

## UNRAVELING OF MEN'S FRAGRANCE PREFERENCES ON ONLINE MARKETPLACES: A MACHINE LEARNING STUDY USING DBSCAN CLUSTERING AND LINEAR REGRESSION

Zakiyyan Zain Alkaf<sup>\*1</sup>, Muhammad Ihsan Fawzi<sup>2</sup>, Murti Wisnu Ragil Sastyawan<sup>3</sup>, Radita Dwi Putera<sup>4</sup>

<sup>1,3,4</sup>Industrial Engineering, Faculty of Engineering, Universitas Jenderal Soedirman, Indonesia

<sup>2</sup>Informatics Engineering, Faculty of Engineering, Universitas Jenderal Soedirman, Indonesia

Email: <sup>1</sup>[zakiyyan.alkaf@unsoed.ac.id](mailto:zakiyyan.alkaf@unsoed.ac.id), <sup>2</sup>[ihsfwz@incomso.com](mailto:ihsfwz@incomso.com), <sup>3</sup>[murti.sastyawan@unsoed.ac.id](mailto:murti.sastyawan@unsoed.ac.id),

<sup>4</sup>[radita.putera@unsoed.ac.id](mailto:radita.putera@unsoed.ac.id)

(Article received: December 07, 2024; Revision: January 08, 2025; published: January 17, 2025)

### Abstract

*The perfume industry is undergoing significant growth, driving the need to understand consumer preferences, particularly in men's fragrances, to optimize business strategies. This study aims to analyze and uncover men's fragrance preferences, using machine learning techniques. A dataset of approximately 1,000 men's perfume records from Kaggle was utilized, where systematic methodologies were employed. Data preprocessing involved handling missing values, removing duplicates, standardizing categorical entries, and performing feature engineering by extracting geographic information from item locations. Exploratory Data Analysis (EDA) was conducted to uncover data distribution. Clustering analysis using DBSCAN revealed consumer segments. Additionally, regression analysis was used to predict sales based on price and location, employing a linear regression model evaluated with metrics like Mean Squared Error (MSE). The findings indicate that price exhibits a complex relationship with sales; while affordable products drive higher sales volumes, premium-priced items cater to a niche yet impactful market segment. Geographic location plays a pivotal role in sales patterns. Clustering analysis reveals two distinct consumer segments: one driven by price sensitivity and another oriented towards premium preferences, influenced by regional factors. Regression analysis demonstrated a negative correlation between price and sales volume, with a coefficient of -1.81, while availability positively influenced sales with a coefficient of 8.36. Despite a moderate model fit ( $R^2 = 0.17$ ), the analysis highlights key market dynamics. These insights emphasize the importance of leveraging data-driven strategies to develop targeted marketing campaigns, optimize inventory management and refine market segmentation.*

**Keywords:** Clustering Analysis, Customer Behaviour, Customer Preferences, Perfume, Regression Analysis.

## 1. INTRODUCTION

The men's fragrance industry has become a significant segment of the global market, characterized by rising consumer interest and unpredictable shifts in purchasing preferences [1]. As the sector continues to expand rapidly, understanding consumer behavior has become crucial for brands seeking to cater to the unique demands of their customers [2]. Numerous factors—cultural, social, psychological, and more—shape consumers' scent preferences and purchasing decisions [3]–[5]. However, traditional research methods often fall short in identifying the underlying drivers behind these choices.

Advanced machine learning techniques have access to vast amounts of data, enabling them to uncover unique consumer preferences in men's fragrances [6]. Specifically, clustering methods allow professionals to categorize scents by similarity, identify target audiences, and classify fragrance types. For example, these methods help businesses determine which demographic is most likely to favor

a particular scent or fragrance component. Similarly, predictive models are valuable for anticipating consumer reactions or preferences. These techniques and their practical applications provide valuable insights for professionals in fields such as marketing [7].

Machine learning clustering enables businesses to categorize different consumers based on their fragrance preferences, allowing them to tailor their strategies accordingly [8]. For instance, consumers who prefer fresh, citrusy scents can be grouped separately from those who favor warmer notes like woody or oriental fragrances. Such segmentation provides valuable insights into consumer behavior and helps companies identify specific product collections to target each group [9]–[11]. Understanding these preferences also aids in developing tailored fragrance lines for each consumer segment [12]. By gaining a deeper understanding of consumer behavior within each group, businesses can enhance customer satisfaction and ensure that new products, such as perfumes, are aligned with the desires of different target markets [13], [14].

Machine learning regression analysis can further refine these insights by predicting future trends based on historical data, allowing brands to stay ahead of the curve and innovate effectively [15]. This proactive approach not only fosters brand loyalty but also encourages a more personalized shopping experience, ultimately driving sales and enhancing market presence [16]. Additionally, integrating customer feedback loops into the development process can provide real-time insights, enabling brands to adapt quickly to changing preferences and emerging trends [17].

The aim of this study is to analyze and uncover men's fragrance preferences by utilizing machine learning methodologies. By utilizing the insights gained from the research, companies can ensure that their products align with consumer desires and that their marketing efforts effectively target the right audiences. The study employs techniques such as clustering and regression, which extract valuable information from data on scent preferences. This approach not only enhances product development but also fosters a deeper connection between brands and their consumers, ultimately leading to increased loyalty and repeat purchases.

## 2. METHOD

Data science methodologies are essential for extracting meaningful insights from complex datasets. The adoption of structured approaches is increasing as more organizations rely on data-driven decision-making [18]. By following systematic processes, data scientists can provide robust analyses that inform business strategies. This article explores various techniques utilized in data science projects, emphasizing the critical steps that lead to actionable findings. The stages carried out in this research according to figure 1 include: Data Preprocessing, Exploratory Data Analysis (EDA), Clustering Analysis, Regression Analysis, Data Visualization.

### 2.1. Data Collection

The data we use we take from open source data: kaggle.com, specifically in [Perfume E-Commerce Dataset 2024](#), which has a usability level of 10 which means it has 100% score of completeness, credibility, and compatibility. In addition, the dataset is recommended by several users for research materials. Then this data has an Open Data Commons Attribution License (ODC-By) v1.0 license. Then the number of data rows that existed at the beginning before preprocessing was approximately 1000 data.

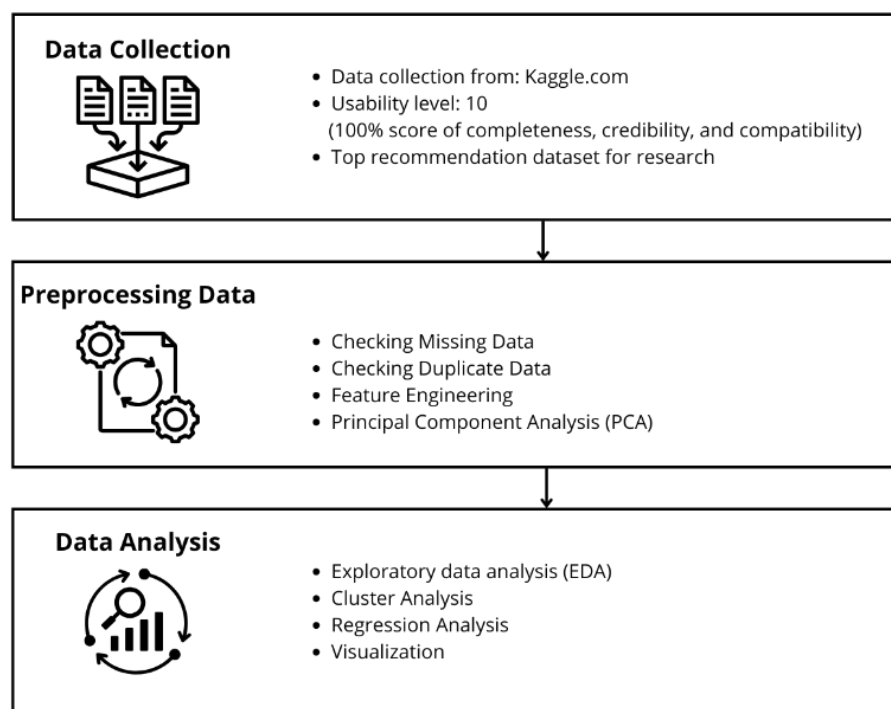


Figure 1. Method

### 2.2. Preprocessing Data

The cleaning step is key because incorrect or strange info is thrown away. Missing values were identified using the `.isnull().sum()` method, followed by the removal of rows containing missing values through `.dropna(inplace=True)`, which helped maintain the integrity of the dataset. For duplicates,

the `.duplicated()` function was employed to detect and count duplicate entries, which were subsequently reviewed and removed. Inconsistencies in the 'type' column were addressed by removing rows with non-standard entries such as 'Does Not Apply' and other irregular descriptions. After the cleanup, changes in data and size followed to allow us to change the info, so everything was in agreement. A step-by-step plan

is suitable for achieving valid and significant results [19]. Creating new features is key between feature selection and change, which improves the skill and correctness of the model. Common ways used to enhance datasets include alignment of properties, turning things into numerical shapes, and ways of cutting measures such as Principal Component Analysis (PCA). This allows experts to find and make the most obvious points for a given dataset, improving the reliability of the findings [20]. In this study, this feature engineering stage is modifications were made to the 'itemLocation' column, where commas were added to ensure uniformity across entries. The city names were then extracted from this column and stored in a new 'kota' column. Rows containing empty strings, None, or NaN in the 'kota' column were also removed to finalize the data cleaning process. Thus, solid results are ensured because the invention is not warped or deceptive.

### 2.3. Data Analysis

Exploratory data analysis (EDA), which aims to summarize the main characteristics of information, often uses the image method. By definition, EDA is all about finding hidden trends, spotting oddities, and testing guesses. It is a plan that provides understanding and prepares info for formal review [21]. In this study, the researcher sought to identify outliers using z-scores, then looked for the distribution of data that had been preprocessed previously using summary statistics (mean, median and standard deviation).

Clustering plays a critical role in data analysis, as it helps to group similar items, making the interpretation of patterns more straightforward. Various clustering techniques exist, and DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is one of the effective methods for identifying clusters [22]. Unlike K-means, which divides items based on distance and averages, DBSCAN is particularly useful for detecting clusters of varying shapes and sizes by focusing on the density of points in the dataset. In this study, DBSCAN is employed to identify clusters from the data, particularly focusing on sales data, using three clusters to uncover similarities.

Regression analysis is used to predict the number of sales based on price and location. The method or approach used is linear regression. Furthermore, a linear regression algorithm is applied using the scikit-learn library in Python to build a predictive model. The trained model uses 80% of the total data as training data and 20% as test data to evaluate the model's performance. The evaluation results are reviewed based on metrics such as Mean Squared Error (MSE) to determine the accuracy and predictive capabilities of the model.

Data visualization is key between sharing elusive things successfully. This allows the crowd to

get useful ideas. In this study, data visualization in the form of histograms, box plots, scatter plots, and bar charts is used. It is hoped that this data visualization can provide easier understanding to readers.

## 3. RESULT

### 3.1. Average price distribution

Figure 2 illustrates the average price distribution across various types of men's fragrances. The data reveals significant price disparities among different product categories. Eau De Parfum emerges as the most premium offering, with an average price of approximately 260 units, followed closely by Gift Sets at about 200 units. In contrast, Body Oil represents the most economical option, with an average price just above 10 units. Other notable categories include Fragrance and Extrait De Parfum, both commanding average prices exceeding 150 units. Traditional men's fragrance types such as Eau De Toilette, Cologne Intense, and Aftershave occupy the mid-range price segment, varying between 40 and 150 units. This price stratification clearly delineates a hierarchical structure within the men's fragrance market, reflecting diverse product formulations and target consumer segments.

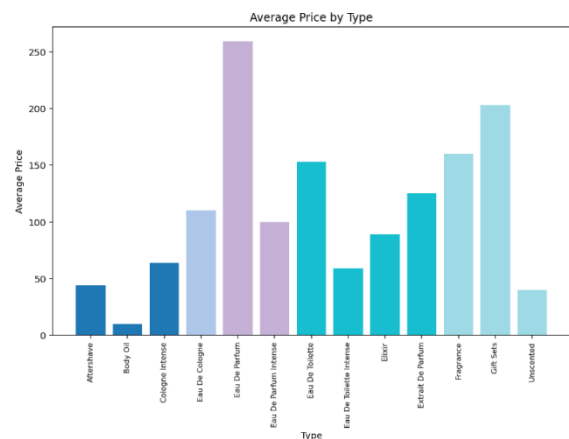


Figure 2. Average Price by Type

### 3.2. Distribution of Price, Available and Sold

The histogram in Figure 3 illustrates the distribution of prices, available and sold for men's perfumes. The majority of products were priced in the lower range, suggesting a market bias towards more affordable options. However, a notable number of high-priced perfumes were also observed, indicating the presence of a premium segment. The availability of products, as shown in Figure 3, was generally sufficient, with a normal distribution pattern. Nevertheless, some products had limited or no stock, potentially due to high demand or recent product launches. The sales distribution in Figure 3 exhibited a wide range, from products with very low sales to those with exceptionally high sales. This indicates significant variations in product popularity.

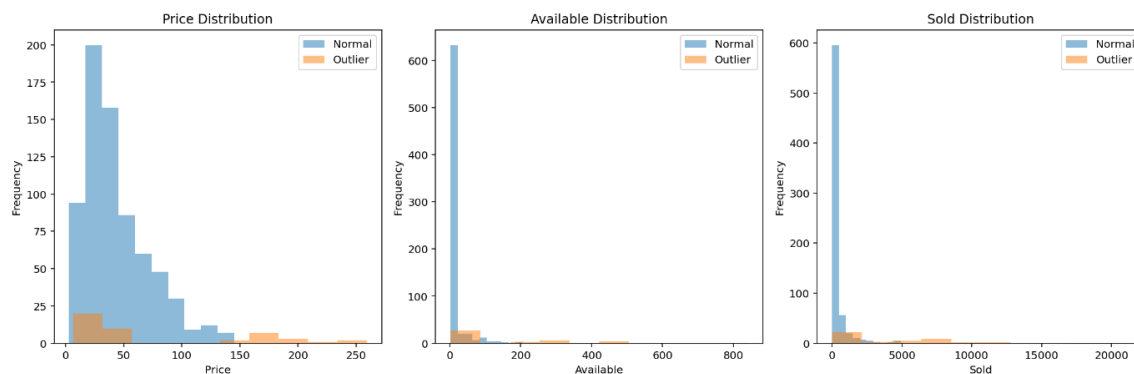


Figure 3. Distribution of Price, Available and Sold

### 3.3. Cluster Analysis

Cluster analysis revealed two distinct segments in the men's perfume market. Cluster 1, comprising the majority of data points, was characterized by a more affordable price range, with an average price of \$43.87. The dominant brands in this cluster were

Giorgio Armani, and the most common product type was Eau de Toilette. In contrast, Cluster 2 was a smaller segment with a significantly higher average price of \$28.83. Versace was the most prominent brand in this cluster, and the products were primarily sold in Hackensack.

Table 1. Cluster Analysis

No	Cluster	Amount of Data	Average by Price	Average by Sold	Dominant Brand	Dominant Perfume Types	Dominant Location
1	Cluster 1	697	43.867	249.278	Giorgio Armani	Eau De Toilette	Dallas
2	Cluster 2	7	28.837	4555.857	Versace	Eau De Toilette	Hackensack

The analysis of the Silhouette score across various epsilon values revealed that dividing the data into two clusters yielded the most optimal results. The Silhouette score in figure 4, which quantifies the quality of clustering, demonstrated a general upward trend as the epsilon value increased, reaching a maximum of approximately 0.89. This indicates that the two-cluster solution provides a high degree of intra-cluster cohesion and inter-cluster separation. Notably, attempts to partition the data into more than two clusters resulted in several sparsely populated clusters beyond the primary cluster, suggesting that a two-cluster approach most effectively captures the underlying structure of the data.

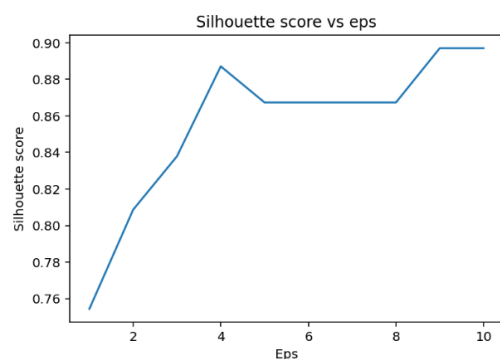


Figure 4. Silhouette score

### 3.4. Correlation between Price and Sold

The scatter plot in figure 5 illustrating the correlation between price and sales volume of men's

perfumes on eBay reveals a complex relationship. The majority of data points, classified as "Data Normal," are concentrated in the lower price range (below \$100) with sales volumes primarily under 5,000 units. A notable cluster of higher-selling items is observed in the \$20-\$50 price range. The plot also identifies several "Data Outlier" points, particularly in the lower price brackets, showing exceptionally high sales volumes (up to 20,000 units). Interestingly, as the price increases beyond \$150, the sales volume generally decreases, with few outliers present in this higher price range. This distribution suggests a non-linear relationship between price and sales, with peak sales occurring in the mid-price range.

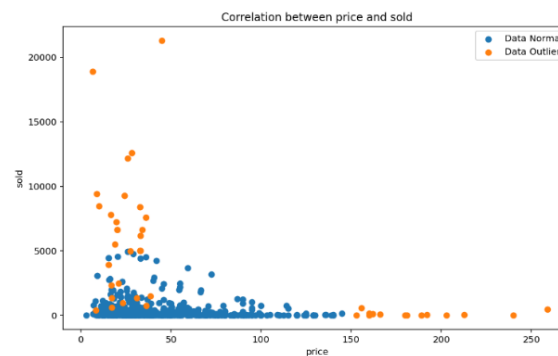


Figure 5. Correlation between Price and Sold

### 3.5. Regression Analysis

The regression analysis in table 2 reveals a Mean Squared Error (MSE) of 451721.07 and an R-squared value of 0.17, indicating that the model

accounts for only 17% of the variance in the number of products sold (dependent variable). The regression coefficients show a positive relationship between the variable "available" and "sold" with a coefficient of 8.36, suggesting that an increase in the availability of men's cologne corresponds to an increase in the number of units sold. Conversely, there is a negative relationship between "price" and "sold," as indicated by the regression coefficient of -1.81. This suggests that higher prices are associated with a reduction in sales. The regression intercept is 248.025, representing the predicted number of units sold when both availability and price are held constant.

Table 2. Regression Analysis

No	Mean Squared Error	R-Squared	Regression Coefficient	Regression Intercept
1	451721.07	0.17	[8.36095472 -1.81135758]	248.025

#### 4. DISCUSSION

Research has shown that male fragrance preferences can vary significantly by age group, income level, and cultural background [23]. Younger consumers may gravitate towards trendier, lower-priced products, while older consumers with higher disposable income may be more inclined to purchase premium options like Eau De Parfum. Cultural influences can shape scent preferences, with certain fragrances resonating more in specific regions due to historical or societal trends [24]. This implies that pricing strategies should not only reflect product quality but also align with the cultural and socioeconomic factors that drive consumer behavior in different markets.

Studies suggest that price-sensitive consumers are more likely to respond to discounts and promotions, which can drive higher sales volume for lower-priced products [25]. Conversely, premium products often rely on brand loyalty and perceived value, which means price changes may have less impact on sales volume. This dichotomy highlights the importance of employing differentiated pricing strategies based on product type and target market. Moreover, the variation in product availability may suggest that certain items are subject to supply chain challenges, such as seasonal demand fluctuations or production constraints [26]. Addressing these supply chain issues through better inventory management and demand forecasting could help optimize sales and avoid stockouts.

Based on research by Rogrigues [6], consumer preference clusters regarding perfume are assessed based on its aroma, which explains that the average rating of perfumes containing raspberry and saffron aromas; the heart aromas are jasmine and frankincense; the base aromas are leather, amber, and wood are more preferred. According to Kapadia's research, women prefer floral, fruity, and citrus scents, while men also prefer floral scents, followed

by woody, fruity, and citrus. According to Kholibrina and Aswandi [27], respondents prioritize three factors, namely aroma, grade, and price. With aroma having the greatest influence and followed by its grade. Which means that the cluster results in this study are appropriate, where grade/quality has a stronger influence compared to its price which is very difficult to analyze.

Pentus et al. [28] discussed from the advertising side, comparing perfumes that advertised sexual elements and non-sexual elements. With the same speed and time of advertising, the results were that advertisements containing sexual elements did not outperform advertisements that did not contain sexual elements in terms of viewing time. Research by Kim [29] presents the results of their survey, that after the COVID-19 pandemic, respondents preferred light perfumes, with disposable packaging, and perfumes were preferred over colognes.

According to consumer behavior theories, factors such as brand reputation, product reviews, and marketing efforts can significantly influence sales outcomes [30], [31]. For instance, a highly rated cologne may outperform competitors with lower prices due to its perceived value in terms of quality and brand prestige. Dwijayanti and Windasari [32] discussed the relationship between influencers and consumer purchasing decisions, the results showed that if consumers feel that influencers who have a high level of trust, high expertise and high brand image will help consumers' decisions to buy perfume. In another study by Isamudin and Islam [33], the main factor influencing consumer trust is affect-based trust which consists of seller reputation, customer feedback, and recommendations from experts.

From the research that discusses the influence of perfume packaging on consumer purchasing behavior, the results of the study are that packaging information, color, design, background images, and innovation have an effect on consumer purchasing behavior [34], [35]. However, the packaging material has no effect. This can be an additional insight if in the future the perfume industry will produce types of aromas that are in high demand.

As the regression analysis fails to account for these qualitative factors, future research could integrate sentiment analysis of reviews or measure the impact of visual presentation on sales performance. This is in line with research by Ardelet et al. that it is difficult to predict market trends, so it is necessary to highlight newly released products [1]. Moreover, the inclusion of time-based variables, such as the age of the product listing or seasonal demand patterns, could improve the predictive power of the regression model.

#### 5. CONCLUSION

This study aimed to explore consumer behavior and preferences in the men's fragrance market, leveraging machine learning techniques such as

clustering and regression analysis. By analyzing a dataset of men's perfumes, the research uncovered key insights into the dynamics of pricing, availability, and sales.

The findings revealed a distinct segmentation in consumer preferences: one cluster driven by affordability and another by premium product offerings influenced by regional factors. Regression analysis demonstrated that price negatively impacts sales volume, while availability positively correlates with sales. These insights underscore the importance of pricing strategies that cater to both budget-conscious consumers and niche premium markets. Additionally, the results highlight the critical role of geographic and cultural factors in shaping consumer behavior, suggesting that businesses should tailor their marketing and inventory strategies accordingly.

The implications of this study are manifold. First, it provides actionable insights for businesses to optimize pricing and inventory management, ensuring alignment with consumer demand patterns. Second, the identification of consumer segments can assist marketers in developing targeted campaigns to enhance engagement and customer loyalty. Third, the study offers valuable guidance for businesses seeking to improve their operational efficiency and maximize sales potential in a competitive market.

Future research should aim to integrate qualitative factors such as consumer sentiment, product reviews, and advertising impact to enhance the predictive accuracy of sales models. Furthermore, incorporating time-based variables and exploring the influence of seasonal trends could provide deeper insights into market dynamics. This approach will enable a more comprehensive understanding of consumer preferences, facilitating the development of more effective business strategies in the men's fragrance industry.

## REFERENCES

- [1] C. Ardelet, B. Slavich, and G. de Kerviler, "Self-referencing narratives to predict consumers' preferences in the luxury industry: A longitudinal study," *J. Bus. Res.*, vol. 68, no. 9, pp. 2037–2044, 2015, doi: <https://doi.org/10.1016/j.jbusres.2015.02.017>
- [2] R. N. Joshi and D. Y. C. Joshi, "Construction and Validation of Indian Consumer Ethnocentrism Scale: The Ice-Scale," *Gadjah Mada Int. J. Bus.*, vol. 23, no. 2, p. 173, May 2021, doi: [10.22146/gamaijb.58039](https://doi.org/10.22146/gamaijb.58039).
- [3] F. Begnaud, "Mass Spectrometry: An Essential Tool for the Flavor and Fragrance Industry," *Chimia (Aarau)*, vol. 76, no. 1–2, p. 68, Feb. 2022, doi: [10.2533/chimia.2022.68](https://doi.org/10.2533/chimia.2022.68).
- [4] K. Ponsree and P. Naruetharadhol, "Unveiling the determinants of alternative payment adoption: exploring the factors shaping generation Z's intentions in Thailand," *Int. Entrep. Manag. J.*, vol. 21, no. 1, p. 45, Dec. 2025, doi: [10.1007/s11365-024-01057-2](https://doi.org/10.1007/s11365-024-01057-2).
- [5] K. Adalarasu, K. Ghousiya Begum, M. Vishnu Priyan, C. Devendranath, and G. V. Sriram, "Neuro-signaling techniques in advertisement endorsements: Unveiling consumer responses and behavioral trends," *J. Retail. Consum. Serv.*, vol. 84, p. 104175, May 2025, doi: [10.1016/j.jretconser.2024.104175](https://doi.org/10.1016/j.jretconser.2024.104175).
- [6] B. C. L. Rodrigues, V. V. Santana, L. P. Queiroz, C. M. Rebello, and I. B. R. Nogueira, "Harnessing graph neural networks to craft fragrances based on consumer feedback," *Comput. Chem. Eng.*, vol. 185, p. 108674, Jun. 2024, doi: [10.1016/j.compchemeng.2024.108674](https://doi.org/10.1016/j.compchemeng.2024.108674).
- [7] S. Bogacki, P. Rymarczyk, T. Smutek, M. Rutkowski, and A. Chmielowska-Marmucka, "Advanced methods for target audience identification: enhancing marketing strategies through machine learning and data analytics," *J. Mod. Sci.*, vol. 57, no. 3, pp. 417–435, Aug. 2024, doi: [10.13166/jms/191180](https://doi.org/10.13166/jms/191180).
- [8] Y. Cai, Q. Shu, Z. Zhou, and H. Meng, "Incomplete multi-view clustering based on information fusion with self-supervised learning," *Inf. Fusion*, vol. 117, p. 102849, May 2025, doi: [10.1016/j.inffus.2024.102849](https://doi.org/10.1016/j.inffus.2024.102849).
- [9] S. Sobantu and O. E. Isafiade, "Cascaded RFM-Based Fuzzy Clustering Model for Dynamic Customer Segmentation in the Retail Sector," 2025, pp. 53–73. doi: [10.1007/978-3-031-78255-8\\_4](https://doi.org/10.1007/978-3-031-78255-8_4).
- [10] N. M. Nhat, "Applied Density-Based Clustering Techniques for Classifying High-Risk Customers: A Case Study of Commercial Banks in Vietnam," *J. Appl. Data Sci.*, vol. 5, no. 4, pp. 1639–1653, Dec. 2024, doi: [10.47738/jads.v5i4.344](https://doi.org/10.47738/jads.v5i4.344).
- [11] P. Castricone and L. Long, "Deriving consumer insights with segmentation for identity and consent," *Appl. Mark. Anal. Peer-Reviewed J.*, vol. 10, no. 3, p. 245, Dec. 2024, doi: [10.69554/JZEK9481](https://doi.org/10.69554/JZEK9481).
- [12] K. P. Suhaas, B. G. Deepa, D. Shashank, and M. Narender, "Millets Industry Dynamics: Leveraging Sales Projection and Customer Segmentation," *SN Comput. Sci.*, vol. 5, no. 8, p. 1063, Nov. 2024, doi: [10.1007/s42979-024-03437-5](https://doi.org/10.1007/s42979-024-03437-5).
- [13] A. R. Oskooei and T. E. Adak, "B2B Customer Engagement Customer Behaviour Forecast Application," *Orclever Proc. Res.*



- Dev.*, vol. 3, no. 1, pp. 188–196, Dec. 2023, doi: 10.56038/oprd.v3i1.323.
- [14] I. Maulana, A. M. Siregar, S. A. P. Lestari, and S. Faisal, "OPTIMAL STUDY OF REAL-ESTATE PRICE PREDICTION MODELS USING MACHINE LEARNING," *J. Tek. Inform.*, vol. 5, no. 4, pp. 1149–1164, 2024, doi: <https://doi.org/10.52436/1.jutif.2024.5.4.2565>.
- [15] C. W. Cheah, M. N. A. Karim, Y. S. Lee, and M. D. i Md Ghazal, "The Accuracy and Error of Ground Penetrating Radar System with Machine Learning Support Vector Regression Technique," *J. Adv. Res. Appl. Sci. Eng. Technol.*, vol. 50, no. 1, pp. 191–202, Aug. 2024, doi: 10.37934/araset.50.1.191202.
- [16] M. Wan, M. A. Alshahrani, N. M. Aloraini, A. A. Alkhathami, and H. Alqahtani, "On predictive modeling of the twitter-based sales data using a new probabilistic model and machine learning methods," *Alexandria Eng. J.*, vol. 113, pp. 661–671, Feb. 2025, doi: 10.1016/j.aej.2024.11.041.
- [17] M. K. S. Rajalakshmi, A. Sriraman, and M. Krishnamoorthy, "Regression analysis for software development and management using machine learning." Google Patents, Jun. 15, 2021.
- [18] R. Cerezo, J.-A. Lara, R. Azevedo, and C. Romero, "Reviewing the differences between learning analytics and educational data mining: Towards educational data science," *Comput. Human Behav.*, vol. 154, p. 108155, 2024, doi: <https://doi.org/10.1016/j.chb.2024.108155>.
- [19] C. Fan, M. Chen, X. Wang, J. Wang, and B. Huang, "A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data," *Front. Energy Res.*, vol. 9, p. 652801, Mar. 2021, doi: 10.3389/fenrg.2021.652801.
- [20] S. Galli, *Python feature engineering cookbook*. Packt Publishing Ltd, 2024.
- [21] T. Milo and A. Somech, "Automating Exploratory Data Analysis via Machine Learning: An Overview," in *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, New York, NY, USA: ACM, Jun. 2020, pp. 2617–2622. doi: 10.1145/3318464.3383126.
- [22] J. Kim, H. Lee, and Y. M. Ko, "Constrained Density-Based Spatial Clustering of Applications with Noise (DBSCAN) using hyperparameter optimization," *Knowledge-Based Syst.*, vol. 303, p. 112436, 2024, doi: <https://doi.org/10.1016/j.knosys.2024.112436>.
- [23] C. Chen *et al.*, "Aroma-sensory properties of Gouda cheeses based on young Chinese consumers' preferences," *J. Dairy Sci.*, vol. 107, no. 7, pp. 4333–4343, 2024, doi: <https://doi.org/10.3168/jds.2023-23525>.
- [24] S. G. H. Meyerding and B. H. Trinh, "A cross-cultural study: How product attributes and cultural values influence chocolate preferences," *J. Agric. Food Res.*, vol. 19, p. 101579, Mar. 2025, doi: 10.1016/j.jafr.2024.101579.
- [25] J. van der Plas, M. G. Dekimpe, and I. Geyskens, "Mind the gap: National brands' sensitivity to price and line-length differentials at hard discounters versus conventional retailers," *J. Retail.*, vol. 100, no. 2, pp. 199–216, 2024, doi: <https://doi.org/10.1016/j.jretai.2024.05.005>.
- [26] R. Anguzu and F. Aila, "Sources of Supply Chain Volatility: A Literature Review," *Int. J. Supply Chain Manag.*, vol. 9, no. 2, pp. 20–36, Feb. 2024, doi: 10.47604/ijscm.2370.
- [27] C. R. Kholibrina and A. Aswandi, "The consumer preferences for new styra based perfume products using a conjoint analysis approach," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 935, no. 1, p. 012016, Sep. 2020, doi: 10.1088/1757-899X/935/1/012016.
- [28] K. Pentus, M. Ruusu, A. Kuusik, L. Dorokhova, and K. Ploom, "How Sexualised Images in Advertisements Influence the Attention and Preference of Consumers with a Modern View," *Organ. Mark. Emerg. Econ.*, vol. 14, no. 2(28), pp. 366–385, Jul. 2023, doi: 10.15388/omee.2023.14.97.
- [29] C. G. Kim, "The Fragrance of Cyber World 2080: A Perfume Forecast," in *Proceedings of the 24th International RAIS Conference on Social Sciences and Humanities*, Scientia Moralitas Research Institute, 2021, pp. 125–132.
- [30] C.-L. Hung *et al.*, "Enhancing healthcare services and brand engagement through social media marketing: Integration of Kotler's 5A framework with IDEA process," *Inf. Process. Manag.*, vol. 60, no. 4, p. 103379, 2023, doi: <https://doi.org/10.1016/j.ipm.2023.103379>.
- [31] L. Nikitina, "Fragrance lexicon for analysis of consumer-generated perfume reviews in Russian and English," *MethodsX*, vol. 12, p. 102627, 2024, doi: <https://doi.org/10.1016/j.mex.2024.102627>.
- [32] Y. Dwijayanty and N. A. W. S.A., MBA., Ph.D., "Perfume Influencers Impact on Customers' Purchase Decision of HMNS

- Perfumery,” *Int. J. Curr. Sci. Res. Rev.*, vol. 05, no. 12, Dec. 2022, doi: 10.47191/ijcsrr/V5-i12-31.
- [33] N. F. Isamudin and I. B. Islam, “Factors Influencing Young Consumer Trust in Online Purchase Decision of Luxury Perfume in Malaysia: A Preliminary Qualitative Study,” *Int. J. Econ. Manag. Account.*, vol. 31, no. 2, pp. 299–323, Dec. 2023, doi: 10.31436/ijema.v31i2.1062.
- [34] H. Hussien, M. A. Saleh, R. Al Saed, and N. Sweidan, “The Impact of Perfume Packaging on Consumer Buying Behavior of Jordanian Female,” *Int. J. Bus. Excell.*, vol. 1, no. 1, p. 1, 2020, doi: 10.1504/IJBEX.2020.10030847.
- [35] L. Udayanga, N. Subashini, M. Udugama, P. Silva, and T. Ranathunge, “Knowledge, perceptions, and consumption behaviour of cosmetics among undergraduates of Sri Lanka: a descriptive cross-sectional study,” *Front. Public Heal.*, vol. 11, Jan. 2024, doi: 10.3389/fpubh.2023.1184398.