Determining Marketing Mix of CV Nutri Pro by Using Big Data Analytics

Vira Maulina Putri[1], Manahan Siallagan[2]

[1,2] Master of Business Administration, Institut Teknologi Bandung

ABSTRACT: Technological growth supports the acceleration of the health industry. Technology provides an opportunity for business actors to convey product advantages to be disseminated widely through digital media. Apart from providing benefits in the easy dissemination of information, digital media can be a sales media for the health industry. Based on data from Tokopedia (e-commerce with the most users in Indonesia), the biggest sales are dominated by health products. Large amounts of data (big data) available in e-commerce can be extracted using the Web Scraping method. Big data can be processed to gain certain insights in achieving competitive advantage. CV Nutri Pro as a medium-sized business has limited data which causes the marketing mix that has been prepared beforehand to be incomplete. This condition causes out of sync, where there are demands that cannot be fulfilled. Based on the opportunity to utilize big data, CV Nutri Pro can determine a comprehensive renewable marketing mix. Each aspect of the marketing mix (4Ps) will be processed using big data analytics. The methods used include Pivot Data, K-Means, and Multidimensional Scaling (MDS). This research provides new insights for the company to renew marketing mix.

KEYWORDS: Data Analytics, E-Commerce, K-Means, Marketing Mix, MDS, Pivot Data, Web Scraping.

INTRODUCTION

The current technological growth is very influential on the health product industry. The need for health products increases with the easy dissemination of information related to product benefits. Data from Tokopedia (2022) as an Indonesian's e-commerce, state that the most product sales are health products, food and beverages, household appliances, body care, and automotive. The large demand for health products on Tokopedia, prompted this industry has a big opportunity to sustain and grow their company. The high consumers demand for health products on Tokopedia has opened up opportunities for the health product industry to develop marketing strategies that provide a competitive advantage. In addition, based on e-commerce User Data in Indonesia in 2022, it was found that Tokopedia also became e-commerce with the highest number of users, 157.2 million. The big data that available in e-commerce uses to determine certain strategies. The large amounts of digital data are able to assist in making marketing strategies that increase competitive advantage [1]. In business, the benefits of digital big data allow companies to carry out marketing analysis of markets, customers, products and competitors [2].

Based on the opportunity and the use of big data, CV Nutri Pro as a health product producer of snakehead fish extract wants to gain an advantage of that condition. CV Nutri Pro produces a snakehead fish extract with low durability compared to its competitors. The product can only last 10 months in the freezer, 2 days in the fridge, and 24 hours at room temperature. However, the product has the advantage of combining the benefits of snakehead fish extract with selected spices. CV Nutri Pro's product contains albumin from snakehead fish plus the nutritional value of spices that have antioxidants. The antioxidants in spices and herbs are very effective because they possess excellent activity of antioxidant [3]. Antioxidants help protein absorption to provide its role in the immune system, growth and repair of wound healing fiber [4]. That situation makes CV Nutri Pro's product excellence than competitors' products which are purely made from snakehead fish extract. Competitors that have been indicated at this time include serum albumin, capsules, syrup and pure essence. The competitor product is a product that has a high durability. Albumin infusion is often used by Hospitals to increase albumin in hypoalbuminemia patients. However, serum albumin has a poor correlation with low nutrient intake levels which causes a rapid increase in albumin levels but also a rapid decline [5]. Another competitor's product is albumin capsules, syrup and pure essence which are often sold in Pharmacies and e-commerce. In the existing conditions, CV Nutri Pro has carried out marketing through the Website and Instagram. Expansion of product information has occurred very quickly and extensively. However, the company cannot fulfil end customer demands if they are not in the same city as the producer or hospital partner as a reseller and also an additional inventory.

BUSINESS ISSUE

CV Nutri Pro is a company that wants to increase its competitive advantage in the health product market. The company team believes that their excellence in terms of product composition is able to encourage wider sales. On a certain side, the marketing staff of company has carried out several marketing activities to reach a wider market. These activities include Word of Mouth (WoM), brochure and especially digital marketing (Instagram and Website). The marketing that has been carried out has the effect of expanding the market, but sometimes market demands cannot be fulfilled because there are no partners in the end customer's city. Therefore, wider sales can be performed by expanding sales partners (hospitals, pharmacies, or other resellers). This condition causes fluctuations in sales as shown in Figure I.
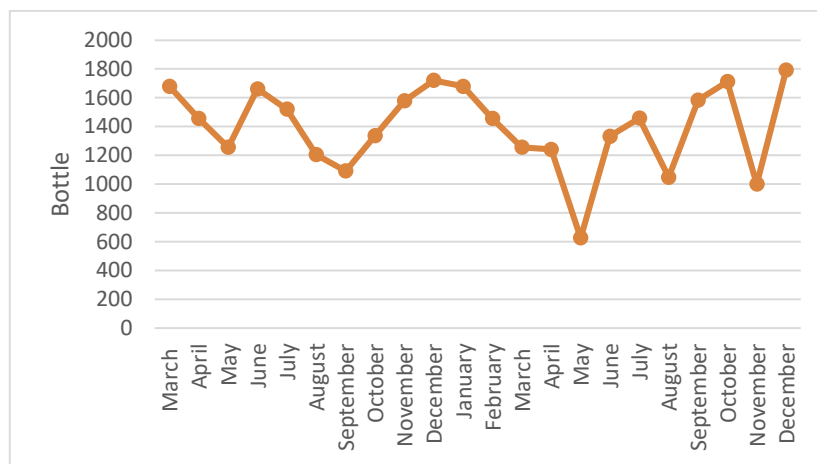

Figure I. Sales on March 2021 - December 2022

Figure I. describe significant fluctuations in the period March 2021 – December 2022. Fluctuating sales is a driving force of the company's desire to increase market growth by fulfilling individual purchases beyond the current reached of producer and partners. It is necessary to add partners (hospital, pharmacy, and reseller) by looking at the market needs regarding the products, prices, places with much interest or high demand. If doing Problem Exploration about the causes of sales fluctuations, the root causes will as show in Figure II.
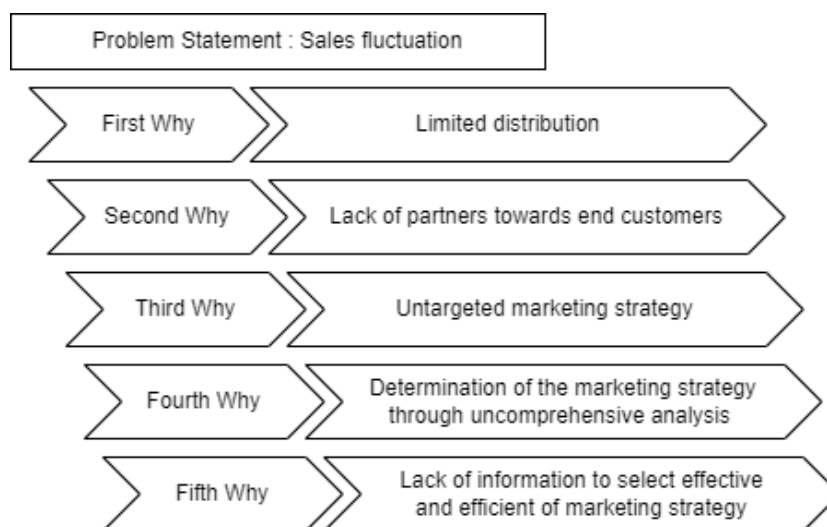

Figure II. Five Whys Analysis

Based on Figure II, the Five Whys Analysis state that the root of the problem is the lack of information to determine an effective and efficient marketing strategy. The marketing strategy covers all aspects of the Marketing Mix (Product, Price, Place, and Promotion). The currently implemented Marketing Mix is shown in Table I.

Table I. Existing Marketing Mix

No	Marketing Mix	Existing Condition
1.	Product	The available products have a combined composition of snakehead fish extract and selected spices in essence form. This product has low durability.
2.	Price	Prices are determined based on Cost of Goods Sold (COGS) regardless of the average market price.
3.	Place	Limited sales in cities that have acquaintances with the owner regardless of high or low demand.
4.	Promotion	The selling price is standardized without any concessions at certain times.

Existing conditions indicate a lack of research and data to determine a more objective strategy in determining the marketing mix. Facing that situation, it is necessary to renew the existing marketing mix which is not competent yet. As a medium-sized company, CV Nutri Pro has limited data to determine an effective marketing mix strategy. This research aims to provide solutions using methods which can analyze digital data that is widespread in e-commerce. The e-commerce data that will be used is Tokopedia. Based on these data, appropriate segmentation will be carried out in determining the product, price, place, and promotion accepted by the market.

LITERATURE REVIEW
Marketing Mix
For company which offer a product, Marketing Mix is a market determination strategy with a combination of controlled variables consisting of product, place, price, and promotion (Eti & Bari, 2020).

1. Product

   Product is things that will be offered to the market in accordance with what customers are interested in and are willing to pay for.

2. Price

   Price is the value of the product in exchange for a certain amount of currency. Pricing takes into account the factors of fixed costs, competition, company objectives, and the proposed position (competitive advantage).

3. Place

   Place means for products to be accessed or used easily. Place is also a distribution channel that is significant in the means of delivering product value. In addition, storage maintenance, inventory, and product distribution costs must be presented at the right place, at the right time and in the right amount.

4. Promotion

   Promotion is a way to build relationships with customers that encourage customers to make purchases at the right time. Promotion includes advertising, sales promotion, events & experiences, public relations & publicity, direct marketing, word of mouth, and personal selling. In this research, only focuses on sales promotion.

The Intelligence-based era, provide big data with huge amount of data that allows to create and improve every aspect of the Marketing Mix [7]. Based on research of Saini & Rajesh (2022), creating a framework for implementing the use of big data to improve the function of the marketing mix. Utilization of big data in increasing the effectiveness and efficiency of implementing the Marketing Mix will be applied in this research.

Web Scraping

Web scraping is the extraction method of raw data which includes product, text, and graphic images [8]. Primary Data Collection can be solved by doing web scraping as an automated data collection using the Python programming language [9]. The extracted data can be in JSON, XML, CSV, XLS or RSS format [10]. In this research, web scraping was carried out on Tokopedia in XLS format.

Knowledge Discovery in Database (KDD)

KDD is the process of identifying a hidden pattern in certain data into actionable information as data analytics process [11]. KDD provides a descriptive analysis of each existing feature to find out the cause and further explanation [12]. As descriptive analytics, the KDD Framework is the most suitable research methodology to be implemented for this research. KDD is process consist of selection, pre-processing, transformation, data mining, and interpretation or evaluation [13].

Selection

Selection is processed by taking the target data in the data store. The selected data is data that is relevant to the knowledge discovery process that has been determined [11].

Pre-processing

Improving the quality of dataset. In certain analyses, if there are data that have significant distances, they can be eliminated using the outlier data formula. The outlier data formula developed by Tukey identifies the 25th percentile as Q1, the 75th percentile as Q3, and the difference between Q3 and Q1 is the Inter-Quartile Range [14].

Inter-Quartile Range (IQR) = Q3 - Q1
Q3 = Percentile 0.75, Q1 = Percentile 0.25
Q1-1.5 x IQR | Q3+1.5 x IQR

Transformation

Analyzing the variables, importance and correlations. There are several steps in transforming data including data cleaning, data grouping, changing data types, dimension reduction, and Principal Component Analysis (PCA). In grouping data can be done by using Pivot Table. Pivot Table is a way to summarize and group large data sets into concise reports by quickly filtering data [15]. Transformation data in K-Means Algorithm is usually used PCA to reduce the data into two categories. On research Clayman et al. (2020), using PCA on the K-Means method with the aim of selecting relevant features and helping K-Means clustering work well by using low and effective dimensions. The application of PCA helps to overcome the weakness of the K-Means algorithm with the attribute reduction process by eliminating irrelevant features but has a low risk of losing information [16].

Data Mining

Conducting machine learning algorithms to extract structures, correlations, patterns and rules based on data (data-driven). Clustering method is conducted on this research. Clustering support to segment the data that has been obtained and each cluster requires a different way of responding [17]. The data mining method will be conduct by using algorithm are Pivot Data, K-Means and MDS methods. Pivot Tables are the easiest way to cluster and analyze the big data that provide easy-to-understand interpretations [18]. Then, K-Means and MDS that uses certain algorithm formula that will be represented by command in Python.

RESEARCH METHODOLOGY

Data Collection Method

In the past, to analyze those aspect, the researcher only has an alternative which is survey to gain large information. Otherwise, in this era, strategic decision making can utilize the advantage of big data and the results will be more objective. Big data on e-commerce will be obtained by web scraping method. The primary data used comes from external big data in Tokopedia. The selection of keywords for data retrieval is based on the top keywords in the search engine on Tokopedia ("Ekstrak Ikan Gabus"). Data stored from Tokopedia until January 5, 2023. Table II. shows the data features to be used in this research.

Table II. Features in Data Collection

No	Features	Description
1.	Store	Store name
2.	Location	Location of store
3.	Product Name	Title of product in Tokopedia
4.	Selling price	Offered price
5.	Sold	Number of products sold
6.	Normal Price	Price before discount
7.	Rating	Valuation of products sold

Data collection for each data feature is based on the need for analysis of each aspect of the Marketing Mix which will be improved. The marketing mix that has been going on in order to meet the needs of customer comprehensively.

Data Analysis Method

Data analysis as descriptive research uses methodology of Knowledge Discovery in Database (KDD). It consists of five steps as shown in Figure III.
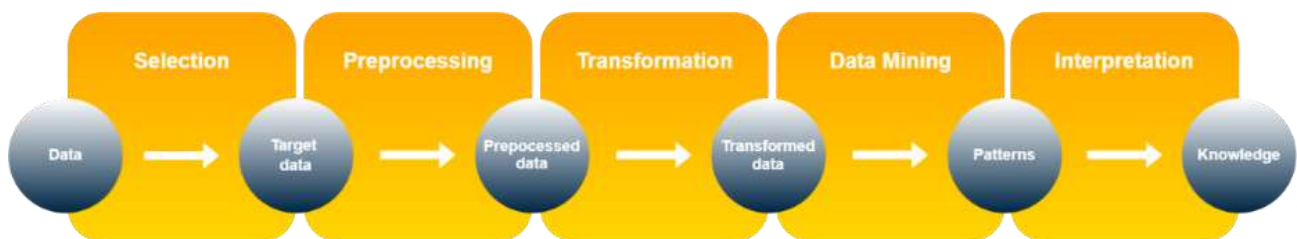

Figure III. Data Analysis Method[13]

This research uses KDD with the following stages:
1. Data and Source Selection
   This stage carries out the large data collection as known as big data. Each feature of data collection will be useful for further analysis.
2. Data Pre-processing
   The process in Data Pre-processing consists of several steps which include completing missing values, removing duplicates that may exist in the raw data, resolving data inconsistencies, removing outlier data and generating new attributes.
3. Data Transformation
   Data Transformation analyzes the variables related to the objectives for every aspect of the Marketing Mix. The importance, interaction, and correlation of each variable are determined at this stage.
4. Data Mining
   Data Mining is a core stage in KDD. This stage involves machine learning algorithms to obtain patterns based on data. The conducting method are Pivot Table, K-Means, and MDS
5. Interpretation
   This stage reveals the results of Data Mining in the form of Data Visualization. It aims to make it easy for understanding and containing insightful information or a knowledge.

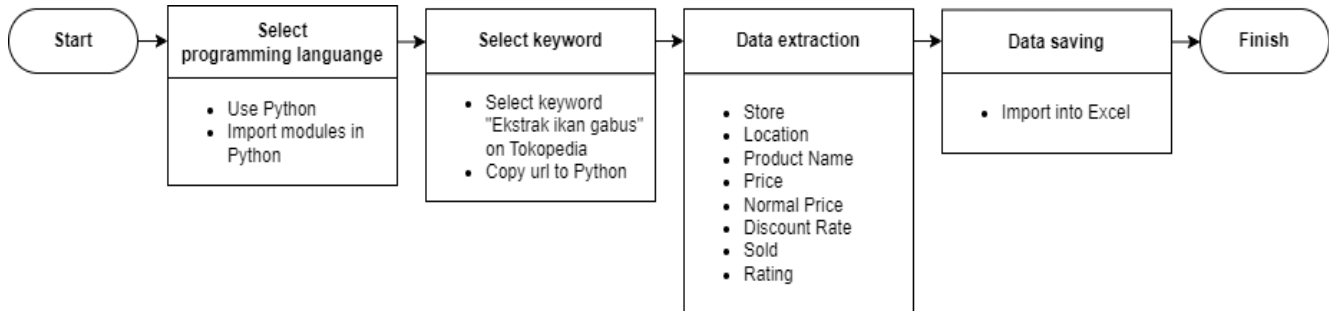## FINDING AND DISCUSSION

Selection

Figure IV. Selection Process

The selection stage initially by selecting the data retrieval. This research uses the python programming language in data collection. Data collection process started by importing several modules related to the Web Scraping method. In applying the Web Scraping Method, the keyword "Ekstrak Ikan Gabus" is used based on the top keywords on the Tokopedia search engine. The next step is data extraction performed by taking the data 'Store', 'Location', 'Product Name', 'Price', 'Normal Price', 'Discount Rate', 'Sold', 'Rating'. Each column of data obtained has a role in each aspect of the analysis to be carried out.

First, 'Product Name' and 'Sold' data are used to analyze product aspects. After product analysis, a new column category 'Product Categories' is obtained which will be used to analyze aspects of the other 4Ps. Second, 'Product Categories', 'Sold', and 'Price' data are used to analyze the price aspect. Third, 'Location', 'Store', and 'Sold' data are used to analyze the place aspect. Fourth, the 'Product Categories', 'Sold', 'Discount Rate' data obtained from the 'Price' calculation is divided by 'Normal Price' which is used for promotion analysis.

The 'Rating' data is used for finding the relevance of market interest and product image from the sale of Snakehead Fish Extract. 'Rating' data gives a good value, which is above 4.5. Information was obtained that products that have sales data get a rating above 4.5, which means the product has a good image or reputation on digital point of view. That condition explains the products with the keyword "Ekstrak Ikan Gabus" are eligible for further processing. The final step is data saving into Excel file (.XLSX). The data obtained consists of 7 columns and 1888 rows including category titles. It means there are 1887 information data that can be continued to the next stage with each respective function.
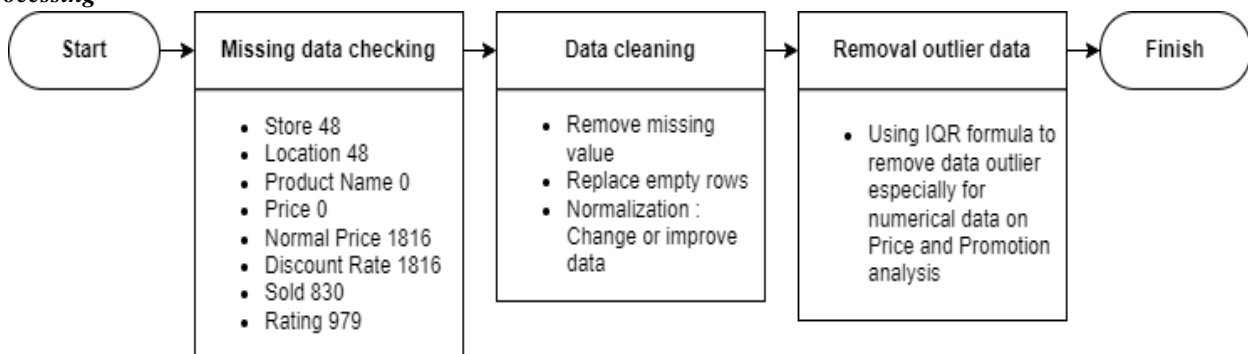
Pre-processing

Figure V. Pre-processing Process

The second stage is pre-processing. Export the Excel file to python. At this stage, the missing data is checked which can be eliminated or improved in quality. 'Store' and 'Location' have 48 missing data. 'Normal Price' and 'Discount Rate' have 1816 missing data. 'Sold' has 830 missing data. 'Rating' has 979 missing data. Then, data cleaning is carried out which includes the process of eliminating missing data, replacing empty columns, changing data type and improving data quality as in Table III.

Table III. Elimination and Improvement Data

No	Column Title	Raw Data	Elimination and Improvement Data
1.	Store	NaN	"not define"
2.	Location	NaN	"not define"
		"Kab."	" "
		"Badung"	"Bandung"
3.	Price	"Rp"	" "
		"."	" "
		String	Integer
4.	Rating	NaN	5.0
5.	Sold	NaN	0
		"Terjual"	" "
		"+"	" "
		"rb"	"000"
		String	Integer
6.	Normal Price	NaN	0
		"Rp"	" "
		"."	" "
		String	Integer

The steps of elimination and data improvement indicate there are several raw data that need to be improved. In the store column, each NaN data will be changed to not define, which means it is not involved in the next analysis. In the place column, each NaN data will be replaced with not define, the word "Kab." removed to support the analysis run well, and the wrong spelling or typo "Badung" was upgraded to "Bandung". In the price column, the currency is "Rp" and the sign "." removed to assist in changing the data type from a string or set of characters to an integer or numerical. In the rating column, each NaN data is changed to 5.0. NaN in the rating indicates no purchase yet. Change to 5.0 aims to normalizing the distribution of ratings. In the sold column, each NaN data is replaced with 0. Then, the words "Sold" and "+" are removed to support the changing data type from string to integer. Abbreviation "rb" are changed into "000". In the normal price column, each NaN data is replaced with 0. Then, the currency "Rp" and the sign "." are removed to support the changing data type from string to integer.

After cleaning the data, the next step is removing outlier data using the IQR formula (IQR = Q3 - Q1). Removal is performed on numerical data that will be used in price and promotion analysis in order to obtain normally distributed data. For prices, the data of 'Price' was had originally 1888 rows becomes 1848 rows. For promotion, the data of 'Discount Rate' was had originally 1888 rows becomes 1568 rows.

Transformation and Data Mining
1. Product

Product data processing has a target to determine product categories in the market. In determining product categories, the features will be used are 'Product Name' and 'Sold'. Category determination is based on the name of the product which has a description of the type of product. In the first stage, data transformation is carried out by replacing products that have double categories. In this research, there was one case of double categories. The product categories are "Kapsul" and "Sari". There are 48 products that have both categories. Based on observations on Tokopedia, products that have the information "Kapsul" and "Sari" are products of the "Kapsul" type filled with "Sari" powder of snakehead fish. Thus, the product is included in the "Kapsul" category. In the second stage, doing data mining to find out product categories by using Pivot Data. There are four categories that will be grouped: "Kapsul", "Madu", "Sari", and "None". Based on the categorization results obtained results as in Table IV.

Table IV. Product Categories

No.	Product Categories	Total
1.	"Kapsul" or Capsule	1307
2.	"None"	473
3.	"Madu" or Syrup	72
4.	"Sari" or Essence	36

2. Price

Price data processing has a target to segment prices that have prices with "Low demand" and prices with "High demand". Determination of price segmentation uses the features 'Price', 'Sold' (sales data), and product categories. The price data processing also considers the total income. Total income is obtained from the price multiplied by sold. 'Price', 'Sold', and total income are comparable variables. The addition of variables supports in increasing the accuracy of the data processing. In the first stage, the data transformation begins with the determination of codes to each product category. Label encoder applies to price and promotion data processing. The codes as shown in Table V.

Table V. Code of Product Category

Code	Product Category
0	"Kapsul"
1	"Madu"
2	"None"
3	"Sari"

Then, data transformation was carried out using the PCA method. The results of the PCA will generate two new data categories (p1 and p2) from the combined 'Price', 'Sold', total income and product categories as the features. From the new data obtained intervals as in Figure VI.
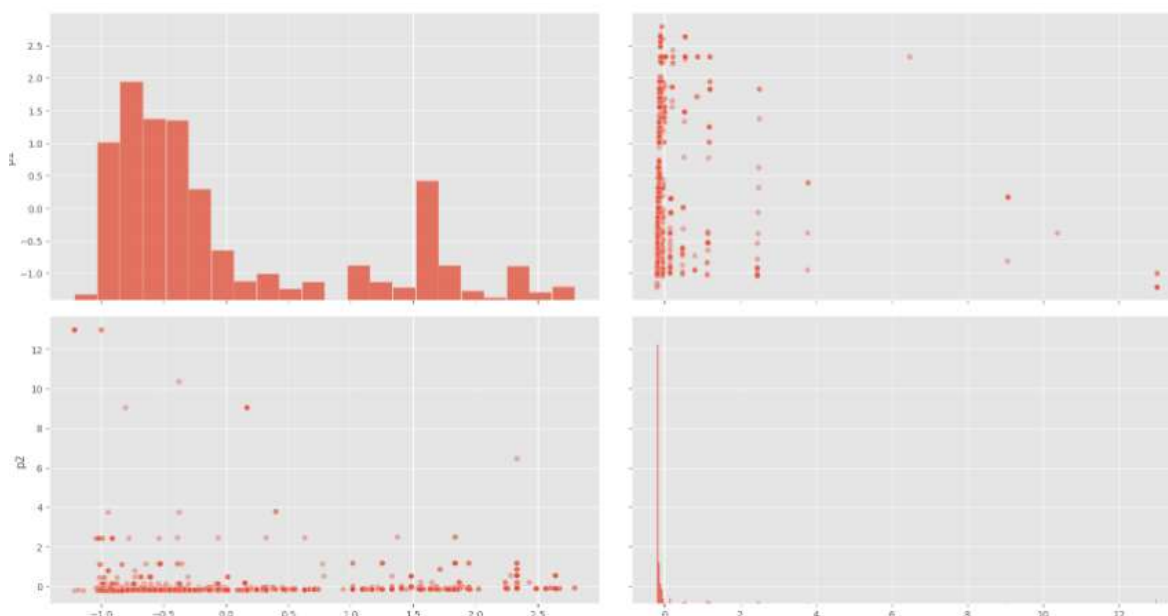

Figure VI. PCA Distribution on Price Segmentation

In the second stage, data mining was carried out using the K-Means method. Cluster determination uses a business approach based on the needs of segmentation analysis. Therefore, at this stage using two clusters: "Low demand" and "High demand".
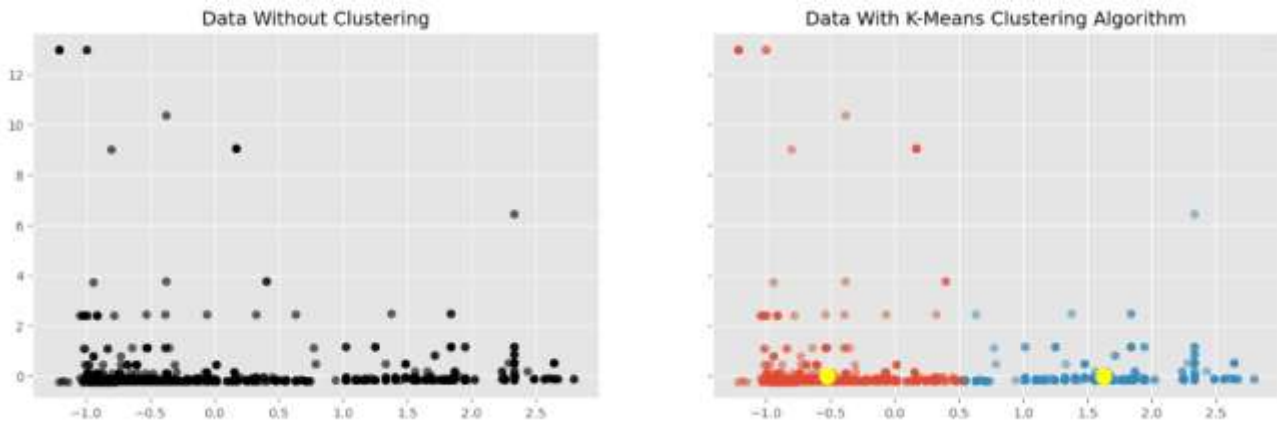
Figure VII. Price Segmentation

3. Place

Place data processing has a target to determine potential market expansion. Determination of potential market expansion uses the feature 'Location', 'Store' and 'Sold'. In the first stage, data transformation is carried out by doing encoder data to change data string into numerical or integer for each 'Location' and 'Store'. For instance, each store in "Jakarta Utara" has encoder 26 and each store "Berkah Natural Bersama" has encoder 121. In the second stage, data mining was carried out using the MDS method. First step is reducing the data to obtain data with lower dimensions (2-dimension). The new data become category "0" and "1" which makes it easier to process in machine learning. Second step, the results of the fusion of data reduction produce location segmentation using a scatter plot as shown in Figure VIII.

Figure VIII. Location Segmentation with MDS Clustering

4. Promotion

Promotion data processing has a target for segmenting discount rates, to define the discount rates that have "Low demand" and "High demand". Determination of discount rate segmentation uses the features 'Discount Rate', 'Sold', and product categories. Transformation is carried out by using PCA. The results of the PCA will generate two new data categories (p1 and p2) from the combined features used. From the new data obtained intervals as in Figure IX.
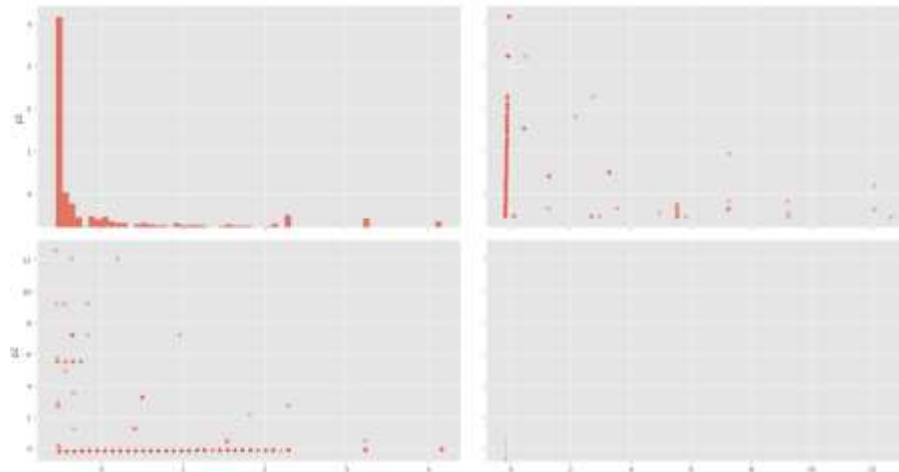
Figure IX. PCA Distribution on Discount Rate Segmentation

In the second stage, data mining was carried out using the K-Means method. Cluster determination uses a business approach based on the needs of segmentation analysis. Therefore, at this stage using two clusters: "Low demand" and "High demand".
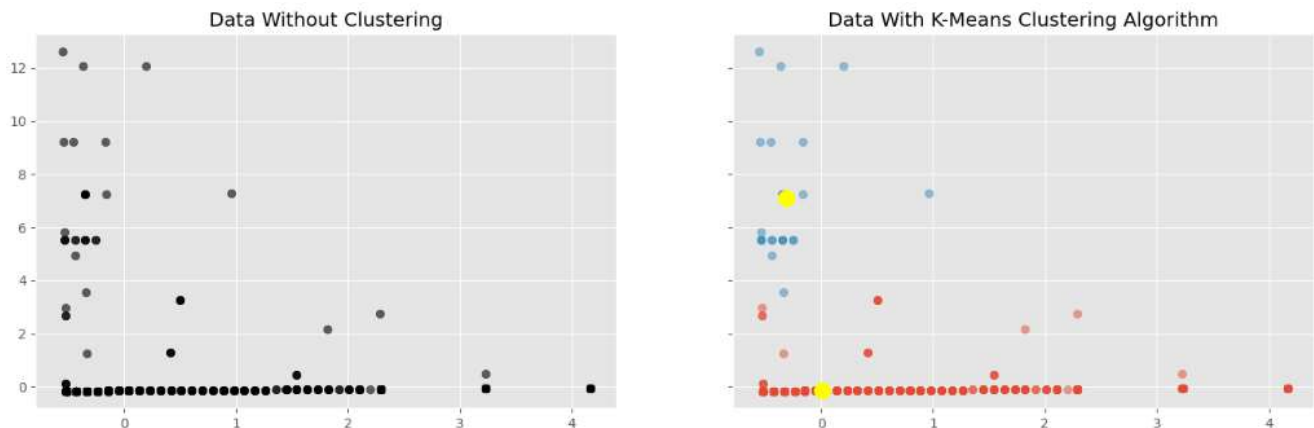

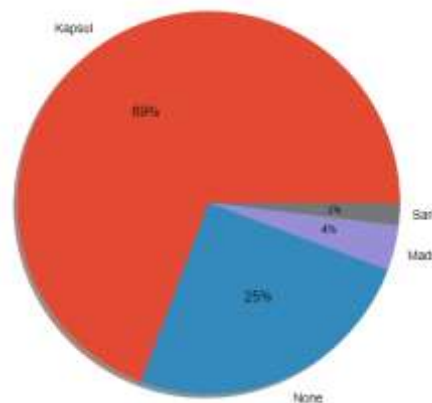**Figure X.** Discount Rate Segmentation

Interpretation
1. Product

Figure XI. Product Interpretation

Based on Figure XI, there are three types of products spreading in the market: "Kapsul", "Madu", dan "Sari". Data shows the product categories with the highest sales order "Kapsul" product category in the market by 69%. Followed by 4% "Madu" and 2% "Sari". The high number of "Kapsul" products in market is due to the large demand for this type of product. It has easy-way to consume and has a longer product durability compared to condensed and liquid products such as "Madu" and "Sari". Currently CV Nutri Pro has a product in the form of essence or "Sari" with superior product composition but has low product durability. The percentage obtained based on the results is intended to be the company's priority level when the company wants to carry out product development.

2. Price

Table VI. Price Segmentation

Cluster 0 : Low Demand			Cluster 1 : High Demand		
Code	Product Category	Price Range	Code	Product Category	Price Range
0	"Kapsul"	Rp114.000 - Rp 260.000	0	"Kapsul"	Rp600 - Rp 112.000
1	"Madu"	Rp115.000 - Rp 250.000	1	"Madu"	Rp28.000 - Rp 79.000
2	"Sari"	Rp125.000 - Rp 250.000	2	"Sari"	Rp5.000 - Rp 120.000
			3	"None"	Rp15.900 - Rp100.000

The data shows products that are selling well are in "Low demand", have a relative high price. Meanwhile, products that are "High demand" a relative low price than "Low demand" from each product.
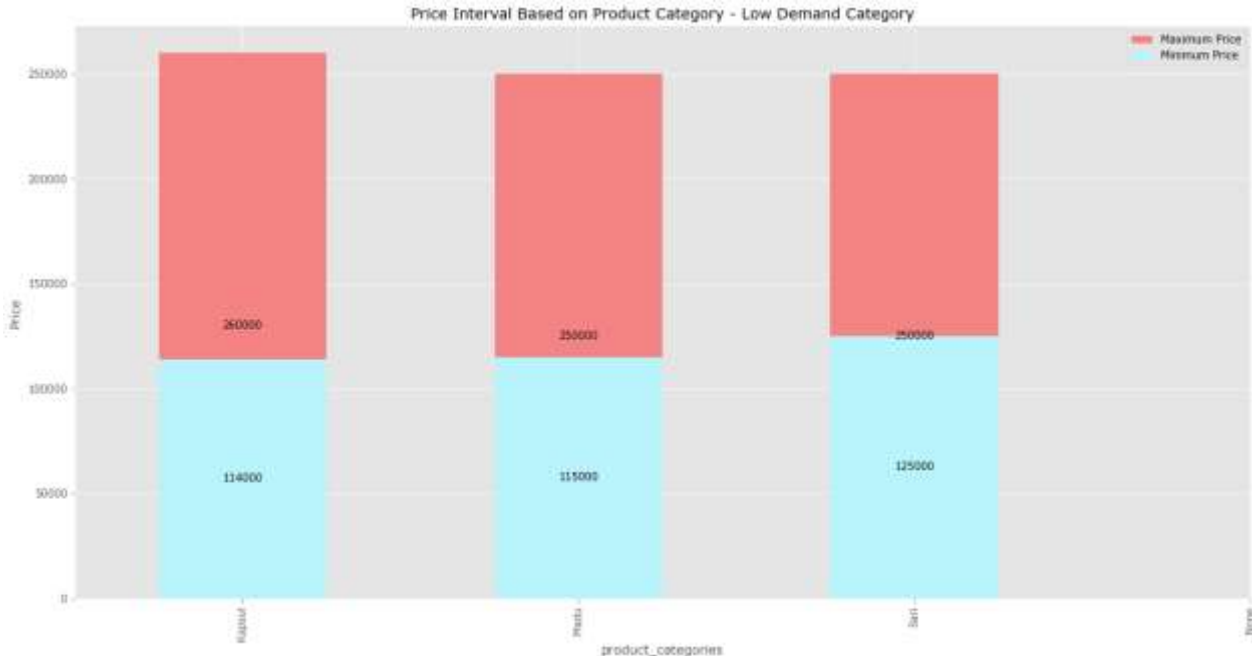

Figure XII. Price Cluster 0 Interpretation

Figure XII, shows the price with low demand. Prices with low demand are influenced by customers knowledge about competitive prices in the market by considering product volume and rating or reviews from previous customers who have made purchases. Based on the consumer's knowledge and their own assessment, the price range in cluster 0 has a low demand.
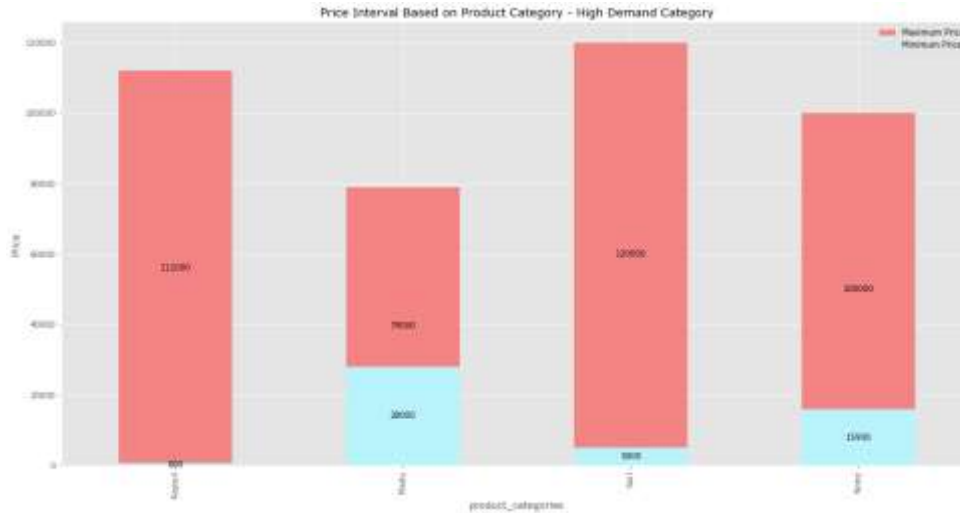
Figure XIII. Price Cluster 1 Interpretation

Figure XIII shows the price that has a high demand. Ignore for data "None", there is no category "None". For "Kapsul" have a price range of Rp600 - Rp112,000. The high range of price influenced by the volume of products sold. There are two types of sales of "Kapsul" products, namely per item and per gram. Products that are sold per item or per gram allow to gain high demand with a price range according to cluster 1. "Madu" with high demand have a price range of Rp28,000 - Rp79,000. Then, for "Sari", the price range is Rp5,000 Rp 120,000. The high range of price for "Sari" is also influenced by the volume of products sold. The company sells "Sari" product in Market with price Rp45,000 - Rp50,000 and volume of 200 ml (200 grams). Based on the superiority of the product composition compared to other competitors, it is still possible to increase the price until Rp120,000. That will be depend on marketing strategy to reach a target market that is willing to buy at a set price. Marketing strategy is the responsibility of marketing staff to create a good brand image for the product. If CV Nutri Pro wants to develop into other types of products, the company can set the price range on the market to get better margins based on market price which has "High demand" or in cluster 1.

3. Place

For place analysis, the most up sales are determined based on location for determining location which has high demand as shown in Figure XIV.
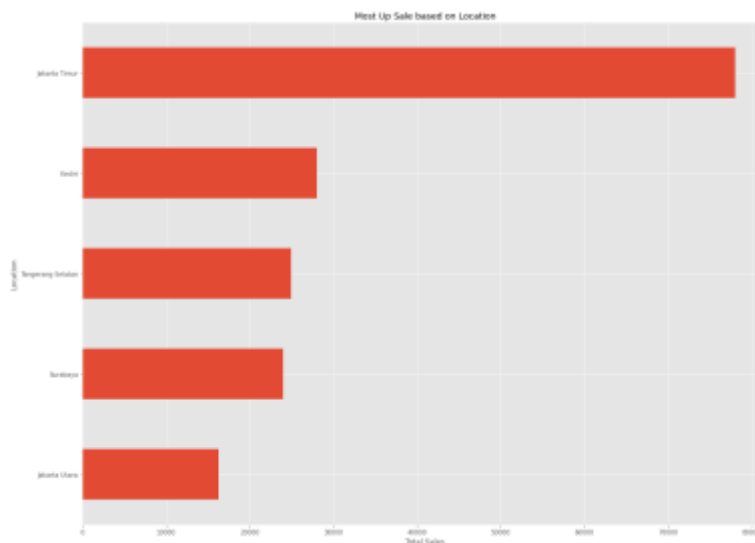

Figure XIV. Most Up Sales Based on Location

Figure XIV shows the highest sales by store location. For looking up the potential sales, the analysis also needs the number of stores from each location that have high sales or demand as shown in Figure XV.
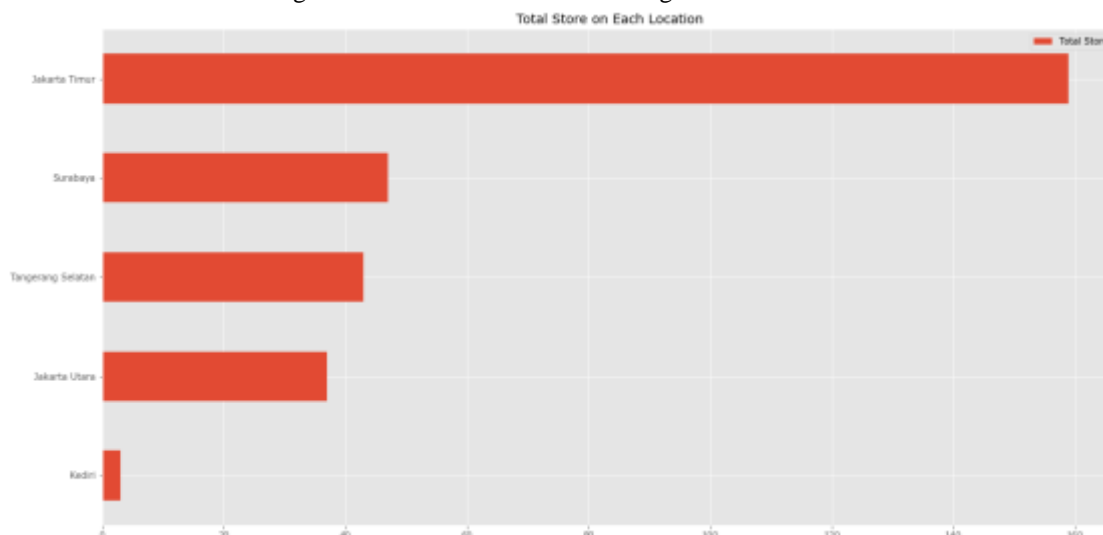

Figure XV. Total Store on Each Location

There are top five cities that will be considered for potential market expansion. The top five are "Jakarta Timur", "Kediri", "Tangerang Selatan", "Surabaya", dan "Jakarta Utara". The next step is to analyze the potential in each city to be able to determine sales expansion opportunities as shown on Table VII.

Table VII. Potential Market Expansion Analysis

No	City	Sold	Number of Sellers	Potential	Rank
1.	"Jakarta Timur"	78.064	159	490,96	4
2.	"Kediri"	28.000	3	9.333,33	1
3.	"Tangerang Selatan"	24.959	43	580,44	2
4.	"Surabaya"	23.991	47	510,45	3
5.	"Jakarta Utara"	16.279	37	439,97	5

Based on analysis, the priority locations have high demand but low supplier will interpret potential market expansion. All of the cities are in Java Island with the rank as follows: (1) Kediri, (2) Tangerang Selatan, (3) Surabaya, (4) Jakarta Timur, (5) Jakarta Utara.

4. Promotion

Table VIII. Discount Rate Segmentation

Cluster 0 : Low Demand			Cluster 1 : High Demand		
Code	Product Category	Discount Rate	Code	Product Category	Discount Rate
0	"Kapsul"	13%	0	"Kapsul"	43%
1	"Madu"	10%	1	"Madu"	33%
2	"Sari"	12%	2	"Sari"	45%
3	"None"	0%			

The existence of a higher discount cuffects the total sales. The data shows that the percentage of discounts still affects sales of health products. Therefore, it is necessary to determine the right price, it can help in determining the right promotion strategy as well based on discount rate.
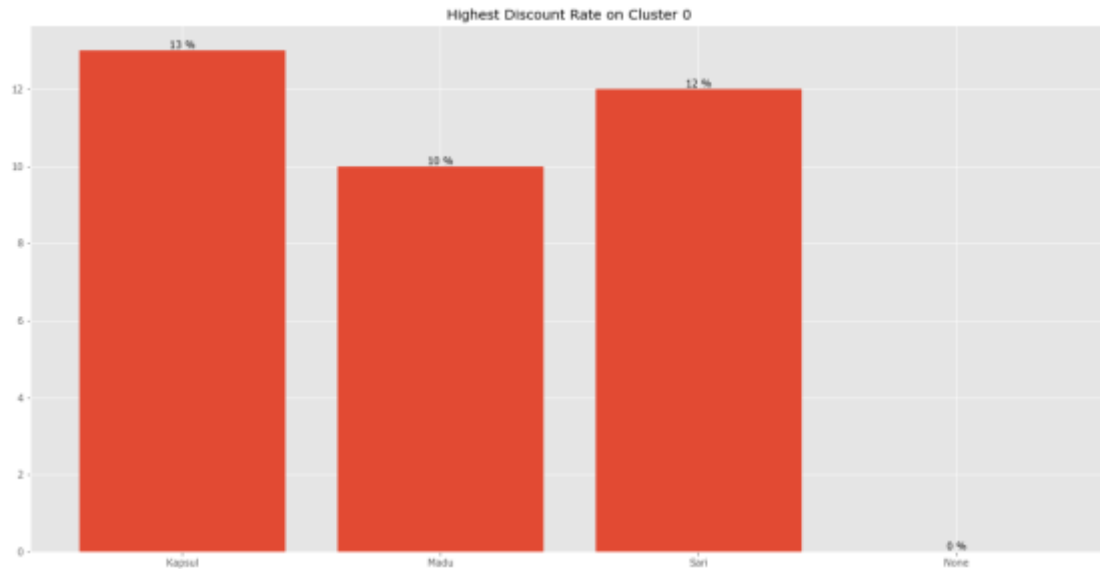
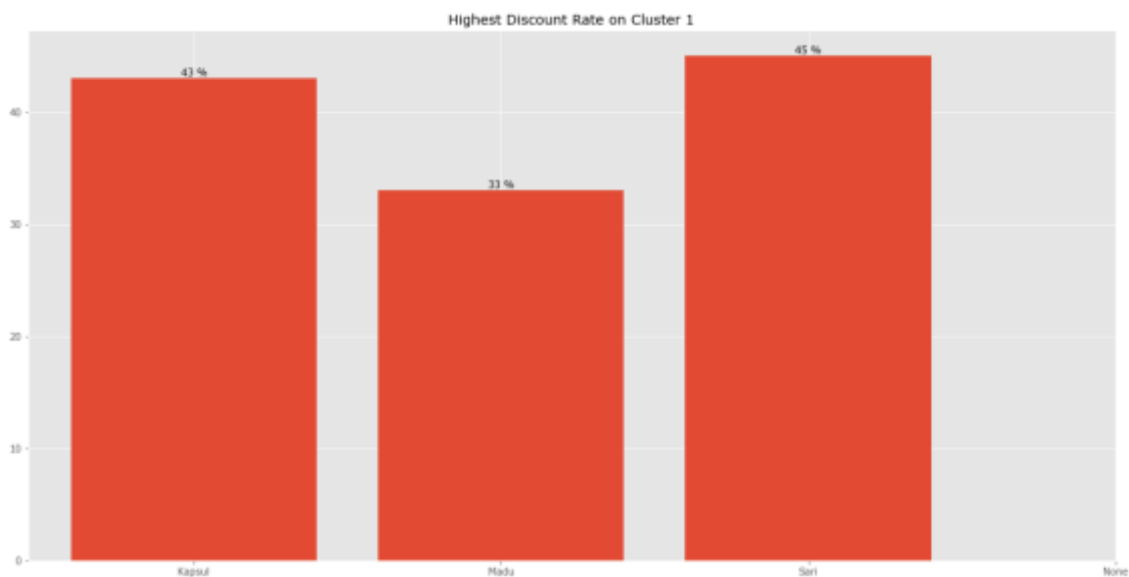**Figure XVI.** Discount Rate Cluster 0 Interpretation

Figure XVII. Discount Rate Cluster 1 Interpretation

If CV Nutri Pro wants to use special promotion on the certain day, the company can set the discount rate around 30% until 45% (in cluster 1) to attract the market demand. It uses as the sales promotion strategy.

BUSINESS SOLUTION

The interpretation results provide a solution in the form of a renewable marketing mix that can be applied by CV Nutri Pro as in Table IX.

Table IX. Renewable Marketing Mix

No	Marketing Mix	Renewal
1.	Product	Product development into several categories based on priority market demand with all having high product durability. The rank of priority is "Kapsul", "Madu", and "Sari". In addition, product development for existing product ("Sari"), the company can do quality improvement by increasing product durability.
2.	Price	Prices can be adjusted according to market prices that have high demand, not only focusing on pricing that covers the COGS. Determination of the marketing strategy is necessary also to be considered in order to the market can accept the price offered.
3.	Place	The top 5 recommendations for locations on Java Island open up opportunities for company to expand the partner or establish production site on this island. Decision to open new production site considering the company has not been able to reach East Java as the province of "Kediri" in same day. The establishing of new production site will also save logistics costs.
4.	Promotion	Selection of discounts at least above 30% until 45% for each product category in certain day or condition to attract purchase interest.

CONCLUSION

Big data that available on Tokopedia as the primary data regarding information of competitors. Primary data are extracted using the Web Scraping method to obtain several column categories as needed. The data that has been obtained then processed and analyzed to determine the renewable marketing mix strategy for the company. Data are processed and analyzed using pivot tables, K-Means and MDS. The processed data provide interpretations that are acceptable to CV Nutri Pro. The decision that the company will make is implementing a renewable marketing mix. Renewable marketing mix includes product, price, place, and promotion.

REFERENCES

1. Kwok L, Tang Y, Yu B. The 7 Ps Marketing Mix of Home-Sharing Services: Mining Travelers' Online Reviews on Airbnb. *Int J Hosp Manag*. 2020;90.
2. Dam K, Dinh L, Menvielle W. Marketing Intelligence from Data Mining Perspective — A Literature Review. *International Journal of Innovation, Management and Technology*. 2019;10(5):184-190.
3. Murugan P. Anti-Inflammatory Activity of Cinnamon (C. zeylanicum and C. cassia) Extracts - Identification of E-cinnamaldehyde and O-methoxy cinnamaldehyde as the Most Potent Bioactive Compounds. *International Journal of Current Research in Life Sciences*. 2021;10(1):3399-3402.
4. Bishop A, Witts S, Martin T. *The Role of Nutrition in Successful Wound Healing*. Vol 32.; 2018. www.nutrition.or.uk;
5. Soeters PB, Wolfe RR, Shenkin A. Hypoalbuminemia: Pathogenesis and Clinical Significance. *Journal of Parenteral and Enteral Nutrition*. 2019;43(2):181-193.
6. Eti IA, Bari M. *Theoretical Inspection and Rethink of Marketing Mix*.; 2020.
7. Dam. Marketing Intelligence from Data Mining Perspective — A Literature Review. *International Journal of Innovation, Management and Technology*. 2019;10(5):184-190.
8. Saini A, Rajesh A. Mapping 4Ps of Marketing with Business Intelligence: The Way to The Future of Marketing. *Central European Management Journal*. 2022;30(3):163-173.
9. Rabiyatou D, Babiga B, Sarr EN, Bousso M, Sall O, Mbaye SN. Web Scraping: State-of-the-Art and Areas of Application. *IEEE International Conference on Big Data (Big Data)*. 2019;17(6):6040-6042.
10. Singrodia V, Mitra A, Paul S. A Review on Web Scrapping and its Applications. In: *2019 International Conference on Computer Communication and Informatics, ICCCI 2019*. Institute of Electrical and Electronics Engineers Inc.; 2019.
11. Rotondo A, Quilligan F. Evolution Paths for Knowledge Discovery and Data Mining Process Models. *SN Comput Sci*. 2020;1(2).
12. Pekar A, Duque-Torres A, Seah WKG, Caicedo O. Knowledge Discovery: Can It Shed New Light on Threshold Definition for Heavy-Hitter Detection? *Journal of Network and Systems Management*. 2021;29(3).

13. Safhi HM, Frikh B, Ouhbi B. Assessing Reliability of Big Data Knowledge Discovery process. In: *Procedia Computer Science*. Vol 148. Elsevier B.V.; 2019:30-36.
14. Wickramasuriya R, Marchiori D. *Automated Detection of Business-Relevant Outliers in E-Commerce Conversion Rate.*; 2019.
15. Domino MA, Schrag D, Webinger M, Troy C. Linking Data Analytics to Real-World Business Issues: the Power of the Pivot Table. *Journal of Accounting Education*. 2021;57.
16. Dana RD, Soilihudin D, Silalahi RH, Kurnia DA, Hayati U. Competency Test Clustering Through the Application of Principal Component Analysis (PCA) and the K-Means Algorithm. *IOP Conf Ser Mater Sci Eng*. 2021;1088(1):012038.
17. Alfonsius A, Depari GS, Huang JP. Marketing Mix and Repurchase Intention of Cafe Industry During Covid-19: A Statistical and Data Mining Analysis. *Jurnal Minds: Manajemen Ide dan Inspirasi*. 2021;8(2):209.
18. Sun Z, Huo Y. The Spectrum of Big Data Analytics. *Journal of Computer Information Systems*. 2021;61(2):154-162.