in data processing and efficiency in computation. Customer data, including customer ID, invoice ID, purchase amount, and transaction date, is processed using the pandas library to calculate scores for each RFM dimension. Scores are assigned based on quantiles, where higher scores indicate better customer performance. After analyzing the data, visualization using Matplotlib helps map the distribution of customers based on their Recency, Frequency, and Monetary scores into the 11 existing segments, facilitating the understanding of customer behavior patterns.

The research shows that Python-based RFM analysis significantly improves the efficiency and accuracy of customer segmentation compared to manual methods. A study by

Febriana & Rosadi (2024) found that using RFM analysis provides deeper insights into the characteristics and value of each customer, allowing companies to design more effective marketing strategies and increase customer retention. Research by Kohavi & Parekh (2004) also highlights the importance of visualization in helping marketing teams understand and convey customer behavior patterns intuitively, which is extremely helpful in developing appropriate strategies.

Based on the segmentation results, the "Recent Buyers" category has the largest number of customers, 1,460 customers, who have recently made a purchase. For this segment, the appropriate strategy is to provide quick follow-up, such as thanking them through messages or emails, and offering discounts for the next purchase. This segment shows active interest in products, so offering discounts or referral programs can increase their chances of becoming more loyal customers. A study by Sudirjo et al. (2023) supports the idea that personalized marketing messages can increase customer retention. The "Hibernating" segment, consisting of 244 customers who have not been active for a long time, requires a reactivation strategy, such as sending reminder messages or offering special discounts to encourage them to shop again. Meanwhile, the "Champions" and "Need Attention" segments, each consisting of two customers, contribute significantly to loyalty and revenue, so loyalty reward strategies and product bundling promotions are highly recommended. This is supported by research by Dewa & Tranggono (2022), which suggests that effective marketing communication strategies can maintain customer loyalty.

Other smaller segments, such as "Promising" and "Potential Loyalists," require special attention to increase their loyalty, through offering loyalty programs such as lottery coupons. For the "About to Sleep" segment, an effective marketing strategy is to offer promotions to encourage them to shop again.

CONCLUSION

The conclusions drawn from this study reveal that the customer segmentation of Dhiza Petshop Tanjung Selor can be effectively identified using RFM (Recency, Frequency, Monetary) analysis, which shows that the majority of customers are purchasing products for pet cats. The behavior patterns of most customers reflect recent purchasing activity, but only a few can be categorized as loyal customers. This suggests that while Dhiza Petshop has great potential to retain a significant customer base, the company faces challenges in establishing strong, long-term relationships with customers. This situation can be interpreted as Dhiza Petshop's success in attracting a large number of new customers, but the customer retention process still requires more attention. Therefore, it is important for the company to focus more on retention and reactivation strategies to strengthen long-term relationships.

The application of RFM analysis using Python and Matplotlib has proven effective in segmenting customers into 11 segments based on their purchasing habits. This approach not only improves efficiency and accuracy but also facilitates a better understanding of customer behavior patterns that can support more targeted and data-driven marketing strategies. The segmentation results are used to design more effective marketing strategies tailored for each customer segment. For the "Recent Buyers" segment, which has the largest number of customers, the marketing strategy applied includes quick follow-up through thank-you messages or offering discounts for future purchases, to increase the likelihood of converting them into more loyal customers. Meanwhile, customers in the "Hibernating" segment, who have not made purchases for a long time, need reactivation strategies such as sending reminder messages, offering special discounts, or providing free promotions to entice them back. Customers in the "Champions" and "Need Attention" segments, though small in number, play an important role in revenue and loyalty, so they should be given more attention through loyalty rewards or exclusive promotions. The "Promising" and "Potential Loyalists" segments require loyalty programs to increase engagement and encourage them to become loyal customers.

Strategies for the "About to Sleep" segment include offering promotions to encourage them to make another purchase. Conversely, no customers were found in the "Loyal", "Cannot Lose Them", or "At Risk" segments because no behaviors matched the criteria for those segments.

This study has several limitations, including the use of customer transaction data that only includes customer IDs, invoice numbers, transaction amounts, and transaction dates, without incorporating demographic data or customer preferences. Nevertheless, focusing on transaction data alone is sufficient to generate relevant segmentation through RFM analysis. Additionally, data visualization only utilized bar charts and pie charts without exploring other types of visualizations, although these simple visualizations were adequate for supporting the interpretation of the analysis results. This study also relies on historical transaction data, so it does not fully reflect real-time changes in customer behavior. Therefore, while this analysis provides an overview of customer behavior patterns, it is more relevant for short-term and medium-term marketing strategy planning. Furthermore, while the marketing strategies developed based on RFM segmentation have been tailored to previous research findings, their effectiveness in improving sales or customer loyalty has not been directly tested. This study also did not compare the use of Python and Matplotlib with other software, although both have proven to be efficient and flexible tools for processing data and generating relevant visualizations.

The implications of this study can be divided into practical and theoretical categories. Practically, this study provides direct guidance for Dhiza Petshop's managers in designing marketing strategies based on the customer segmentation conducted. By understanding the characteristics of customers, such as Recent Buyers, Champions, and Hibernating, managers can implement more specific and targeted strategies, such as loyalty programs, discounts, or more effective reactivation strategies. This has the potential to increase marketing efficiency and customer loyalty. RFM-based segmentation also helps managers allocate marketing resources more effectively by focusing attention on segments with significant potential contributions to revenue.

Theoretically, this study enriches the literature on the implementation of RFM analysis in the retail industry, particularly in the petshop sector, which has not been widely studied. By using Python and Matplotlib as analytical tools, this study demonstrates how data analytics technology can be applied in small- to medium-sized businesses. The findings of this study also support and strengthen the theory of data-driven decision-making in marketing, which emphasizes the importance of understanding customer behavior through historical data analysis to develop relevant, evidence-based marketing strategies.

REFERENCE

Agustini, N. K. Y. (2003). Segmentasi Pasar, Penentuan Target dan Penentuan Posisi. *Equilibrium, Jurnal Ekonomi-Manajemen-Akuntansi*, *1*(2), 91–106. http://journal.uwks.ac.id/index.php/equilibrium/article/download/169/159

Ajzen, I. (2005). *Attitudes, Personality and Behavior*. USA: Open University Press.

Ananda, T. A., Dewi, N. K., & Saleh, M. Z. (2023). Fenomena Perubahan Strategi Pemasaran dalam Menghadapi Tantangan di Era Digital. *Jurnal Publikasi Ilmu Manajemen (JUPIMAN)*, *2*(4), 98–107.

Asgandi, S. (2020). *Analisis Strategi Pemasaran Dalam Meningkatkan Penjualan Pada Toko Kurnia Petshop Banjarmasin*. http://eprints.uniska-bjm.ac.id/2151/1/ARTIKEL ILMIAH.Pdf, 57–67

Astandi. (2024). Penerapan Matplotlib dalam Visualisasi Data untuk Analisis Hubungan Penggunaan Gadget dan Hasil Belajar. *Journal of Digital Business and Information Technology*, 29–39. https://doi.org/https://doi.org/10.23971/jobit.v1i1.204

Astuti, D. P. (2018). *Strategi Pemasaran Dalam Menarik Minat Konsumen Di Tinjau Dari Etika Bisnis Islam (Studi Kasus pada Pedagang Bakso dan Mie Ayam di Desa*

Sidodadi Sekampung Lampung Timur).

Bayhaqi, H., & Aslami, N. (2022). Indentifikasi Pasar, Segmen Dan Target Pasar Sasaran Bisnis Asuransi. *MAMEN: Jurnal Manajemen*, *1*(1), 111–118. https://doi.org/https://doi.org/10.55123/mamen.v1i1.48

Blattberg, R. C., Malthouse, E. C., & Neslin, S. A. (2009). Customer Lifetime Value: Empirical Generalizations and Some Conceptual Questions. *Journal of Interactive Marketing*, *23*(2), 157–168. https://doi.org/https://doi.org/10.1016/j.intmar.2009.02.005

Chaffey, D., & Chadwick, F. E. (2024). Digital Marketing Implementation and Practice. *In Digital Marketing Technologies*. https://doi.org/https://doi.org/10.1007/978-981-97-0607-5_3

Christian, Y., Yap, K. O., & Qi, R. (2022). Analisa Efisiensi Pendekatan Data-Driven Dalam Proses Segmentasi Pasar Dengan Studi Kasus Startup. *Infotech Journal*, *8*(2), 147–156. https://doi.org/https://doi.org/10.31949/infotech.v8I2.3732

Devenport, T. H. (2006). Competing on Analytics. *Harvard Business Review*, *84*(3), 144. http://journal.uwks.ac.id/index.php/equilibrium/article/download/169/159

Dewa, R. P., & Tranggono, D. (2022). Strategi Komunikasi Pemasaran Pit-Stop Kopi KIG Gresik dalam Mempertahankan Loyalitas Pelanggan. *Jurnal Nomosleca*, *8*(1), 30–45. https://doi.org/https://doi.org/10.26905/nomosleca.v8i1.6569

Elvida, R., Wandi Al-Hafiz, N., Hasim Siregar, M., & Kuantan Singingi, I. (2021). Sistem Informasi Rekam Medis Hewan Peliharaan Berbasis Web. *Prosiding Seminar Nasional Hasil Penelitian Dan Pengabdian Kepada Masyarakat*, 46–52. http://ejournal.uniks.ac.id/index.php/ProsidingUniks/article/view/1963

Febriana, T. F., & Rosadi, M. I. (2024). Analisis Segmentasis Pelanggan Toko Sepatu Di E-Commerce Menggunakan Algoritma K-Means Klastering. *JATI(Jurnal Mahasiswa Teknik Informatika)*, *8*(6), 12069–12072.

Gazali, S. (2017). Analisis Pengaruh Bauran Pemasaran Jasa Terhadap Keputusan Mahasiswa Dalam Memilih Sekolah Tinggi Ilmu Ekonomi Madani Balikpapan. *Jurnal Akuntansi Manajemen Madani*, *1*(3), 83–95.

Hamdat, A. (2020). *Buku Manajemen Pemasaran Dan Perilaku Konsumen, Manajemen Dan Strategi Pemasaran Dalam Bisnis*.

Jamal, J., & Yanto, D. (2019). Analisis RFM dan Algoritma K-Means untuk Clustering Loyalitas Customer. *Fakultas Teknik Universitas Panca Marga Probolinggo*, *9*(1), 1–8.

Jamal, Jamal, & Yanto, D. (2019). Analisis RFM dan Algoritma K-Means untuk Clustering Loyalitas Customer. *Fakultas Teknik Universitas Panca Marga Probolinggo*, *9*(1), 1–8.

Jeffery, M. (2010). *Data-Driven Marketing: The 15 Metrics Everyone in Marketing Should Know*. papers2://publication/uuid/1095FC9F-30FD-4F01-8201-26E748727D55

Junaidi, S., Devegi, M., & Kurniawan, H. (2023). Pelatihan Pengolahan dan Visualisasi Data Penduduk menggunakan Python. *ADMA : Jurnal Pengabdian Dan Pemberdayaan Masyarakat*, *4*(1), 151–162. https://doi.org/https://doi.org/10.30812/adma.v4i1.2963

Kohavi, R., & Parekh, R. (2004). Visualizing RFM segmentation. *SIAM Proceedings Series*, 391–399. https://doi.org/https://doi.org/10.1137/1.9781611972740.36

Kotler, P., & Keller, K. L. (2016). Marketing Management. *Boletin Cultural e Informativo - Consejo General de Colegios Medicos de España*, *22*.

Krisnawati, D. (2018). Peran Perkembangan Teknologi Digital Pada Strategi Pemasaran Dan Jalur Distribusi Umkm Di Indonesia. *Jurnal Manajemen Bisnis Krisnadwipayana*, *6*(1). https://doi.org/https://doi.org/10.35137/jmbk.v6i1.175

Kurniawan, D. (2020). Pengenalan Machine Leaning dengan Python. Solusi untuk Permasalahan Big Data. In *Alex Media Komputindo*.

Lubis, N. S., & Nasution, M. I. P. (2023). Perkembangan Teknologi Informasi Dan Dampaknya

Pada Masyarakat. *KOHESI: Jurnal Multidisplin Saintek*, *1*(12), 41–50. https://ejournal.warunayama.org/index.php/kohesi/article/view/1311

Maidla, M. (2023). *No TitleUtilising Machine Learning And Rfm Analysis For Customer Retention In an Online Grocery Delivery Startup*. University of Taru.

Maskanah, I. (2020). Segmentasi Pelanggan Toko Purnama dengan Algoritma K-Means dan Model RFM untuk Perancangan Strategi Pemasaran. *INOVTEK Polbeng - Seri Informatika*, *5*(2), 218. https://doi.org/https://doi.org/10.35314/isi.v5i2.1443

Maulani, & Nadhifatul Nur, A. A. W. (2024). Analisis Pemanfaatan Data Analytics Dalam Pengambilan Keputusan Bisnis Di Pt Xyz. *Jurnal Ekonomi Revolusioner*, *7*(7), 164–170. https://oaj.jurnalhst.com/index.php/jer/article/view/86/84

Muliyah, P., Aminatun, D., Nasution, S. S., Hastomo, T., Wahyuni, S. S., & Sitepu, T. (2020). Studi kelayakan Bisnis. *Journal GEEJ*, *7*(2).

Prasetyo, S. S., Hakim, A. R., Statistika, D., Sains, F., & Diponegoro, U. (2020). *3 1,2,3*. *9*, 421–433.

Putri Nugraha, J., Alfiah, D., Sinulingga, G., Rojiati, U., Saloom, G., Rosmawati, F., Johannes, R., Kristia, Batin, M., Jati Lestari, W., Khatimah, H., & Beribe, M. (2021). *Perilaku Perilaku Konsumen Teori*.

Romzi, M., & Kurniawan, B. (2020). Pembelajaran Pemrograman Python Dengan Pendekatan Logika Algoritma. *JTIM: Jurnal Teknik Informatika Mahakarya*, *03*(2), 37–44.

Salmah, H. (2018). Analisis Market Segmentation, Targeting, Dan Positioning Transportasi Online Grab Di Kota Bogor (Studi Kasus Pada Mahasiswa Universitas Pakuan). *JIMFE (Jurnal Ilmiah Manajemen Fakultas Ekonomi)*, *4*(2), 81–96.

Siahainenia, S. E. T. (2020). Analisis Pengaruh Orientasi Pesaing, Diferensiasi Produk, Positioning Produk Terhadap Kinerja Produk Gula Rendah Kalori Merek Tropicana Slim Di Kota Ambon. *Jurnal Masohi*, *1*(1), 42–57.

Sitorus, E. R., & Nugraha, I. (2025). Analisis segmentasi pelanggan dengan model RFM (Recency , Frequency , Monetary) dan K-Means Clustering (Studi kasus : PT XYZ). *JUTIN: Jurnal Teknik Industri Terintegrasi*, *8*(1).

Smith, W. R. (1956). Product Differentiation and Market Segmentation as Alternative Marketing Strategies. *Journal of Marketing*, *21*(1), 3. https://doi.org/https://doi.org/10.2307/1247695

Sudirjo, F., Purwati, T., Widyastuti, W., & Budiman, Y. U. (2023). Analisis Dampak Strategi Pemasaran Digital dalam Meningkatkan Loyalitas Pelanggan: Perspektif Industri E-commerce. *Jurnal Pendidikan*, *7*, 7524–7532. https://jptam.org/index.php/jptam/article/view/7422%0Ahttps://jptam.org/index.php/jptam/article/download/7422/6136

Sugiyono. (2020). *Metodologi Penelitian Kuantitatif, Kualitatif dan R & D*.

Sutresno, S. A., Iriani, A., & Sediyono, E. (2018). Metode K-Means Clustering dengan Atribut RFM untuk Mempertahankan Pelanggan. *JuTISI*, *4*(3), 433–440. https://doi.org/http://dx.doi.org/10.28932/jutisi.v4i3.878

Syah, A. (2020). Penerapan Web Walkers Sebagai Media Informasi Untuk Perbandingan Manual Brewing Coffee Di Indonesia. *JSiI (Jurnal Sistem Informasi)*, *7*(2), 132–137. https://doi.org/https://doi.org/10.30656/jsii.v7i2.2507

Syah, A. (2023). Strategi Pemasaran. In *NBER Working Papers*. http://www.nber.org/papers/w16019

Utama, A. A., & Ardiansah, I. (2023). Analisis Segmentasi Menggunakan Metode Rfm Dalam Mengidentifikasi Perilaku Pelanggan. *SemanTIK : Teknik Informasi*, *9*(2), 107. https://doi.org/https://doi.org/10.55679/semantik.v9i2.44945

Zaki, M. (2018). Pengaruh Influencer Marketing Sebagai Strategi Pemasaran Digital Era Moderen. *Jurnal Manajemen Dan Inovasi (MANOVA)*, *1*(2), 14–23. https://doi.org/https://doi.org/10.15642/manova.v1i2.350

APPENDIX

Python Code:

1. Install and Import Libraries
2. Import Dataset

```
# Baca dataset
df = pd.read_excel('/content/DATA-PENELITIAN-_2_.xlsx')

print(df.head(5))  # 5 baris pertama
print("...\n")  # Pemisah
print(df.tail(5))  # 5 baris terakhir
```

3. Determining the Reference Date

```
# Tanggal referensi (tanggal terakhir dalam data)
reference_date = df['Invoice Date'].max()

print("Tanggal referensi:", reference_date)
```

4. RFM Score Calculation

```
# Menghitung Recency, Frequency, dan Monetary
rfm = df.groupby('Customer ID').agg({
    'Invoice Date': ['count', 'max'],  # Frequency dan max untuk Recency
    'Amount': 'sum'                    # Monetary
}).reset_index()

# Membuat kolom Recency berdasarkan Invoice Date max
rfm['Recency'] = (reference_date - rfm[('Invoice Date', 'max')]).dt.days

# Rename kolom agar lebih jelas
rfm.columns = ['Customer ID', 'Frequency', 'Max Invoice Date', 'Monetary', 'Recency']

# Drop kolom yang tidak diperlukan
rfm = rfm.drop(columns=['Max Invoice Date'])

print(rfm.head(5))  # 5 data pertama
print("...\n")  # Pemisah
print(rfm.tail(5))  # 5 data terakhir
```

5. RFM Data Scoring

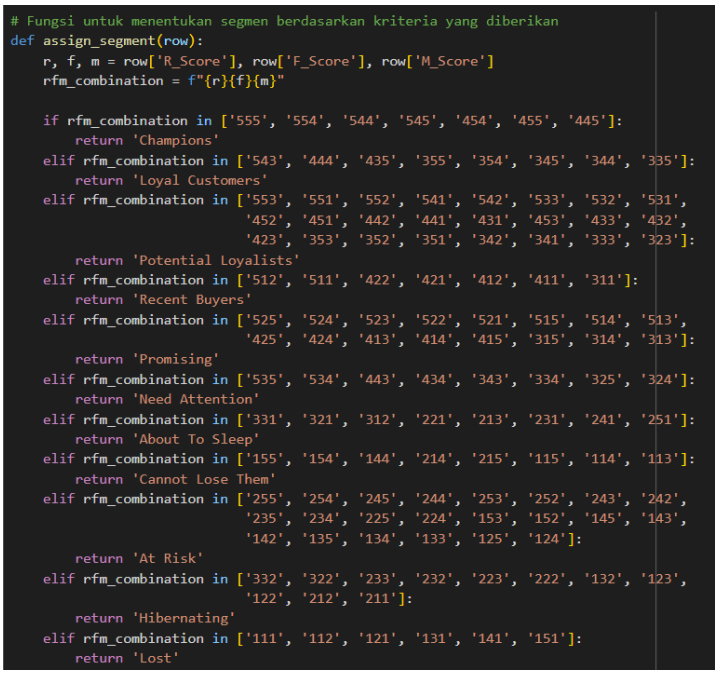
```
# Fungsi scoring menggunakan pd.cut
rfm['R_Score'] = pd.cut(rfm['Recency'], bins=5, labels=[5, 4, 3, 2, 1], include_lowest=True)
rfm['F_Score'] = pd.cut(rfm['Frequency'], bins=5, labels=[1, 2, 3, 4, 5], include_lowest=True)
rfm['M_Score'] = pd.cut(rfm['Monetary'], bins=5, labels=[1, 2, 3, 4, 5], include_lowest=True)

print(rfm.head(5))  # 5 data pertama
print("...\n")  # Pemisah
print(rfm.tail(5))  # 5 data terakhir
```

```
# Install library
!pip install pandas matplotlib seaborn

# Import library
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

6. Customer Segmentation Process



```
# Fungsi untuk menentukan segmen berdasarkan kriteria yang diberikan
def assign_segment(row):
    r, f, m = row['R_Score'], row['F_Score'], row['M_Score']
    rfm_combination = f"{r}{f}{m}"

    if rfm_combination in ['555', '554', '544', '545', '454', '455', '445']:
        return 'Champions'
    elif rfm_combination in ['543', '444', '435', '355', '354', '345', '344', '335']:
        return 'Loyal Customers'
    elif rfm_combination in ['553', '551', '552', '541', '542', '533', '532', '531',
                             '452', '451', '442', '441', '431', '453', '433', '432',
                             '423', '353', '352', '351', '342', '341', '333', '323']:
        return 'Potential Loyalists'
    elif rfm_combination in ['512', '511', '422', '421', '412', '411', '311']:
        return 'Recent Buyers'
    elif rfm_combination in ['525', '524', '523', '522', '521', '515', '514', '513',
                             '425', '424', '413', '414', '415', '315', '314', '313']:
        return 'Promising'
    elif rfm_combination in ['535', '534', '443', '434', '343', '334', '325', '324']:
        return 'Need Attention'
    elif rfm_combination in ['331', '321', '312', '221', '213', '231', '241', '251']:
        return 'About To Sleep'
    elif rfm_combination in ['155', '154', '144', '214', '215', '115', '114', '113']:
        return 'Cannot Lose Them'
    elif rfm_combination in ['255', '254', '245', '244', '253', '252', '243', '242',
                             '235', '234', '225', '224', '153', '152', '145', '143',
                             '142', '135', '134', '133', '125', '124']:
        return 'At Risk'
    elif rfm_combination in ['332', '322', '233', '232', '223', '222', '132', '123',
                             '122', '212', '211']:
        return 'Hibernating'
    elif rfm_combination in ['111', '112', '121', '131', '141', '151']:
        return 'Lost'
```