

The global demand for beauty products continues to grow due to raised public awareness of applying cosmetics, with a 1.45 % to 3.34 % growth annually. Unfortunately, the COVID-19 outbreak broke out globally in December 2019, affecting face-to-face businesses such as the beauty industry falling until -7.11 % in 2020. This study aims to analyze the impact of the COVID-19 outbreak on Indonesia's beauty industry and the shift in the beauty consumer segment during the pandemic.

This study adopts the react-cope-adapt (RCA) framework to construct the COVID-19 pandemic periodization in Indonesia. The correlation analysis was used to investigate the impact of the COVID-19 pandemic on the beauty industry. In addition, clustering techniques were employed to identify hidden consumer segments and product preferences throughout the COVID-19 outbreak.

The study shows that COVID-19 cases positively impact beauty company's sales during the reacting phase. A strong negative relationship between COVID-19 and company revenue was observed in the coping phase. In the adapt phase, the negative impact of COVID-19 on the company's sales has decreased. Our finding also confirms the shift in consumer buying behavior during the pandemic. Consumers prefer to buy cosmetics products online than offline during the reaction phase. In the coping phase, consumers slowly begin to purchase in-store. Finally, consumers return to buying cosmetics offline in the adapting phase, similar to before the pandemic. The clustering results showed three hidden consumer segments: the loyal consumer segment, the impulsive consumer segment, and the compulsive consumer segment. In addition, during the pandemic, consumers prefer to buy skincare products over make-up products since government policies forced people to stay, work, and study at home.

Our study has theoretical and practical implications. Theoretically, our results support the usefulness of the RCA model and clustering techniques in analyzing the change in consumer buying behavior during a time of crisis, such as COVID-19 pandemic. Practically, beauty industries can anticipate this shift by accelerating the digital business transformation and focusing on the most preferred product to sustain their business.

Keywords: COVID-19 pandemic, consumer behavior, product preference, beauty industry, clustering technique

UDC 578

DOI: 10.15587/1729-4061.2023.274299

IMPLEMENTATION OF CLUSTERING TECHNIQUE FOR ANALYZING CONSUMER BUYING BEHAVIOR DURING THE COVID-19 PANDEMIC: A CASE IN THE BEAUTY INDUSTRY

Beny Maulana Achsan

Corresponding author

Master Student of Information Technology*

E-mail: beny.maulana@ui.ac.id

Achmad Nizar Hidayanto

Professor of Information Systems, Vice Dean for Resources, Venture, and General Administration*

*University of Indonesia

Salemba Raya str., 2, Jakarta, Indonesia, 10430

Received date 27.03.2023

Accepted date 29.05.2023

Published date 30.06.2023

How to Cite: Achsan, B. M., Hidayanto, A. N. (2023). Implementation of clustering technique for analyzing consumer buying behavior during the COVID-19 pandemic: a case in the beauty industry. *Eastern-European Journal of Enterprise Technologies*, 3 (2 (123)), 14–26. doi: <https://doi.org/10.15587/1729-4061.2023.274299>

1. Introduction

The global beauty market continues to grow due to increased public awareness of applying cosmetics, with a growth of 3.34 % in 2018 and 1.45 % in 2019. Nevertheless, the COVID-19 pandemic broke out globally in December 2019, causing the beauty market significantly fall to -7.11 % in 2020 [1]. Many regulations have been issued to handle the spread of the COVID-19 virus, such as lockdowns, social restrictions, and stay-at-home. These policies forced some offline stores to close at the pandemic's beginning due to safety and health concerns [2]. As a result, it negatively affected face-to-face businesses such as the beauty industry [3].

During the COVID-19 pandemic, the global beauty market slowly recovered, growing by 6.72 % in 2021 and 4.96 % in 2022. This growth aligns with increasing online channels contribution by 23.6 % in 2021 to 25.4 % in 2022 [1]. Consumers prefer buying cosmetics online than offline throughout the crisis. Store regulations such as sanitizing hands, wearing a mask, and social distancing influence consumers to purchase products offline. In addition, hygienic packag-

ing has become a crucial factor for consumer purchasing decisions during the pandemic [4]. Further, consumers considered skin care as an essential aspect throughout the pandemic, whereas the importance of make-up decreased [5].

The spread of COVID-19 itself, as well as government policies, caused a shift in consumer purchasing behavior. In the Czech Republic, online shopping increased from 12 % before the COVID-19 pandemic to 43 % during the COVID-19 pandemic [6]. Online purchasing also increased in the United States, Canada, and Germany, from 29.7 % before the pandemic to 48.3 % throughout the pandemic. At the end of the pandemic crisis, 42.2 % of people were reported to plan to continue online shopping [7].

As the world has not experienced a significant pandemic since the Spanish Flu in the early 20th century, few studies on consumer behavior exist in this context as references. At the same time, it is more complicated because the COVID-19 pandemic might be over for a long time due to the SARS-CoV-2 virus still evolving [8]. It adds business and economic problems, such as unemployment and poverty [9]. In contrast, understanding consumer behavior and managing

their relationship with enterprises is essential to increase consumer purchase intention, loyalty, and participation [10], ultimately boosting company revenue. Therefore, a study of the impact of the COVID-19 pandemic on the beauty industry and how consumer buying behaviors shifted during the pandemic in the context of the beauty market's growth is essential for academics, business people, and public policymakers.

2. Literature review and problem statement

The study [11] identifies offline consumer buying behavior in the beauty industry. This study reported that consumers' moods, personal situations, store layout, product promotion, and product attributes positively impact impulsive buying. However, this research only uses surveys and is descriptive in nature, which may lead to bias. In addition, external factor such as the COVID-19 pandemic that spread over the world is not considered in detail. These limitations can be overcome by complementing online interviews and considering external factors for a more robust analysis.

The paper [12] examines changes in consumer behavior during the outbreak of COVID-19. The authors prove that consumer behavior is affected by the mass infection of COVID-19 and public policy implementation. Factors such as the resident's capability to cope with the policies and adherence to quarantine rules also influence consumer behavior. However, this study period has a temporal range of six months, which does not cover the possibility of consumer behavior shifts at the pandemic's end. Another paper [13] employed descriptive analysis and discrete choice modeling to identify trends in online shopping during and after the COVID-19 pandemic and to analyze shifting patterns in online and offline shopping behaviors. However, how pre-COVID, during COVID, and post-COVID pandemic periods are divided in this work is unclear.

The paper [14] explores how consumer behavior changed throughout the COVID-19 pandemic. The authors adopt the react, cope, and adapt (RCA) framework to identify consumer behavior at the pandemic's beginning. However, there were unresolved issues related to how the pandemic phase was divided over time. In addition, how consumers behave in the cope and adapt phases needs to be explained in detail. It might be because time was too short to observe consumer behavior during the early days of the pandemic. The way to overcome these difficulties is to extend the research observation period. This approach was used by [15] to identify how online purchasing behavior evolved during the COVID-19 crisis. It shows that during the reaction phase, consumers start reacting by expanding the average basket size. In the cope phase, consumers start focusing on themselves to take their minds away from the COVID-19 problem by increasing purchasing of skincare categories, such as hair care, body care, and face care. Finally, consumers would be less reactive by reducing purchasing protective products throughout the adapting phase. However, this study is descriptive in nature, and the period has been too short to identify how consumers adapt during the pandemic. In addition, consumer segmentation during the crisis was not observed.

Recent studies show that unsupervised learning is more attractive in analyzing consumer behavior. It can identify hidden patterns in a group based on similarity. The study [16] identifies potential customers segment using

k-Means clustering based on recent, frequency, and monetary (RFM) models. Nevertheless, this study did not explain product segmentation analysis, such as the most preferred products. In addition, determining the number of clusters is limited to two options, $k=3$ and $k=5$, which can lead to the potential loss of the most optimal number of groups.

The research [17] used unsupervised learning to analyze customer behavior by segmenting customers based on product sales categories. The authors combine the classic k-Means algorithm with affinity propagation algorithms (AP+k-Means) and semi-supervised affinity propagation algorithms (SAPK+k-Means). It shows that traditional k-Means has the fastest execution time but has the lowest accuracy comparing the two other algorithms. However, this research has an unresolved issue related to the volatility effect, making the classification difficult. A standardization or smoothing process is necessary to overcome this difficulty [18].

The paper [19] used unsupervised learning and adopted Stern's buying theory principles to identify consumer segments. This research confirms the existence of planned, unplanned, and impulsive behavior when purchasing organic products. However, this study did not consider external factors such as the COVID-19 pandemic, which began to spread then. In addition, this study has yet to provide real market transactions to support a more realistic and robust analysis.

Another paper [18] identifies customer-product segmentation in the retail industry using k-Means and Louvain algorithms. It is shown that the Louvain algorithm has a better response rate than the k-Means algorithm in determining customer-product segments. Louvain algorithm is explicitly interpreted in this paper to fulfill the need for transparency of the algorithm as it generates clusters similar to classic k-Means clustering methods. However, this study has unresolved issues due to the unlabeled customer segment for further analysis.

The work [20] adopted RFM variables and used time-series clustering to identify consumer segments in the retail industry. The result shows that Hierarchical with Complexity-Invariant Distance evaluation performs better than Spectral and k-Shape clustering techniques. However, this study has limitations due to the initial number of clusters set to three options, $k=4$, $k=5$, and $k=6$, which could potentially lose the optimal number of formed groups. In addition, the utilization of the dataset for eleven months needs to be longer to gain deeper insight into the change in consumer behavior over time.

The main issue from the literature above is that exploratory and descriptive analysis still dominates in identifying consumer buying behavior during crises such COVID-19 pandemic. Although some previous studies show similar results, few analyzed the shift in consumer buying behavior during the COVID-19 outbreak, particularly in Indonesia's beauty industry. Unsupervised learning, such as clustering techniques, still needs to be improved in its application to identify consumer and product segmentation. At the same time, COVID-19 infection is still occurring globally and affecting business operations and consumer activities. In addition, it adds other problems, such as unemployment and poverty. It is essential that research containing specific recommendations will help management to decide on the right business strategy in the future. Therefore, studying the impact of COVID-19 on the beauty industry and how the buying behavior of beauty consumers shifted during

the COVID-19 pandemic using clustering techniques is an urgent task that requires continuous improvement.

3. The aim and objectives of the study

The study aims to analyze the impact of the COVID-19 pandemic on Indonesia's beauty industry and the shift in the beauty consumer segment using clustering techniques. This will make it possible to identify potential consumer clusters and product preferences shifted more accurately, enabling the beauty industry to anticipate this change with the right strategies for sustaining their business throughout the crisis.

To achieve this aim, the following objectives are accomplished:

- to determine the impact of COVID-19 cases on Indonesia's beauty industry during the react, cope, and adapt phases;
- to identify beauty consumer segments and product preferences throughout the pandemic phases.

4. The study materials and methods

The object of this study is one of Indonesia's largest beauty companies. The main hypothesis is that the COVID-19 pandemic impacts consumer behavior in purchasing beauty products. The proposed methodology is illustrated in Fig. 1 below.

Several scientific methods are generally applied when conducting research and confirming assumptions. This case study employs descriptive analysis and adopts the RCA model to construct the COVID-19 pandemic periodization in Indonesia based on the timeline of COVID-19 events. This pandemic phase was then verified using the national online transaction volume from the Bank of Indonesia (bi.go.id). PowerBI v.2.115.842.0 was used to visualize the result of the pandemic phases.

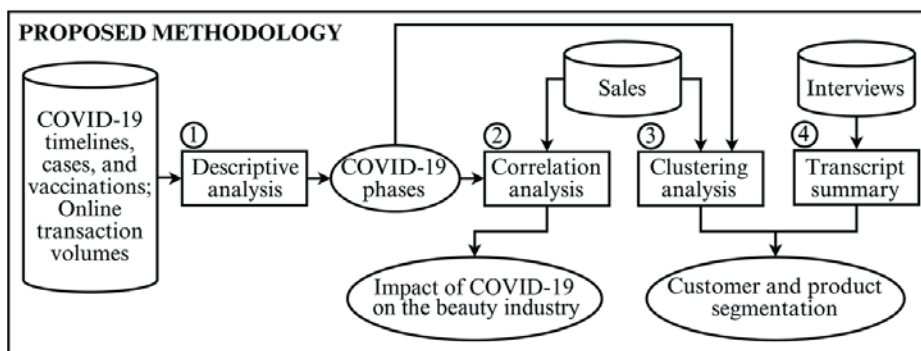


Fig. 1. Proposed methodology

Correlation analyses are conducted to investigate the effect of COVID-19 cases on the beauty industry sales. The authors consider Pearson's correlation coefficient (R) smaller than -0.4 and greater than $+0.4$ as a strong correlation, while those between $(-0.4$ and $-0.2)$ and $(0.2$ and $0.4)$ as a moderate correlation [19]. Python v.3.9.16 was used to calculate Pearson's correlation coefficient. Mathematically, Pearson's correlation coefficient can be calculated using the following eq. (1) [21]:

$$R(x_i, y_i) = \frac{\text{cov}(x_i, y_i)}{\sqrt{\text{var}(x_i) \cdot \text{var}(y_i)}}, \quad (1)$$

where x_i and y_i represent the dataset, $\text{cov}(x_i, y_i)$ and $\text{var}(\bullet)$ represent covariance and variance.

Clustering techniques, one of the most popular unsupervised learning [22], were used to identify hidden consumer and product clusters during the COVID-19 outbreak. A cluster is a group of objects with similar characteristics [16], while clustering is a technique for finding the optimal number of clusters based on similarity within the same cluster and the dissimilarity between objects in different clusters [23]. The k-Means algorithm was chosen in this study to determine the number of consumer segments during the pandemic crisis. Compared to other algorithms, the k-Means algorithm gives faster computation and works well on large data sets. In addition, the k-Means algorithm only requires one input parameter, k , compared to other methods to reduce the rate of data classification errors [16]. On the other hand, the Louvain algorithm is used in this study as a comparison method because it also gives fast computation even for large graphs consisting of millions of consumers and products [18]. Orange v.3.34 was used in this study for data modeling.

Silhouette analysis was carried out in this study to evaluate the number of clusters by calculating the similarity distance between objects. There are many ways to measure the similarity distance, such as the Minkowski and Mahalanobis Distance. However, the Euclidean Distance is the most popular and widely used in the literature [24, 25]. This research used Euclidean Distance to measure the similarity of the cluster member, which can be calculated using eq. (2) below [16, 24]:

$$d(i, j) = \sqrt{(i_1 - j_1)^2 + (i_2 - j_2)^2 + \dots + (i_n - j_n)^2}, \quad (2)$$

where $d(i, j)$ is the distance between objects i and j , and n is the total object in the cluster. Furthermore, the silhouette score value can be calculated using eq. (3) below [16]. The silhouette score ranges $[-1, 1]$, where high values imply that the object fits well into its group and poorly fits into other groups [23]:

$$s(i) = \begin{cases} 1 - \frac{a_i}{b_i}, & a_i < b_i, \\ 0, & a_i = b_i, \\ \frac{a_i}{b_i} - 1, & a_i > b_i, \end{cases} \quad (3)$$

where $a(i)$ represents cohesion and $b(i)$ represents separation as the following equations:

$$a(i) = \frac{1}{|c_i| - 1} \sum_{j \in c_i, i \neq j} d(i, j), \quad (4)$$

$$b(i) = \min_k \frac{1}{|c_k|} \sum_{j \in c_k} d(i, j). \quad (5)$$

Furthermore, the quality of the clusters was evaluated by calculating the average silhouette score as eq. (6) below [26]:

$$\hat{s} = \frac{1}{n} \sum_{i=1}^n s_i, \quad (6)$$

where \hat{s} is the average silhouette score, and n is the number of clusters. The process of data modeling is illustrated in Fig. 2 below.

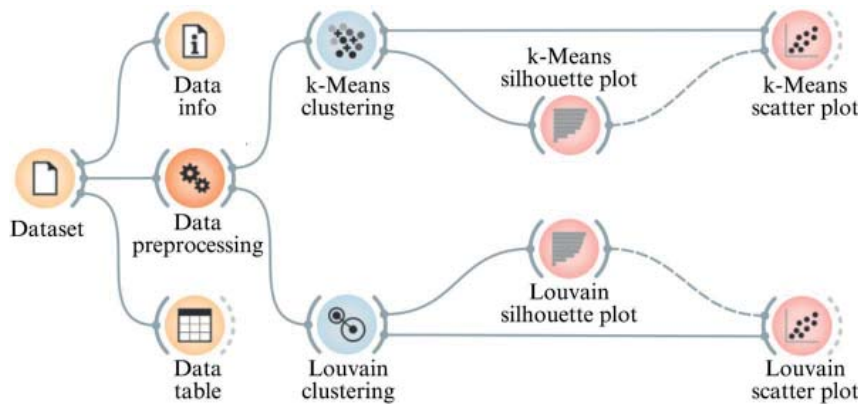


Fig. 2. Data modeling using Orange

Average basket sizes and COVID-19 cases were used in this research as the features for segmenting consumers during the pandemic. The average basket size data is gathered from March 2020 to June 2022 with a total of 28 months, while the monthly COVID-19 cases data is obtained from Indonesia’s COVID-19 portal (covid19.go.id). Data preprocessing consists of normalization $[-1, 1]$ to measure continuous variable differences. K-Means and Louvain algorithms are conducted to define the number of clusters (k). The k-Means algorithm determined the clusters by examining twenty-six iterations from $k=2$ to $k=27$. Meanwhile, the number of clusters in the Louvain algorithm is determined by varying the number of k-neighbors from 2 to 27. The number of clusters with the highest silhouette score was selected. In addition, product segmentation is determined using sales growth and sales average features of eight product categories, such as body care (BC), face care (FC), deodorant (DE), fragrance (FR), hair care (HA), hand care (HN), make-up (MU), and others categories (OT) to identify the change in consumer buying preference during each pandemic phase. In addition, online interviews were conducted with the management and executive to give another perspective on the shift in consumer buying behavior during the crisis. Finally, the method of theoretical generalization was used to write the conclusions of this study.

5. Research results of consumer buying behavior during the COVID-19 pandemic

5.1. The impact of the COVID-19 pandemic on Indonesia’s beauty industry during the react, cope, and adapt phases

The COVID-19 pandemic in Indonesia is inseparable from a pneumonia case in Wuhan, China, on December 31, 2019 [27]. With these findings, the Indonesian government immediately evacuated 243 citizens from Wuhan to the Natuna Islands [28]. A month after the evacuation,

President Joko Widodo announced the first infection of COVID-19 in Indonesia on March 2, 2020 [29], as illustrated in Fig. 3, and stated that the COVID-19 pandemic is a national disaster on April 13, 2020 [30, 31]. Some regulations were imposed by the government to minimize the spread of COVID-19, such as suspending incoming foreign visitors from COVID-19-confirmed countries [32] and enforcing large-scale social restrictions (PSBB) [1].

Good news has emerged with the presence of the COVID-19 vaccine in Indonesia. The first dose was received by President Joko Widodo on January 13, 2021 (Fig. 3), as a step toward mass

vaccination in Indonesia [33]. Even though the mass vaccine was already started, the first wave of COVID-19 broke out in Indonesia 12 days later with 89,902 new cases in a day, an increase of 284 % or almost four times in 13 weeks. The second wave happened on June 21, 2021, with 125,396 cases in one day, an increase of 381 % or almost five times in 6 weeks. As such, the government enforced strict restrictions on community activities (PPKM) from 3–31 July 2021 [34].

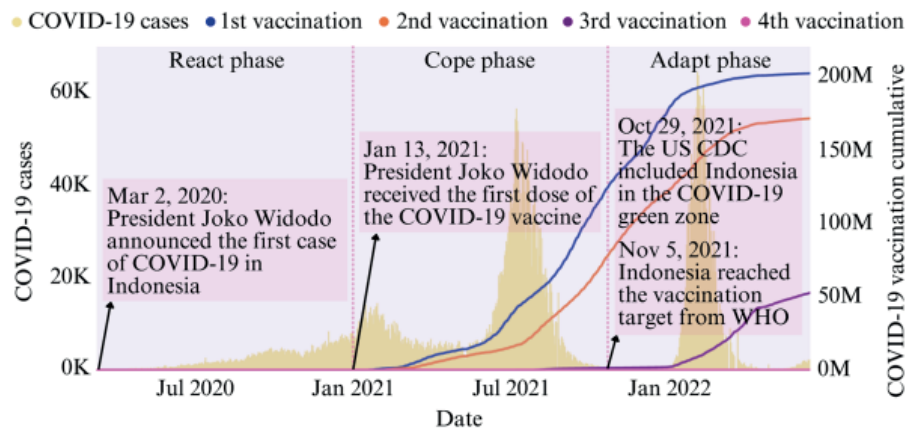


Fig. 3. COVID-19 phase in Indonesia

Finally, US Centers for Disease Control and Prevention declared Indonesia in the green zone for COVID-19 and is safe to visit on October 29, 2021 [35]. This is also supported by the achievement of mass vaccination in Indonesia, which can exceed the WHO target of 200 million doses on November 5, 2021 [36]. Due to the increasing number of mass vaccinations, President Joko Widodo allowed Indonesians to take off their masks in public areas on May 17, 2022 [37], after the third wave of COVID-19 was declining.

Based on the events related to COVID-19 in Indonesia above, we proposed the periodization of the COVID-19 pandemic as follows:

- 1) the react phase was marked by President Joko Widodo’s first announcement of COVID-19 cases in Indonesia;
- 2) the cope phase was signed by the first COVID-19 vaccine injected into President Joko Widodo to boost the body’s immunity;
- 3) the adapt phase was marked by Indonesia’s entry into the green zone of COVID-19 and is safe for visits by other countries. Furthermore, the government can achieve

the WHO’s vaccination target with more than 200 million doses. The periodizations of the COVID-19 pandemic in Indonesia are as follows:

- react phase: March 2020 – December 2020;
- cope phase: January 2021 – October 2021;
- adapt phase: November 2021 – June 2022.

These pandemic phases were then verified by observing the trend of online transaction volume in Indonesia before and during the COVID-19 pandemic, as illustrated in Fig. 4 below.

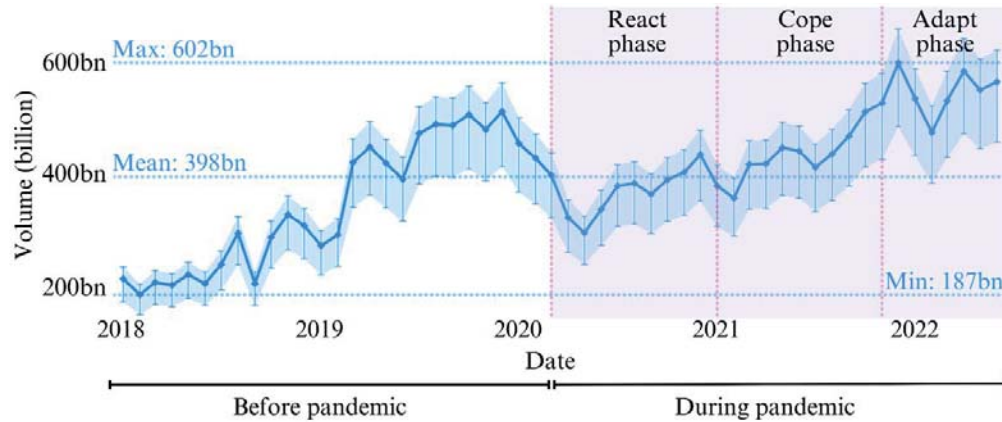


Fig. 4. Online transaction volume in Indonesia from Q1 2018 to Q2 2022

Before the pandemic, online transaction volume in Indonesia had a positive trend, with an average of 378 billion monthly transactions. This trend decreased when entering the first pandemic phase. Online transactions in Indonesia evolved during the COVID-19 pandemic. In the react phase, online transactions gradually increased until the phase ended in December 2020. In this phase, consumers started to change their behavior in buying products through online channels. It was observed by an increase in online transactions a few months after COVID-19 spread in Indonesia, which previously experienced a sharp decline due to the COVID-19 issue. During the cope phase, the online transaction trend increased by 15.5 % compared to the react phase. It shows that consumers continue to buy products or services online during this phase. In the adapt phase, the trend of online transactions begins to stabilize, with an average of 549 billion transactions per month. It shows that consumers have started to adopt online purchasing. In this phase, online transactions experienced the highest record in the last four and a half years, with online transaction volume reaching 602 billion per month. Descriptive statistics of online transaction volume in Indonesia are summarized in Table 1.

Table 1
Descriptive statistics of online transaction volume in Indonesia (billion)

Period	Min	Max	Mean
Before pandemic	187	515	378
During pandemic	298	602	429
React phase	298	438	374
Cope phase	360	514	432
Adapt phase	478	602	549

Based on Table 1 above, online transaction volume during the COVID-19 pandemic increased by 13.5 % compared to before the pandemic. This indicates that consumers have started to adopt online shopping during the pandemic.

During the pandemic, the company sales decreased by –13.6 % compared to before the pandemic. Offline sales tend to fall when COVID-19 cases hit the wave’s peak, while online sales are stable during the period. The channel contribution and growth against the last period can be noticed in Table 2 below.

Based on Table 2 above, the contribution of online sales increased by 9.3 % during the pandemic, in line with the interview result with the Head of Business Innovation and Development (HBID) as follows:

“The trend of online sales in Indonesia was increasing before the pandemic. The company started opening market-places in 2018 with positive growth. This growth has increased rapidly during the COVID-19 outbreak.” HBID, 15/10/22.

Table 2

Channel Contribution and Growth

Period	Online contribution	Offline contribution	Online growth vs. last period	Offline growth vs. last period
Before pandemic	3.3 %	96.7 %	–	–
During pandemic	12.6 %	87.4 %	–	–
React phase	11.7 %	88.3 %	178.8 %	–27.4 %
Cope phase	15.5 %	86.5 %	18.1 %	3.4 %
Adapt phase	12.5 %	87.5 %	7.7 %	18.4 %

During the react phase, the sales growth on the online channel increased significantly, up to 178.8 %. This growth declined during the cope and adapted phases. In contrast, the contribution of offline channels decreased by –27.4 %. This growth slowly recovered during the cope and adapted phases.

Correlation analysis was conducted to measure the degree of association between COVID-19 and company sales. The correlation between COVID-19 and company sales was observed during the react, cope, and adapt phases, as shown in Table 3 below.

During the reacting phase, COVID-19 cases positively correlate to the beauty company’s sales on online and offline channels. A strong relationship between COVID-19 and sales of categories such as deodorant, face care, and fragrance was observed. In this phase, consumers prefer to buy make-up and hair care online. It is caused by the fear of being infected with COVID-19 when purchasing offline.

Table 3
Correlation between COVID-19 and company sales during pandemic phases (underlined values represent strong correlations)

Sales	React phase (R)	Cope phase (R)	Adapt phase (R)
Total	<u>0.628</u>	<u>-0.480</u>	-0.339
Online channel	<u>0.592</u>	0.056	<u>-0.461</u>
Body care	0.076	0.060	<u>-0.593</u>
Deodorant	<u>0.433</u>	0.190	<u>-0.569</u>
Face care	<u>0.705</u>	0.113	-0.384
Fragrance	<u>0.585</u>	-0.150	-0.338
Hair care	<u>0.520</u>	<u>0.573</u>	-0.361
Hand care	0.158	0.047	0.033
Make-up	<u>0.506</u>	-0.052	<u>-0.520</u>
Others	<u>0.511</u>	-0.075	-0.355
Offline channel	<u>0.503</u>	<u>-0.561</u>	-0.299
Body care	<u>-0.700</u>	<u>-0.482</u>	-0.364
Deodorant	<u>0.476</u>	-0.165	<u>-0.428</u>
Face care	<u>0.756</u>	<u>-0.451</u>	-0.060
Fragrance	<u>0.464</u>	<u>-0.464</u>	<u>-0.755</u>
Hair care	0.140	<u>-0.431</u>	-0.045
Hand care	0.158	0.047	0.033
Make-up	0.275	<u>-0.500</u>	<u>-0.556</u>
Others	0.061	-0.235	-0.351

A strong negative relationship between COVID-19 and sales on offline channels was observed in the coping phase. In this phase, consumers begin to abandon offline purchases. This is driven by government policies such as PPKM, which instructs people to stay at home. In addition, the closure of shopping centers causes consumers to be reluctant to shop in-store. In contrast, COVID-19 did not affect sales on online channels. Hair care sales on offline channels still have a positive correlation with COVID-19. On the other hand, sales of body care, face care, fragrance, hair care, and make-up are starting to feel a negative impact due to COVID-19.

In the adaptation phase, COVID-19 negatively affects sales in online channels. In contrast, the impact of COVID-19 on offline sales has started to decline. It shows that consumers are returning to buying products in-store. In this phase, the company’s sales are recovering, aligned with the decline in COVID-19 infections.

5.2. Beauty consumer segments and product preferences during the COVID-19 pandemic

Here, we identify the consumer segments and product preferences throughout the COVID-19 outbreak. K-Means and Louvain algorithms are conducted to determine the optimal number of clusters. The optimal consumer segment from each algorithm is illustrated in Fig. 5, 6 below.

In online channels (Fig. 5), the k-Means algorithm identified three hidden consumer segments. Cluster one (C1) represents the consumer segment not affected by COVID-19 in purchasing products, where consumers tend to be stable in online purchases despite the high number of COVID-19 cases. Cluster two (C2) is a consumer segment that purchases in large basket sizes when COVID-19 cases are low. Meanwhile, cluster three (C3) mainly represents the consumer segment that buys in small basket sizes when COVID-19 cases are low. In contrast, Louvain’s algorithm produces two consumer segments where cluster one (C1) is a union of cluster one and two

in the k-Means algorithm, while cluster two (C2) is similar to cluster three from the k-Means algorithm.

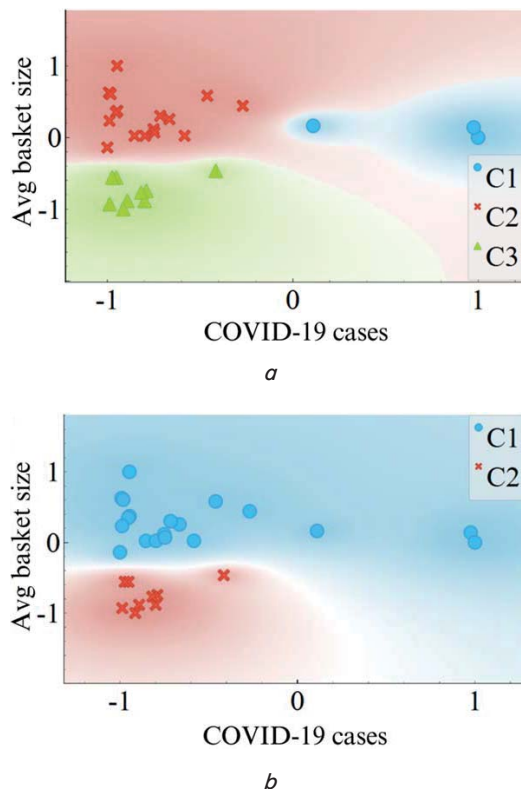


Fig. 5. Online consumer segmentation using: a – the k-Means algorithm; b – the Louvain algorithm

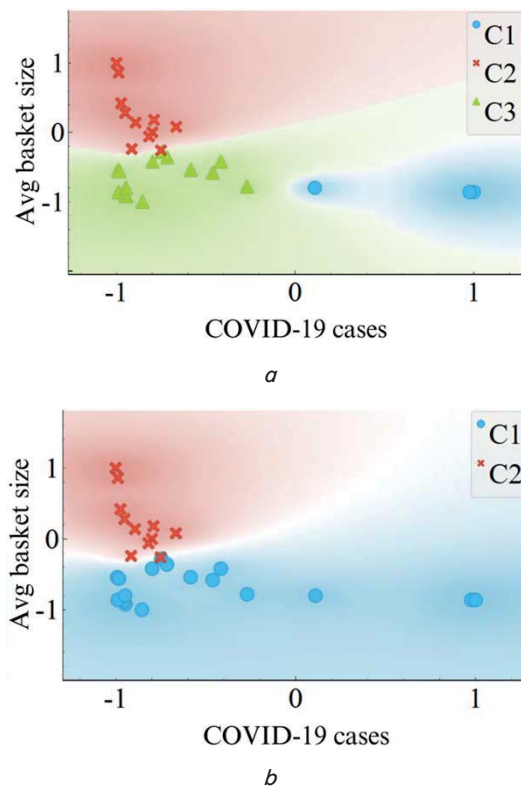


Fig. 6. Offline consumer segmentation using: a – the k-Means algorithm; b – the Louvain algorithm

In offline channels (Fig. 6), the k-Means algorithm produces three consumer segments. Cluster one (C1) is a consumer segment that tends to continue buying products even though COVID-19 cases are high by reducing the basket size. Cluster two (C2) is a consumer segment that purchases offline in large basket sizes only when COVID-19 cases are low. Meanwhile, cluster three (C3) is a consumer segment that continues to make offline purchases in small basket sizes when COVID-19 cases are medium to low. In contrast, the Louvain algorithm identified two consumer segments. Cluster one (C1) is a combination of cluster one and cluster three in the k-Means algorithm. Meanwhile, cluster two (C2) is similar to cluster two of the k-Means algorithm. The evaluation of cluster formation is presented in Table 4 below.

Table 4

Consumer cluster evaluation

Algorithm	Channel	Number of clusters (<i>k</i>)	Silhouette score (<i>s</i>)	Avg silhouette score (\hat{s})
k-Means	Online	3	0.593	0.530
	Offline	3	0.467	
Louvain	Online	2	0.521	0.479
	Offline	2	0.436	

Based on the comparison of the silhouette score above (Table 4), the k-Means algorithm ($\hat{s}=0.530$) outperforms the Louvain algorithm ($\hat{s}=0.479$) in determining consumer segments in both online and offline channels. The higher the silhouette score, the more compact consumers within a cluster and the more separated consumers to other

clusters. In addition, changes in consumer buying behavior were also observed by segmenting the product category during the pandemic, as illustrated by Fig. 7, 8 (online channel) and Fig. 9, 10 (offline channel) below.

In the react phase (Fig. 7), the k-Means algorithm produces three clusters of products on the online channel. Cluster one (C1) consists of the hair care, deodorant, fragrance, hand care, and others categories; cluster two (C2) only consists of the make-up category and cluster three (C3) consists of body care and face care categories. On the other hand (Fig. 8), Louvain’s algorithm groups fragrance, hand care, and other categories into one cluster; cluster two consists of body care, hair care, and deodorant; and the last cluster consists of face care and make-up categories.

In the cope phase (Fig. 7, 8), the k-Means algorithm produces more clusters (4 clusters) than the Louvain algorithm (3 clusters). The k-Means algorithm concludes that the others category differs from the fragrance, deodorant, and hair care categories. In contrast, the Louvain algorithm considers those four categories to have high similarities.

In the adapt phase (Fig. 7, 8), the k-Means and Louvain algorithms construct the same number of product segments as three clusters. However, there is a difference between them where the k-Means algorithm concludes that the deodorant category is more similar to the body care, hair care, and others categories. In contrast, the Louvain algorithm concludes that the deodorant category is closer to fragrance and hand care. The number of product clusters in the online channel with the corresponding average silhouette score is shown in Table 5 below.

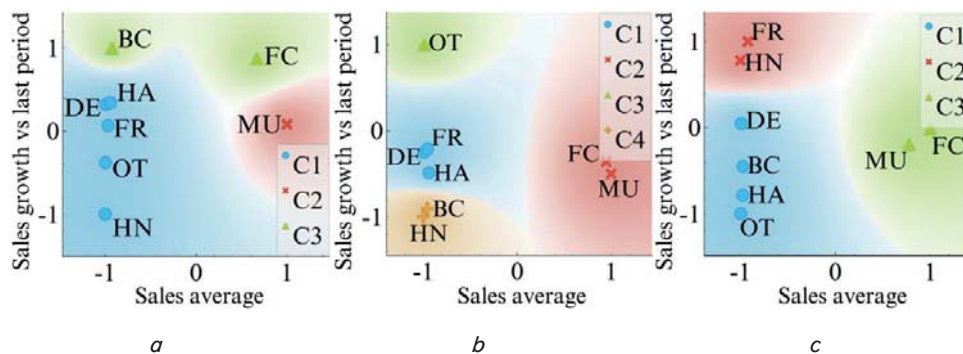


Fig. 7. Product segmentation in the online channel using the k-Means algorithm: a – react phase; b – cope phase; c – adapt phase

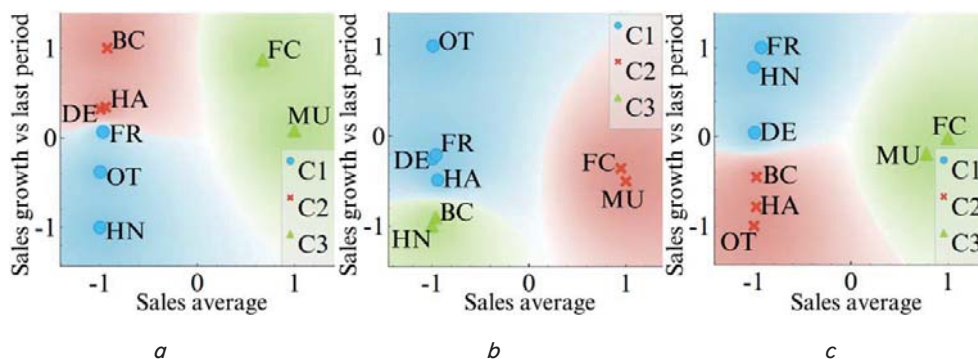


Fig. 8. Product segmentation in the online channel using the Louvain algorithm: a – react phase; b – cope phase; c – adapt phase

Table 5

Online product cluster evaluation

Algorithm	Phase	Number of clusters (k)	Silhouette score (s)	Avg silhouette score (\hat{s})
k-Means	React	3	0.310	0.544
	Cope	4	0.675	
	Adapt	3	0.647	
Louvain	React	3	0.281	0.530
	Cope	3	0.647	
	Adapt	3	0.660	

According to the cluster evaluation above (Table 5), the k-Means algorithm ($\hat{s}=0.544$) outperforms the Louvain algorithm ($\hat{s}=0.530$) in identifying hidden product segments in online channels during the COVID-19 periodization. Further, those algorithms were then used to determine hidden product patterns in offline channels, as illustrated in Fig. 9, 10 below.

In the react phase (Fig. 9, 10) k-Means and Louvain algorithms construct the same three clusters of products. Cluster one (C1) consists of hair care, deodorant, fragrance, and other categories; cluster two (C2) consists of face care and make-up categories, and cluster three (C3) consists of hand care and body care categories.

In the cope phase (Fig. 9, 10) k-Means and Louvain algorithms produce the same number of product segments as three clusters. However, there is a difference between

them where the k-Means algorithm states that others category is dissimilar with fragrance, deodorant, hand care, hair care, and body care. Meanwhile, the Louvain algorithm separates body care as a different cluster from fragrance, deodorant, hand care, hair care, and others category.

In the adapt phase (Fig. 9, 10), the k-Means and Louvain algorithms also result in the same number of product clusters as three clusters. However, there is a difference between them where the k-Means algorithm concludes that the body care category is more similar to the deodorant, fragrance, and hair care categories. In contrast, the Louvain algorithm concludes that the body care category is similar to hand care and other categories. Table 6 summarizes the number of product clusters in the offline channel with the corresponding average silhouette score.

Table 6

Offline product cluster evaluation

Algorithm	Phase	Number of clusters (k)	Silhouette score (s)	Avg silhouette score (\hat{s})
k-Means	React	3	0.567	0.554
	Cope	3	0.612	
	Adapt	3	0.484	
Louvain	React	3	0.596	0.474
	Cope	3	0.354	
	Adapt	3	0.472	

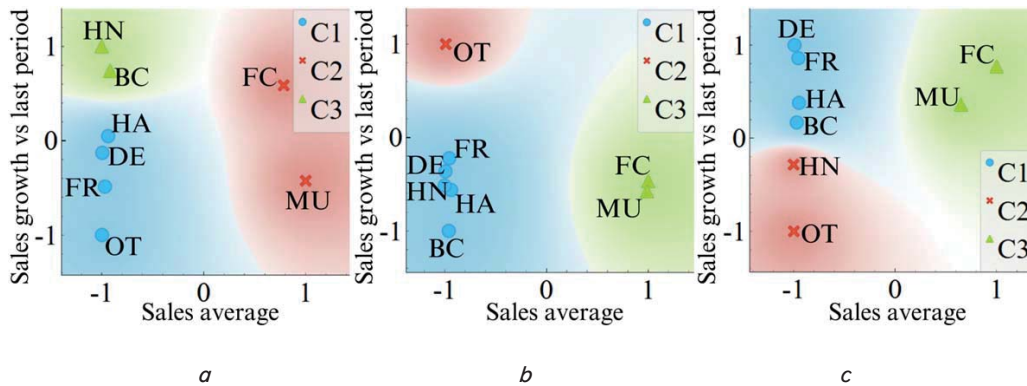


Fig. 9. Product segmentation in the offline channel using the k-Means algorithm: a – react phase; b – cope phase; c – adapt phase

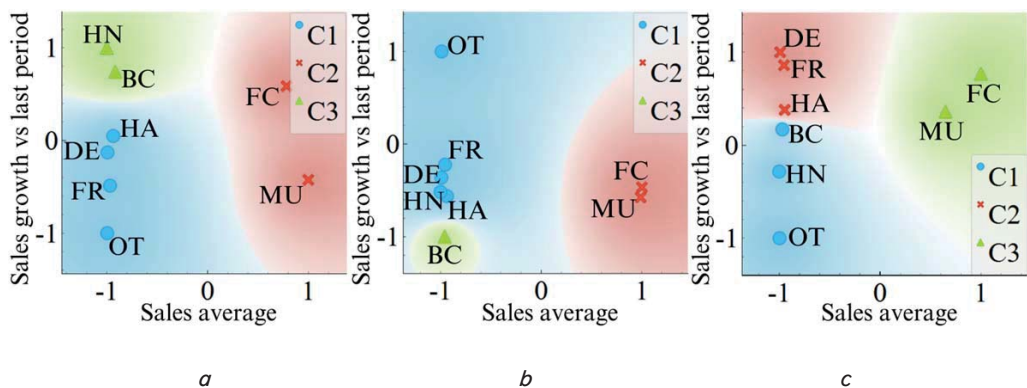


Fig. 10. Product segmentation in the offline channel using the Louvain algorithm: a – react phase; b – cope phase; c – adapt phase

Comparing the average silhouette score (Table 6) above, the k-Means algorithm ($\hat{s}=0.554$) outperforms the Louvain algorithm ($\hat{s}=0.474$) in identifying hidden product clusters in offline channels during the pandemic. As a result, the product clustering using k-Means is more compact within a cluster and more separated from other clusters to Louvain's algorithms.

6. Discussion of the results of the study on the impact of the COVID-19 pandemic on beauty consumer behavior and product preference

The result of the case study supports the RCA model's usefulness in dividing Indonesia's COVID-19 pandemic periodization into react, cope, and adapt phases. The COVID-19 outbreak, as well as its policies, impact the growth of the beauty industry. During the pandemic, sales growth decreased by -13.6% compared to before the pandemic. The COVID-19 cases positively impact the company's revenue during the reacting phase. A strong negative relationship between COVID-19 and company sales was observed in the coping phase. In the adapt phase, the negative impact of COVID-19 on company sales has reduced.

Our finding also confirms the change in consumer buying behavior during the pandemic, where consumers prefer buying cosmetics products online than offline due to the restriction regulations. The shift in this buying behavior is supported by the online channel contribution, which increased by 9.3% during the crisis (Table 2). This increase is in line with the contribution of online transactions in Indonesia, which grew by 13.5% throughout the pandemic (Table 3). It also aligned with the interview undertaken with the Executive of Financial Planning and Analysis (EFPA) as follows:

"Before the pandemic, consumers prefer to buy cosmetics directly at stores than buying them in e-commerce. However, when COVID-19 spread, people started experimenting with online shopping due to the PPKM policy. In addition, the discount offers on the marketplace encourage the consumer to shift from offline to online purchasing." EFPA, 2/12/22.

Consumer shopping preferences shifted from offline to online during the reaction phase. This change is due to the threat of being infected with COVID-19 when purchasing in-store, in line with [15]. In the cope phase, consumers slowly return to buying in-store. This can be observed by increasing the contribution of offline sales up to 3.4% (Table 2). This finding implies that consumers have started to return to activities outside due to easing health protocols and successfully implementing mass vaccinations. In the adapt phase, consumers return to buying cosmetics offline, similar to before the pandemic, with the offline channel contribution increasing by 18.4% . Consumers prefer to return to the store to purchase products that need to be tried instantly. It is in line with [13] that consumers tend to return to buy at stores after the COVID-19 outbreak ends. Even though online channels remain a good option during this phase due to positive growth (7.7%), it is challenging to say that online shopping can totally replace offline shopping during the adapt stage.

Our work also contributes to the growth of research on purchasing decisions such as planned, unplanned, and impulsive buying and the development of the application of unsupervised learning for consumers and product segmen-

tation. Stern's buying theory is able to explain the hidden consumer segment constructed using the k-Means algorithm as the loyal consumer segment (C1), where consumers are not affected by the presence of COVID-19 in buying products; impulsive consumer segment (C2), where consumers purchase products in large quantities when COVID-19 cases are low; and compulsive consumer segment (C3), where consumers are not influenced to increase the number of products purchased when COVID-19 cases are low (Fig. 5, 6). Impulsive buying behavior was observed in cluster two (C2), where consumers purchased items spontaneously due to fear of COVID-19 infection. It is in line with [11] that personal situations such as fear can trigger impulsive buying.

Based on product category (Table 3), COVID-19 positively impacts skincare products but negatively impacts make-up products (in line with [5]). The interest of consumers in make-up has declined since government policies (e.g., PSBB, PPKM) forced people to stay at home. Conversely, consumers tend to buy more skin care products due to having more time at home (e. g., work at home, study at home, and relax at home).

In comparison, the paper [16] has similar aims and methods as the proposed study, which is to identify consumer buying behavior. However, the proposed algorithms give more compact and separated clusters of consumers ($\hat{s}=0.530$) than that in [16], which is 0.362 . This is because the proposed method calculates all possible combinations of clusters, unlike the method in [16], which explicitly determines the fixed number of clusters $k=3$ and $k=5$ at the beginning. This is also observed in [20], which only specifies three possible numbers of clusters, $k=4$, $k=5$, and $k=6$, with a maximum silhouette score of 0.460 .

Our study has recommendations for management. First, the RCA model indicates buying behavior patterns during a crisis. Management can anticipate shifts in consumer buying behavior during the pandemic and adapt accordingly. Accelerating digital business transformation is critical to boosting company revenue during the COVID-19 pandemic. This is proven by consumers switching to buying beauty products from offline to online channels.

Second, establishing and integrating a consumer relationship management (CRM) system with the enterprise system. The CRM system allows decision-makers to create better marketing strategy decisions and improve company innovation during the COVID-19 pandemic [38]. Third, strengthening the brand positioning for the face care category because this segment grew positively during the crisis. This category is slowly overtaking the make-up category, which has experienced a decline due to the impact of COVID-19 infections. This is in line with [5], where consumers tend to buy skincare products compared to make-up products during the pandemic outbreak. Fourth, utilizing social media as an effective channel to offer and sell the product. This is proven by the increase in online channel contributions of up to 9.3% (Table 2) during the restriction.

Last, the clustering results provide strong evidence of the impact of COVID-19 on the purchase of beauty products in three different consumer segments. The authors would highlight that consumers prefer to buy fragrances and hand care categories through online channels during this adaptive phase (Fig. 7). Moreover, deodorant, fragrance, hair care, and body care have become potential categories in offline channels (Fig. 9). Finally, consumers tend to buy face care

products online or offline compared to the make-up category during the remaining phase of this pandemic (Fig. 7, 9).

This work has several limitations. First, external data for verification of the RCA model only uses online transaction volume data. Second, factors related to government regulation throughout the COVID-19 outbreak were not discussed in depth in this study. Last, the interviews were limited to management and executive. In the future, the presence of offline transaction volume data can be used to complement the RCA model validation. For a more diverse analysis, all regulations related to COVID-19 are essential to be observed. Further, interviews with other parties also need to be conducted to enrich perspectives on changes in consumer buying behavior during times of crisis, such as the COVID-19 outbreak.

This paper has disadvantages due to calculating all possible clusters to be constructed. The time for computation is longer than determines the fixed number of clusters in the initial stage. Thus, to reduce this disadvantage, initializing *k* clusters can be chosen randomly at first, and the following *k* clusters were chosen proportional to the square of the distance from the closest center using k-Means++.

The difficulties of this work are that offline interviews are not allowed due to the pandemic, which can lead to misinterpretations to validate changes in consumers' buying behavior and product preferences during the crisis. In addition, because this study is related to consumer behavior, many internal and external factors besides the COVID-19 pandemic have caused the change in their purchasing behavior.

7. Conclusions

1. The periodization of the COVID-19 pandemic in Indonesia is divided into three phases: the react phase from March 2020 to December 2020, the cope phase from January 2021 to October 2021, and the adapt phase from November 2021 to June 2022. The COVID-19 outbreak, as well as its policies, impact the growth of the beauty industry. The COVID-19 cases positively impact the company's revenue during the reacting phase. A strong negative relationship between COVID-19 and company sales was observed in the coping phase. Finally, the negative impact of COVID-19 on company sales has reduced in the adapt phase.

2. The k-Means algorithm ($\hat{s}=0.530$) outperforms the Louvain algorithm ($\hat{s}=0.479$) in identifying consumer

segments. The k-Means algorithm constructed three hidden consumer segments during the pandemic: the loyal consumer segment (C1), where consumers are not affected by the presence of COVID-19 in buying products; impulsive consumers segment (C2), where consumers purchase products in large quantities when COVID-19 cases are low; and compulsive consumer segment (C3), where consumers are not influenced to increase the number of products purchased when COVID-19 cases are low. In addition, the k-Means algorithm also works better in identifying hidden product segments than Louvain's algorithm. Consumers return to buying cosmetics offline at the end of this pandemic phase, similar to before the pandemic. In this phase, consumers prefer to purchase fragrance and hand care categories in online channels. In addition, deodorant, fragrance, hair care, and body care have become potential categories in offline channels. Finally, consumers tend to buy face care products online or offline compared to the make-up category at the end of this pandemic phase.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper.

Financing

The study was performed without financial support.

Data availability

Data will be made available on reasonable request.

Acknowledgments

The authors would like to thank the Faculty of Computer Science, University of Indonesia, for their facilities, which have helped to enhance the quality of this paper.

References

1. Beauty & Personal Care - Worldwide (2022). Statista. Available at: <https://www.statista.com/outlook/cmo/beauty-personal-care/worldwide>
2. Mega farisha, Hartoyo, Safari, A. (2022). Does Covid-19 Pandemic Change the Consumer Purchase Behavior Towards Cosmetic Products? *Journal of Consumer Sciences*, 7 (1), 1–19. doi: <https://doi.org/10.29244/jcs.7.1.1-19>
3. Kim, M., Kwon, K. H. (2022). Significant paradigm of beauty ecosystem after COVID 19 pandemic in Republic of Korea. *Journal of Cosmetic Dermatology*, 21 (10), 4114–4121. doi: <https://doi.org/10.1111/jocd.15192>
4. Arora, N. et al. (2020). A global view of how consumer behavior is changing amid COVID-19. McKinsey & Company. Available at: <https://www.mckinsey.com/-/media/McKinsey/Business%20Functions/Marketing%20and%20Sales/Our%20Insights/A%20global%20view%20of%20how%20consumer%20behavior%20is%20changing%20amid%20COVID%2019/20200707/covid-19-global-consumer-sentiment-20200707.pdf>
5. Choi, Y.-H., Kim, S. E., Lee, K.-H. (2022). Changes in consumers' awareness and interest in cosmetic products during the pandemic. *Fashion and Textiles*, 9 (1). doi: <https://doi.org/10.1186/s40691-021-00271-8>
6. Jílková, P., Králová, P. (2021). Digital Consumer Behaviour and eCommerce Trends during the COVID-19 Crisis. *International Advances in Economic Research*, 27 (1), 83–85. doi: <https://doi.org/10.1007/s11294-021-09817-4>

7. Shaw, N., Eschenbrenner, B., Baier, D. (2022). Online shopping continuance after COVID-19: A comparison of Canada, Germany and the United States. *Journal of Retailing and Consumer Services*, 69, 103100. doi: <https://doi.org/10.1016/j.jretconser.2022.103100>
8. Agarwal, R., Gopinath, G., Farrar, J., Hatchett, R., Sands, P. (2022). A Global Strategy to Manage the Long-Term Risks of COVID-19. *IMF Working Papers*, 2022 (068), 1. doi: <https://doi.org/10.5089/9798400205996.001>
9. Alsaydia, O. M., Saadallah, N. R., Malallah, F. L., AL-Adwany, M. A. S. (2021). Limiting COVID-19 infection by automatic remote face mask monitoring and detection using deep learning with IoT. *Eastern-European Journal of Enterprise Technologies*, 5 (2 (113)), 29–36. doi: <https://doi.org/10.15587/1729-4061.2021.238359>
10. Wibowo, A., Chen, S.-C., Wiangin, U., Ma, Y., Ruangkanjanases, A. (2020). Customer Behavior as an Outcome of Social Media Marketing: The Role of Social Media Marketing Activity and Customer Experience. *Sustainability*, 13 (1), 189. doi: <https://doi.org/10.3390/su13010189>
11. Tanveer, T., Kazmi, S. Q., Rahman, M. U. (2022). Determinants of Impulsive Buying Behavior: An Empirical Analysis of Consumers' Purchase Intentions for Offline Beauty Products. *Nurture*, 16 (2), 75–89. doi: <https://doi.org/10.55951/nurture.v16i2.129>
12. Jo, H., Shin, E., Kim, H. (2020). Changes in Consumer Behaviour in the Post-COVID-19 Era in Seoul, South Korea. *Sustainability*, 13 (1), 136. doi: <https://doi.org/10.3390/su13010136>
13. Adibfar, A., Gulhare, S., Srinivasan, S., Costin, A. (2022). Analysis and modeling of changes in online shopping behavior due to Covid-19 pandemic: A Florida case study. *Transport Policy*, 126, 162–176. doi: <https://doi.org/10.1016/j.tranpol.2022.07.003>
14. Kirk, C. P., Rifkin, L. S. (2020). I'll trade you diamonds for toilet paper: Consumer reacting, coping and adapting behaviors in the COVID-19 pandemic. *Journal of Business Research*, 117, 124–131. doi: <https://doi.org/10.1016/j.jbusres.2020.05.028>
15. Guthrie, C., Fosso-Wamba, S., Arnaud, J. B. (2021). Online consumer resilience during a pandemic: An exploratory study of e-commerce behavior before, during and after a COVID-19 lockdown. *Journal of Retailing and Consumer Services*, 61, 102570. doi: <https://doi.org/10.1016/j.jretconser.2021.102570>
16. Anitha, P., Patil, M. M. (2022). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University - Computer and Information Sciences*, 34 (5), 1785–1792. doi: <https://doi.org/10.1016/j.jksuci.2019.12.011>
17. Tabianan, K., Velu, S., Ravi, V. (2022). K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data. *Sustainability*, 14 (12), 7243. doi: <https://doi.org/10.3390/su14127243>
18. Zhang, L., Priestley, J., DeMaio, J., Ni, S., Tian, X. (2021). Measuring Customer Similarity and Identifying Cross-Selling Products by Community Detection. *Big Data*, 9 (2), 132–143. doi: <https://doi.org/10.1089/big.2020.0044>
19. Taghikhah, F., Voinov, A., Shukla, N., Filatova, T. (2021). Shifts in consumer behavior towards organic products: Theory-driven data analytics. *Journal of Retailing and Consumer Services*, 61, 102516. doi: <https://doi.org/10.1016/j.jretconser.2021.102516>
20. Abbasimehr, H., Bahrini, A. (2022). An analytical framework based on the recency, frequency, and monetary model and time series clustering techniques for dynamic segmentation. *Expert Systems with Applications*, 192, 116373. doi: <https://doi.org/10.1016/j.eswa.2021.116373>
21. Li, Z., Gao, X., Lu, D. (2021). Correlation analysis and statistical assessment of early hydration characteristics and compressive strength for multi-composite cement paste. *Construction and Building Materials*, 310, 125260. doi: <https://doi.org/10.1016/j.conbuildmat.2021.125260>
22. Miloudi, S., Wang, Y., Ding, W. (2021). A Gradient-Based Clustering for Multi-Database Mining. *IEEE Access*, 9, 11144–11172. doi: <https://doi.org/10.1109/access.2021.3050404>
23. Sinaga, K. P., Yang, M.-S. (2020). Unsupervised K-Means Clustering Algorithm. *IEEE Access*, 8, 80716–80727. doi: <https://doi.org/10.1109/access.2020.2988796>
24. Arunachalam, D., Kumar, N. (2018). Benefit-based consumer segmentation and performance evaluation of clustering approaches: An evidence of data-driven decision-making. *Expert Systems with Applications*, 111, 11–34. doi: <https://doi.org/10.1016/j.eswa.2018.03.007>
25. Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhajja, B., Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, 622, 178–210. doi: <https://doi.org/10.1016/j.ins.2022.11.139>
26. Batool, F., Hennig, C. (2021). Clustering with the Average Silhouette Width. *Computational Statistics & Data Analysis*, 158, 107190. doi: <https://doi.org/10.1016/j.csda.2021.107190>
27. Archived: WHO Timeline - COVID-19 (2020). WHO. Available at: <https://www.who.int/news/item/27-04-2020-who-timeline--covid-19>
28. Clear and free: Wuhan evacuees head home from Natuna quarantine (2020). *The Jakarta Post*. Available at: <https://www.thejakartapost.com/news/2020/02/16/clear-and-free-wuhan-evacuees-head-home-from-natuna-quarantine.html>
29. BREAKING: Jokowi announces Indonesia's first two confirmed COVID-19 cases (2020). *The Jakarta Post*. Available at: <https://www.thejakartapost.com/news/2020/03/02/breaking-jokowi-announces-indonesias-first-two-confirmed-covid-19-cases.html>
30. Presiden Tetapkan COVID-19 Sebagai Bencana Nasional (2020). BNPB. Available at: [https://bnpb.go.id/berita/presiden-tetapkan-covid19-sebagai-bencana-nasional#:~:text=JAKARTA%20%2D%20Presiden%20Joko%20Widodo%20secara,%2D19\)%20Sebagai%20Bencana%20Nasional](https://bnpb.go.id/berita/presiden-tetapkan-covid19-sebagai-bencana-nasional#:~:text=JAKARTA%20%2D%20Presiden%20Joko%20Widodo%20secara,%2D19)%20Sebagai%20Bencana%20Nasional)

